# Benchmark Discrepancies and Mutual Fund Performance Evaluation


K.J. Martijn Cremers
mcremers@nd.edu

Mendoza College of Business
University of Notre Dame
Notre Dame, IN 46556

Jon A. Fulkerson
jfulkerson1@udayton.edu

School of Business Administration
University of Dayton
Dayton, OH 45469

Timothy B. Riley*
tbriley@uark.edu

Sam M. Walton College of Business
University of Arkansas
Fayetteville, AR 72701



First Draft: October 2017
This Draft: September 2018



*Abstract*

We introduce a new holdings-based procedure to identify the benchmark discrepancies of mutual funds, which we define as a benchmark other than the prospectus benchmark best matching a fund's investment strategy. Funds with a benchmark discrepancy tend to be riskier than their prospectus benchmarks indicate. As a result, those funds on average outperform their prospectus benchmark—before risk-adjusting—despite generally underperforming the benchmark that best matches their holdings. High active share funds outperform more if there is no benchmark discrepancy, suggesting that managers with more skill are less likely to have a benchmark discrepancy.



* Contact author. Martijn Cremers is currently an independent director at Ariel Investments and a consultant to Touchstone Investments and State Street Associates. All of the authors have performed work on behalf of the Investment Adviser Association. We thank Brad Jordan, Berk Sensoy, seminar participants at the University of Dayton, and members of the CFA Society Arkansas for helpful comments and suggestions. All errors are our own.

## 1. Introduction

The evaluation of the performance of an investment product, such as an actively managed mutual fund, generally involves comparing the performance of that product with some benchmark. That benchmark could be a passive benchmark index that follows the same style as the product's portfolio (e.g., the S&P 500); be based on the portfolio's exposure to different systematic factors (e.g., Fama and French, 1993); or be based on the characteristics of the portfolio's holdings (e.g., Daniel, Grinblatt, Titman, and Wermers, 1997). In practice, as indicated by recent research, investors appear to emphasize simple benchmark comparisons when allocating capital, giving limited attention to a portfolio's exposures to size, book-to-market, and momentum factors.[1] Particularly, Sensoy (2009) finds that mutual fund investors react strongly to performance relative to a fund's prospectus benchmark index—i.e., the benchmark that a fund self-declares in its prospectus—even after controlling for performance relative to a fund's factor exposures.

This tendency of investors to focus on fund performance relative to the prospectus benchmark index would be appropriate if managers pursue strategies with risk similar to the prospectus benchmark. Otherwise, if investors compare portfolio performance to that of the prospectus benchmark without adjusting for differences in risk or factor exposures, investors are likely to over- or under-estimate alpha. For example, managers pursuing a strategy that is riskier than the prospectus benchmark may outperform the prospectus benchmark before risk-adjusting, but underperform after adjusting for the higher risk. Given the tendency of investors to respond to prior performance, these managers could attract additional inflows if investors do not make an appropriate adjustment for risk. Hence, the extent to which the prospectus benchmark captures the fund's investment strategy has important economic implications.

---

[1] See Sensoy (2009); Elton, Gruber, and Blake (2014); Barber, Huang, and Odean (2016); Berk and van Binsbergen (2016); and Agarwal, Green, and Ren (2018).

The main contribution of this study is to introduce a new holdings-based procedure that assesses whether a fund has a benchmark discrepancy, in which case a benchmark other than the prospectus benchmark better matches a fund's actual investment strategy. We implement this procedure using actively managed U.S. equity mutual funds. The SEC requires all mutual funds to self-declare a benchmark in their prospectus (i.e., the "prospectus benchmark") and mandates quarterly disclosure of complete portfolio holdings. Using our procedure, we show that for funds with a benchmark discrepancy, the prospectus benchmark typically understates the factor exposures of the fund and, accordingly, the prospectus benchmark is easier to beat than a benchmark with the same factor exposures as the fund. We show that these benchmark discrepancies have a significant economic impact on performance evaluation and capital allocation, as investors generally focus on fund performance relative to the prospectus benchmark even when a fund has a benchmark discrepancy. Our results suggest that investors could significantly improve their capital allocations by accounting for benchmark discrepancies when evaluating fund performance.

We begin by assessing which benchmark best captures a fund's actual investment strategy. To do this, we identify the benchmark that has the lowest active share with the fund's holdings (hereafter, the "AS benchmark"). If the AS benchmark is different from the prospectus benchmark (as it is in 67% of our sample), we next consider the extent to which the holdings of the prospectus and AS benchmarks differ. We assess that difference by calculating the active share of the prospectus benchmark relative to the AS benchmark (hereafter, *Benchmark Mismatch*). In many cases, *Benchmark Mismatch* is low, as the two benchmarks have holdings that largely overlap (e.g., S&P 500 and S&P 500 Growth have an active share of only 33% with respect to each other). In other cases, *Benchmark Mismatch* is quite high, such as, for example, funds that have a prospectus benchmark of the Russell 2000 and an AS benchmark of the S&P 600 Growth. While both indexes skew towards small-cap stocks, the stocks with the largest weights in the S&P 600 Growth are not even in the Russell 2000. The active share of the benchmarks with respect to each other is 77%, indicating that their holdings are quite different.

We label a fund as having a benchmark discrepancy if its *Benchmark Mismatch* is at least 60% (a criterion similar to that used in Cremers and Petajisto, 2009, to identify active managers). In these cases, fund holdings are not only better captured by the AS benchmark, but the AS benchmark is also substantially different from the prospectus benchmark. Applying this criterion, 26% of funds in our sample have a benchmark discrepancy. A fund is more likely to have a benchmark discrepancy if it has a high active share with respect to its prospectus benchmark and if its strategy focuses on small-cap or mid-cap stocks.

Next, we show that, for the set of funds with a benchmark discrepancy, the prospectus and AS benchmarks have meaningfully different returns. The average return of the AS benchmarks is 1.50% per year (*t*-stat = 3.20) higher than that of the prospectus benchmarks. In contrast, the AS and prospectus benchmark returns are not economically or statistically different for the set of funds with a *Benchmark Mismatch* greater than 0% but less than 60%. As a result, for funds we identify as having a benchmark discrepancy, a substantially lower return suffices to beat the prospectus benchmark compared to that needed to beat the AS benchmark.

The return differences between the prospectus and AS benchmarks lead to different conclusions about fund performance. We find that funds with a benchmark discrepancy have prospectus-benchmark-adjusted returns 1.04% per year (*t*-stat = 3.20) higher than funds without a benchmark discrepancy. However, when we adjust the performance of funds with a benchmark discrepancy using the AS benchmark instead, the benchmark-adjusted returns between the two groups are indistinguishable. If investors do not appropriately adjust for risk or factor exposures, then these benchmark discrepancies materially affect conclusions about *ex post* fund performance.

Benchmark discrepancies are most likely among high active share funds, a group that tends to outperform (Cremers and Petajisto, 2009). For funds with high active share (highest quintile) and a benchmark discrepancy, there is no evidence of outperformance when using the AS benchmark. However, the evidence that high active share funds outperform is considerably stronger when only high active share funds without a benchmark discrepancy are considered, using either suitable benchmark-adjusted returns or when calculating fund alphas using the Cremers, Petajisto, and Zitzewitz (2012) seven-factor model. For example, high active share funds with a benchmark discrepancy have an annualized seven-factor alpha of 0.07% (*t*-stat of 0.14), while high active share funds without a benchmark discrepancy have an alpha of 1.28% per year (*t*-stat = 2.14). Accounting for past performance increases the alpha of the latter group further. Funds that do not have a benchmark discrepancy and that are in the highest quintiles of active share and past benchmark-adjusted return have a seven-factor alpha of 3.21% per year (*t*-stat = 2.37).

Sensoy (2009) was the first study to introduce a procedure to identify funds with benchmark discrepancies, but our procedure differs from his. The procedure in Sensoy (2009) compares the style implied by a fund's prospectus benchmark to a fund's Morningstar (2004) style box, which capture the 3x3 intersection of large-mid-small cap styles with value-blend-growth styles. If a fund's prospectus benchmark style and Morningstar style do not match, then the monthly returns of the fund are regressed on the returns of the prospectus benchmark and, separately, on the returns of the benchmark corresponding to the Morningstar style. If the regression using the Morningstar style benchmark results in a higher $R^2$ than the regression using the prospectus benchmark, then the fund is classified as having a benchmark discrepancy, irrespective of the size of the difference in the $R^2$ between the two regressions. In contrast, our procedure is based on current fund holdings rather than the time-series of returns and considers the economic magnitude of the difference between potential benchmarks. As a result, while the two procedures identify a similar number of benchmark discrepancies in our sample—26% of fund-month observations using our procedure versus 23% using Sensoy's—they often disagree about which particular funds have a benchmark discrepancy.[2]

We compare the two procedures by focusing on funds that have a benchmark discrepancy according to one procedure but not the other. The average benchmark-adjusted performance of funds identified by our procedure, but not Sensoy's, as having a benchmark discrepancy is 1.33% ($t$-stat = 2.25) higher when using the prospectus benchmark instead of the AS benchmark. For funds identified by Sensoy's procedure, but not ours, as having a benchmark discrepancy, there is no difference in benchmark-adjusted return between the two benchmark choices. Therefore, benchmark discrepancies identified by our procedure have a larger impact on fund performance evaluation.

We next show that the higher average returns of the AS benchmarks relative to the prospectus benchmarks, among funds with a benchmark discrepancy, can be explained by the greater systematic factor exposures of the AS benchmarks. Traditional factors (i.e., size, value, and momentum) explain about a third of the average difference in returns between the AS and the prospectus benchmarks. The inclusion of non-traditional factors along with the traditional factors explains about 87% of the average difference in returns. Among the non-traditional factors, the profitability factor (*RMW*) of Fama and French (2015) has the largest impact.

[2] Across the 40% of funds for which at least one of the two procedures identifies a benchmark discrepancy, about half only have a benchmark discrepancy according to our procedure (17% of funds in the sample). Another 14% of funds are classified as having a benchmark discrepancy only according to the procedure in Sensoy (2009).

Subsequently we find that, for the sample of funds with a benchmark discrepancy, benchmark-adjusted fund returns based on the prospectus benchmark have substantial residual factor exposures, whereas benchmark-adjusted returns based on the AS benchmark do not. Therefore, performance evaluation using AS-benchmark-adjusted returns—but not using prospectus-benchmark-adjusted returns—results in conclusions similar to those from employing factor models (i.e., calculating abnormal fund returns based on factor model regressions on excess fund returns). This matters most for large-cap funds with a benchmark discrepancy, where benchmark-adjusting using the prospectus benchmarks results in substantially higher average fund returns compared to using the AS benchmarks (difference of 2.41% per year, *t*-stat of 2.10). All of that difference can be explained by exposure to both traditional and non-traditional factors.

Finally, we consider the impact of different performance measures on fund flows, building on Sensoy (2009). The economic importance of the benchmark discrepancies we document depends on the performance evaluation methods used by investors. The more investors rely on fund performance relative to the prospectus benchmark—rather than relative to the AS benchmark or to a fund's factor exposures—the more capital allocation decisions may be affected by benchmark discrepancies. In line with previous work, we find that investors give substantial weight to fund performance relative to the prospectus benchmark when allocating capital, even when a benchmark discrepancy exists. However, we expand on that result by showing that a prospectus benchmark overstating the performance of the fund by 1% has only about half the effect of an actual 1% increase in performance. We also find that performance relative to the prospectus benchmark has a decreasing impact on investor flows as *Benchmark Mismatch* increases and as the size of the difference in the returns between the AS benchmark and prospectus benchmark increases (although a meaningful effect remains).

In light of our performance results above, investors could considerably improve their capital allocations by focusing their attention on funds with prospectus benchmarks that match the actual investment strategy. Specifically, if investors account for benchmark discrepancies while also considering past performance and active share, they can identify funds with large, positive, statistically significant alphas. Funds that do not have a benchmark discrepancy and that are in the highest quintiles of benchmark-adjusted return and active share have an average seven-factor alpha of 3.21% per year (*t*-stat = 2.37).

## 2. Comparison with prior work

Several studies show that the prospectus benchmarks and declared styles of mutual funds are often inaccurate. DiBartolomeo and Witkowski (1997); Kim, Shukla, and Tomas (2000); Elton, Gruber, and Blake (2003, 2014); Hirt, Tolani, and Philips (2015); Bams, Otten, and Ramezanifar (2017); and Mateus, Mateus, and Todorovic (2017) all show evidence of apparent misclassification, but our study is most comparable to Sensoy (2009). He finds that about 31% of mutual funds have a benchmark discrepancy and that investor flows are influenced by the performance of a fund relative to its prospectus benchmark even when a fund has a benchmark discrepancy.

Our study differs in several important ways from Sensoy (2009). First, our procedure for identifying benchmark discrepancies is substantially different from Sensoy's procedure. We compare fund and benchmark holdings to identify benchmark discrepancies, while Sensoy uses Morningstar (2004) style boxes and fund returns. As explained above, he labels a fund as having a benchmark discrepancy if two conditions are met. First, the fund's Morningstar style must not match the fund's style as implied by the prospectus benchmark. Second, the returns on the benchmark that corresponds to the fund's Morningstar style must have a greater correlation with the full sample of a fund's returns than the returns on the prospectus benchmark.

The benchmark discrepancies from Sensoy's procedure are binary, identified *ex post*, and time invariant. In comparison, our procedure allows us to measure the economic magnitude of the benchmark discrepancy, using *Benchmark Mismatch*; to identify the appropriate benchmark *a priori*, which Sharpe (1992) labels a key component of a benchmark; and to capture time-variation in the appropriate benchmark in response to changes in a fund's reported holdings, which is important given that Huang, Sialm, and Zhang (2011) show significant time variation in fund risk.

Furthermore, our procedure for identifying benchmark discrepancies is more "factor agnostic." Put another way, it is less explicit about which factors a fund should match with respect to its benchmark. Sensoy's procedure can only capture benchmark discrepancies across the two traditional factors of size and value: a fund with a small-cap value prospectus benchmark that is labeled small-cap value by its Morningstar style box cannot have a benchmark discrepancy using his procedure. In comparison, our procedure can identify benchmark discrepancies across other factors despite the fact the benchmarks used in our procedure are nominally built based on the size and value factors. For example, the Russell 2000 Value and S&P 600 Value are both small-cap value benchmarks, but our procedure will identify a benchmark discrepancy for a fund that lists one as their prospectus benchmark and has the other as their AS benchmark, as the holdings of those benchmarks differ significantly from each other (average *Benchmark Mismatch* = 69%).

As a result of all of these differences, our procedure identifies many funds as having a benchmark discrepancy that Sensoy (2009) does not, and vice versa. Among funds that only have a benchmark discrepancy according to our procedure, the average difference in return between the AS benchmark and prospectus benchmark is economically large and statistically significant.

However, when only Sensoy's procedure identifies a benchmark discrepancy, the difference in returns between the alternative benchmark and the prospectus benchmark is economically small and statistically zero. As a result, the benchmark discrepancies identified by our procedure have larger economic implications.

Beyond differences in identification procedures, our study differs from Sensoy (2009) in several other ways. In our analysis, we focus on the magnitude of and explanations for the difference in returns between the prospectus benchmark and the AS benchmark. We find that funds with benchmark discrepancies use prospectus benchmarks that have lower returns than their AS benchmarks and that most of that difference in returns can be attributed to differences in factor exposures. Sensoy does not provide a similar analysis. We also provide novel insights into the responsiveness of investors to performance relative to the prospectus benchmark. For instance, we estimate the marginal impact on investor capital allocations from funds employing a prospectus benchmark that understates actual risk and show how the magnitude of *Benchmark Mismatch* affects that marginal impact. Moreover, our analysis demonstrates how accounting for benchmark discrepancies can improve an investor's ability to identify funds that can be expected to outperform in the future.

## 3. Data

### *3.1. Mutual fund sample*

Our sample of actively managed mutual funds comes from the Center for Research in Security Prices (CRSP) Survivor-Bias-Free U.S. Mutual Fund database. We focus on U.S. equity funds, although our analysis could directly be applied to other styles. To identify actively managed funds that almost exclusively invest in U.S. equities, we first exclude any fund that CRSP identifies as an index fund, ETF, or variable annuity; use only funds with Lipper, Strategic Insight, or

Wiesenberger investment objective codes consistent with following a traditional long-only U.S. equity strategy; and require funds to invest at least 80 percent of their assets in common stock. We filter out funds with names associated with index funds or strategies other than traditional long-only U.S. equity strategies.[3] We address the incubation bias identified by Evans (2010) by excluding a fund from the sample until it is at least two years old and until it first reaches at least $20 million in assets.

All of our analysis is conducted at the fund level. We aggregate information across multiple share classes of a fund using the WFICN variable available from MFLINKS. Fund assets are the sum of the assets across all share classes. All other fund characteristics, including returns and expense ratios, are calculated as the asset-weighted average of the share class values.

We collect information on funds' prospectus benchmarks from Morningstar Direct and match that data to CRSP using ticker and CUSIP. A fund is dropped from the sample if we cannot match it to Morningstar Direct or if Morningstar Direct does not provide a prospectus benchmark. The data on the prospectus benchmarks is cross-sectional, rather than time-series, but changes in the prospectus benchmark are considered very rare.

Securities and Exchange Commission (SEC) rules first required mutual funds to provide a benchmark to investors in certain documents released after July 1, 1993.[4] Since the period used in our analysis is 1991 through 2015, there is the potential for survivor bias in the first few years of

---

[3] The list of terms used in this search is available upon request.

[4] The rule specifically requires all mutual funds provide "a line graph comparing its performance to that of an appropriate broad-based securities market index" as part of "its prospectus or, alternatively, in its annual report to shareholders." It is common to cite December 1, 1998 as the time mutual funds were first required to provide a benchmark to investors; however, that rule only added the requirement that all mutual funds compare "the fund's average annual returns for 1, 5, and 10 years with that of a broad-based securities market index" to the preexisting disclosure. See Final Rule: Disclosure of Mutual Fund Performance and Portfolio Managers, https://www.sec.gov/rules/final/33-6988.pdf, and Final Rule: Registration Form Used by Open-End Management Investment Companies, https://www.sec.gov/rules/final/33-7512r.htm.

the sample. However, we find that the probability of survivorship in the 1991-1993 CRSP sample is not related to having a prospectus benchmark in Morningstar Direct. Furthermore, we find economically negligible differences in our results across the pre-1993 and post-1993 sub-periods, and our conclusions are the same regardless of whether we include the pre-1993 data.[5]

### *3.2. Mutual fund holdings*

We use the Thomson Reuters Mutual Fund Holdings database as our source of mutual fund holdings. As shown in Schwarz and Potter (2016), this data is not always consistent with the data filed by mutual funds with the SEC; however, they find little evidence of systematic bias. The holdings data only contains information on funds' equity positions, so any non-equity positions, including cash, are not reflected. We drop any holdings reports with fewer than 20 equity positions, which is an unusual occurrence and may indicate the holdings report is incomplete.[6]

This data is merged first with the CRSP stock database to obtain price information and adjust for stock splits. It is then merged with the CRSP fund database using MFLINKS. To verify that match, we drop any funds which have asset values in Thomson Reuters and CRSP that are not approximately the same or have implied gross fund returns from Thomson Reuters and net fund returns from CRSP that are not highly correlated.

### *3.3. Benchmark holdings*

Our procedure to determine which funds have a benchmark discrepancy involves a comparison of a fund's holdings to the holdings of a set of benchmark indices (which always includes a fund's prospectus benchmark). We limit our sample of funds to those with the following

---

[5] For example, the difference in returns between the prospectus benchmark and the AS benchmark for funds with a benchmark discrepancy is about the same pre-1993 as post-1993. The difference between the two periods is only 0.01% per year (*t*-stat = 0.01).

[6] If not incomplete, these funds would likely have difficulty satisfying the requirements to be considered diversified under the Investment Company Act of 1940.

prospectus benchmarks: the Russell 1000, Russell 2000, Russell 3000, Russell Midcap, S&P 500, S&P 400, and S&P 600, plus the value and growth components of those seven benchmarks. The primary reason for this condition is that these 21 benchmarks are well-diversified and commonly referenced by investors. This benchmark set contains the prospectus benchmark for 97.4% of the funds in our initial sample, indicating they are among the most popular for funds to self-declare in their prospectus.

By considering just these 21 benchmarks when comparing holdings (i) we do not assign any AS benchmark that is outside of the set normally considered by actual funds and (ii) we generate a more effective interpretation of benchmark discrepancies. If we use a more expanded set of possible benchmarks, including more concentrated and rarely used indices, then it is more likely that a significant overlap with a benchmark other than the prospectus benchmark will be accidental or caused by active stock-picking. Specifically, a fund that is following the style of its prospectus benchmark while also doing a lot of active stock-picking, may, by chance, have holdings similar to a relatively obscure index. Our set of benchmarks includes the set of twelve used in Sensoy (2009), who gives similar reasons for his choice.

Our data on benchmark holdings comes from multiple sources. Russell provided us the constituent weights for their benchmarks, while the constitution weights for the S&P benchmarks come from Compustat. Monthly return data for the benchmarks (with dividends reinvested) comes from Morningstar Direct. Our final sample consists of 197,643 fund-month observations across 1,216 unique funds. The number of funds varies over time. Our sample has 142 unique funds in 1991, 299 in 1996, 633 in 2001, 901 in 2006, 1,053 in 2011, and 931 in 2015.

## 4. Benchmark discrepancy methodology

### *4.1. The active share (AS) benchmark*

We classify a fund as having a benchmark discrepancy if two conditions are satisfied. First, a benchmark in our set matches the fund's actual investment style better than the prospectus benchmark. Second, that alternative benchmark is substantially different from the prospectus benchmark, i.e., the differences between the two benchmarks are economically meaningful.

Our method of determining the best match focuses on holdings. We determine the benchmark that best matches a given fund's actual investment style by finding the benchmark whose holdings have the greatest overlap with that fund's holdings. The extent to which a fund's holdings overlap with a given benchmark is measured using Cremers and Petajisto's (2009) active share, defined as:

$$Active\ Share = \frac{1}{2}\sum_{i=1}^{N}\left|w_{i,f} - w_{i,b}\right| \tag{1}$$

where $w_{i,f}$ is the weight on stock *i* in the fund's portfolio and $w_{i,b}$ is weight on stock *i* in the benchmark. The measure is calculated over all *N* stocks in the investable universe. An alternative formula for active share is given in Cremers (2017):

$$Active\ Share = 1 - \sum_{i=1}^{N} MIN\left(w_{i,f}, w_{i,b}\right) * d[w_{i,f} > 0] \tag{2}$$

where $d[w_{i,f} > 0]$ is a dummy variable equal to one if stock *i* has a positive weight in the fund's portfolio. This version of the active share formula in Eq. (2) produces the same active share values as the prior formula in Eq. (1), as long as the fund does not employ leverage or shorting, but emphasizes that active share is only lowered by overlapping weights (i.e., active share is equal to 1 minus the sum of the overlapping weights).

As active share increases, the fund and a given benchmark are less alike. Assuming all weights are positive (i.e., the fund does not short any shares, which is the case for almost all funds in our sample), an active share of 0% means the fund and a given benchmark are identical and an active share of 100% means the fund and a given benchmark share no stocks in common. Accordingly, we consider the benchmark that best matches a fund to be the benchmark that results in the lowest active share (considered across all 21 of our benchmarks). We label that benchmark the minimum active share benchmark, or simply the 'AS' benchmark.

The AS benchmark for a given fund is re-determined every time our data provides a new report of that fund's holdings, which is quarterly in most instances. Allowing the benchmark to vary over time is important because fund style and risk are not time invariant. Chan, Chen, and Lakonishok (2002), Brown, Harlow, and Zhang (2009), and Cao, Iliev, and Velthuis (2017) all show evidence of style drift, while Brown, Harlow, and Starks (1996) and Huang, Sialm, and Zhang (2011) show variation over time in overall fund risk taking. The AS benchmark is always assigned ex-ante, such that an AS benchmark assigned to a fund at the end of quarter $t$ is used for analyzing the fund in quarter $t + 1$.[7,8]

### *4.2. Benchmark Mismatch*

After identifying the set of funds for which the prospectus benchmark differs from the AS benchmark, we determine whether the AS benchmark is meaningfully different from the prospectus benchmark. This step is important because in many cases the AS benchmark and prospectus benchmark are quite similar, simply because many benchmarks in our set of 21 are

[7] It is not uncommon for a fund's AS benchmark to change. The average fund is in our sample for 13.5 years and changes its AS benchmark 12.7 times (using an average of 3.7 different AS benchmarks). However, many of these changes are not economically meaningful. For example, more than half of all changes are between AS benchmarks that both result in a *Benchmark Mismatch* of less than 60%. If we lessen the number of changes by using the mode of a fund's AS benchmark over the previous three years, our primary results are unchanged.

[8] We find little evidence of market timing by funds related to changes in the AS benchmark. For example, in the month after an AS benchmark change, the average difference in return between the new and old AS benchmark is only 0.94 basis points ($t$-stat = 0.38). That difference grows to just 2.64 basis points ($t$-stat = 0.65) when the sample is limited to changes in which *Benchmark Mismatch* increases and just 4.54 basis points ($t$-stat = 1.18) when limited to changes in which *Benchmark Mismatch* increases by at least 30% (i.e., within the highest quartile of *Benchmark Mismatch* changes).

quite similar to each other. For example, the Russell 1000 and Russell 3000 have an active share of 8.0% relative to each other (averaged across our sample period). Given the similarity in those benchmarks' holdings, a fund with the Russell 1000 as its prospectus benchmark and the Russell 3000 as its AS benchmark may have a difference in benchmarks, but that difference is not economically important.

We measure the extent to which the prospectus and AS benchmarks are different using the lack of overlap in their respective holdings, i.e., using the active share between the two benchmarks. We label the measure *Benchmark Mismatch* and calculate it as follows:

$$Benchmark\ Mismatch = \frac{1}{2}\sum_{\mathrm{i}=1}^{\mathrm{N}}\left|w_{i,p} - w_{i,AS}\right| \tag{3}$$

where $w_{i,p}$ is the weight on stock *i* in the fund's prospectus benchmark and $w_{i,AS}$ is weight on stock *i* in the fund's AS benchmark.[9] When the holdings of the two benchmarks largely overlap, the active share of the prospectus benchmark with respect to the AS benchmark is low and thus *Benchmark Mismatch* will be small. Hence, an increase in *Benchmark Mismatch* represents an increase in the difference between the holdings of the two benchmarks (or a decrease in the overlap of holdings).

Since *Benchmark Mismatch* captures how different the holdings of the prospectus benchmark are from the holdings of the AS benchmark, we can directly interpret *Benchmark Mismatch* as a measure of the economic magnitude of the differences in those benchmarks. For the main results in the paper, we classify funds with *Benchmark Mismatch* above 60% as having significant economic differences in their benchmarks and thus having a benchmark discrepancy.

[9] As with the active share of a fund relative to a given benchmark, the active share of the benchmarks relative to each other can also be calculated using the MIN() specification.

While the 60% cutoff is somewhat arbitrary, we set the threshold there for two reasons. First, a similar 60% cutoff is employed in prior work using active share (e.g., Cremers and Petajisto, 2009, and Cremers and Curtis, 2016). Funds with an active share less than 60% are labeled "closet indexers." Second, as shown in section 6, analysis of the returns on the benchmarks suggests that the economic differences between the prospectus benchmarks and the AS benchmarks of funds with *Benchmark Mismatch* less than 60% are minor on average.

We also consider the difference between the active share of a fund with respect to its prospectus benchmark and the active share of a fund with respect to its AS benchmark. We label this difference the *Active Gap*:

$$Active\ Gap = Active\ Share_{Prospectus} - Active\ Share_{AS} \tag{4}$$

As *Active Gap* increases, the gap between the overlap of a fund's holdings with its AS benchmark and the overlap of a fund's holdings with its prospectus benchmark increases. Unlike *Benchmark Mismatch*, *Active Gap* does not directly measure whether the prospectus benchmark and AS benchmark are meaningfully different. However, conditional on having large *Benchmark Mismatch*, *Active Gap* does indicate the extent to which the activeness of the fund is overstated by using the prospectus benchmark instead of the AS benchmark.

### *4.3. Summary Statistics*

Table 1 shows summary statistics for the key measures used in our study. The dummy variable *Any Mismatch*, which is equal to one if *Benchmark Mismatch* is above 0%, shows that about 67% of quarterly fund observations have a prospectus benchmark different from the AS benchmark. If we require *Benchmark Mismatch* to be greater than 60%, as motivated above, then only 26% of observations have a difference in benchmarks (as indicated by the dummy variable *Large Mismatch*). While the AS benchmark is re-determined with each new holdings report, funds with an AS benchmark that is different from their prospectus benchmark tend to maintain that difference. The correlation between the value of *Any Mismatch* (*Large Mismatch*) in month $t$ and month $t$ – 12 is 0.59 (0.75). Further, if the prospectus benchmark and AS benchmark are different in month $t$, then the probability they will still be different in month $t$ + 12 is 86%. That result is about the same if the analysis is limited to funds with a *Benchmark Mismatch* greater than 60%.

[Table 1 about here]

As shown in Figure 1, the frequency of benchmark differences varies through time. The percentage of funds with a difference of any magnitude is as low as 61% (in 1994, 2005, and 2009) and as high as 78% (in 1995) with no obvious trend. The number of funds with a *Benchmark Mismatch* greater than 60% varies within the range of 14% to 37%. The significant jumps in the number of such large differences in 1992 and 1998 are a result of new benchmarks entering the available set.[10] After 1998, when all 21 benchmarks in our set are available and all funds are legally required to provide a benchmark, the number of funds with a *Benchmark Mismatch* above 60% slowly decreases from 30% to 21%.

[Figure 1 about here]

The average active share is 80.7% using the prospectus benchmark, compared to 78.4% using the AS benchmark. As such, that difference (i.e., the *Active Gap*) has a mean of 2.3%. Active share is persistent over time, as the annual autocorrelation is above 0.90 using either the prospectus or AS benchmark.

Using net fund returns, the average fund underperforms both the prospectus and AS benchmarks (ignoring the costs of investing in those benchmarks); however, the degree of underperformance differs. Relative to the prospectus benchmark, funds on average underperform

[10] Excluding the S&P 500, which is available from the start of our sample, the S&P benchmarks enter into our sample over the period 1992 through 1997. Each of the Russell benchmarks is available from the start of our sample.

by 0.33% per year, which is not statistically distinguishable from zero ($t$-stat = −1.06). In comparison, funds underperform by 0.78% per year relative to their AS benchmark, which is statistically significant ($t$-stat = −2.87). This suggests that the choice of benchmark does affect the evaluation of fund performance if one uses a simple comparison of the performance of the fund relative to the performance of its benchmark, as prospectus benchmarks have noticeably lower returns compared to AS benchmarks.

The average *Benchmark Mismatch* and *Active Gap* are 34.9% and 2.3%, respectively, but those values are pushed downward by funds with the same prospectus and AS benchmark. Figure 2 shows the cumulative density function (CDF) of *Active Gap* for funds with different prospectus and AS benchmarks. For most of these funds, *Active Gap* is small. About 38% of funds have *Active Gap* below 2%, and about 78% have *Active Gap* below 5%. However, among the funds with *Active Gap* above 5% (15% of the full sample of funds), 19% have *Active Gap* greater than 10% (3% of the full sample).

[Figure 2 about here]

Figure 3 shows the CDF of *Benchmark Mismatch* for the same sample of funds. *Benchmark Mismatch* is below 60% for most funds. However, 38% of these funds (26% of the full sample) have *Benchmark Mismatch* above 60%, which implies there is a large economic difference between the prospectus and AS benchmark. Among that group, about half of the funds have a *Benchmark Mismatch* above 80%. If those funds had exactly the same holdings as their AS benchmark (i.e., if they had an active share of 0% with respect to the AS benchmark), then an investor using the prospectus benchmark would conclude those funds have an active share of at least 80%.

[Figure 3 about here]

## 5. Funds with large versus small *Benchmark Mismatch*

Before considering performance, we first analyze the characteristics of funds as a function of *Benchmark Mismatch*. As mentioned before, we separate funds using a *Benchmark Mismatch* cut-off of 60%. We refer to funds with a *Benchmark Mismatch* above 60% as having a benchmark discrepancy.

Comparing the prospectus and AS benchmarks of funds with a positive *Benchmark Mismatch*, funds with and without a benchmark discrepancy differ in several ways. Table 2 shows the five most common benchmark combinations for each group. Funds with small, but non-zero, *Benchmark Mismatch* have prospectus and AS benchmarks that are quite similar. By our construction, the prospectus and AS benchmarks of these funds are closet indexers of each other. The most common difference, an S&P 500 prospectus benchmark and an S&P 500 growth AS benchmark, has a *Benchmark Mismatch* of 33.0%. In most cases when *Benchmark Mismatch* is small, the AS benchmarks is close to or is a complete subset of the prospectus benchmark (or vice versa).

[Table 2 about here]

Conversely, the funds with a benchmark discrepancy have large differences between their prospectus and AS benchmarks, as their *Benchmark Mismatch* exceeds the 60% cut-off. The most common grouping is the set of funds with a Russell 2000 prospectus benchmark and an S&P 600 Growth AS benchmark, with a *Benchmark Mismatch* of 77.1%. Those benchmarks have limited overlap: the Russell 2000 contains all of the stocks with a market capitalization ranking between 1001 and 3000, whereas the S&P 600 Growth contains the growth stocks within the full set of stocks with a market cap ranking between 901 and 1500. As a result, even if all stocks with a market cap ranking between 1001 and 1500 are labeled growth by S&P, the two benchmarks could have at most 500 stocks in common. More importantly, since both benchmarks weight by market capitalization, any overlapping stocks should have relatively large weights in the Russell 2000 and relatively small weights in the S&P 600 Growth. The growth stocks with a market cap ranking between 901 and 1000 will have the largest weights on any stocks in the S&P 600 Growth, but zero weight in the Russell 2000.

Table 3 compares the characteristics of funds with and without a benchmark discrepancy. In this analysis, the group without a benchmark discrepancy includes funds with a *Benchmark Mismatch* of zero. Funds with a benchmark discrepancy tend to be more actively managed. The average active share (with respect to the prospectus benchmark) for those funds is 93.5%, compared to 76.8% for funds without a benchmark discrepancy.[11] Funds with a benchmark discrepancy also have fewer assets, are younger, and charge a greater expense ratio. With respect to style (as defined by the prospectus benchmark), funds with a benchmark discrepancy tend to disproportionately have a style classification of small-cap or mid-cap. About 80.4% of funds with a benchmark discrepancy have a small- or mid-cap style, while only 23.7% of funds without a benchmark discrepancy have those styles. The differences in growth and value style between the two groups are slight in comparison.

[Table 3 about here]

Next, we consider the relation between having a benchmark discrepancy and fund characteristics using the following model:

$$BM > 60\%_{i,t} = \alpha + \beta * Active\ Share_{i,t} + \delta * Chars_{i,t} + \gamma * Style_i + FE + \varepsilon_{i,t} \qquad (5)$$

where $BM > 60\%_{i,t}$ is a dummy variable equal to one if the *Benchmark Mismatch* for fund *i* based on holdings in quarter *t* is greater than 60%. $Active\ Share_{i,t}$ is a vector of information about fund

[11] About 74% of funds in the highest quintile of prospectus active share have a benchmark discrepancy.

*i*'s active share in quarter $t$. It includes the prospectus active share and a dummy variable equal to one if the prospectus active share is among the top 20% in the quarter. $Chars_{i,t}$ is a vector of characteristics for fund *i* available as of quarter *t* and includes the natural log of assets, natural log of age, expense ratio, turnover ratio, the number of equity positions, and the percentage of fund assets held within institutional share classes. $Style_i$ is a vector of information about fund *i*'s style. It includes a large-cap dummy, a blend dummy, and a growth dummy. $FE$ represents year-quarter fixed effects. We estimate the model using a logit regression and the full sample of fund-quarters, including funds with a *Benchmark Mismatch* of zero.

Table 4 presents the results from this regression using *t*-statistics derived from standard errors clustered on both fund and year-month. As prospectus active share increases, the probability of having a benchmark discrepancy increases. After controlling for fund characteristics and style, that relation becomes non-linear. Funds in the top 20% of prospectus active share are more likely to have a benchmark discrepancy than the linear term indicates. The full model without fixed effects in Column 5 predicts that a fund at the 50$^{th}$ percentile of prospectus active share and the mean of all other variables has a 6.1% probability of having a benchmark discrepancy. If that same fund instead had a prospectus active share at the 85th percentile, that probability would increase to 70.1%. The fund characteristics have either limited economic significance or limited statistical significance when considered in the full model, but fund style contains substantial predictive power. Funds with a large-cap style are significantly less likely to have a benchmark discrepancy compared to funds that have a small- or mid-cap style. Returning to the fund at the 85th percentile of prospectus active share, the full model without fixed effects in column 5 predicts that if that fund was a large-cap fund its probability of having a benchmark discrepancy would be 61.3%. In comparison, that probability would be 81.5% if that fund was a small- or mid-cap fund.

[Table 4 about here]

## 6. Differences in prospectus and AS benchmark returns and their performance implications

This section first considers whether the prospectus benchmark gives a fund a "performance boost" when benchmark-adjusting returns, i.e., when evaluating fund performance by comparing it against a benchmark index's performance rather than by using a factor model. We answer that question by comparing the returns on the prospectus and AS benchmarks for funds with a non-zero *Benchmark Mismatch*. We then consider (1) how conclusions about a fund's performance can change depending on the benchmark used and (2) how accounting for benchmark discrepancies affects an investor's ability to select funds that can be expected to outperform in the future.

### *6.1. Comparing benchmark returns*

Table 5 shows the average difference in annualized return between the AS benchmark and prospectus benchmark (i.e., the performance boost) for funds depending on their *Active Gap* and *Benchmark Mismatch*. Panel A divides funds into five ranges of *Benchmark Mismatch* and Panel B divides funds based on whether *Benchmark Mismatch* is above or below 60%. The ranges for *Active Gap* are the same in both panels. Funds with a *Benchmark Mismatch* of zero are excluded from this analysis.

[Table 5 about here]

Focusing first on Panel A, the average performance boost for funds with a non-zero *Benchmark Mismatch* is 0.68% per year ($t$-stat = 2.72). However, the performance boost is considerably higher for funds with a higher *Benchmark Mismatch*. Funds with a *Benchmark Mismatch* greater than 80% have an average performance boost of 1.64% per year ($t$-stat = 2.97), compared to −0.12% per year ($t$-stat = −0.75) for funds with a *Benchmark Mismatch* less than 20%. *Active Gap* matters as well, though on a much more limited basis. Compared to funds with an *Active Gap* less than 1.25%, the average performance boost for funds with an *Active Gap* greater than 5% is 0.42% per year higher (*t*-stat = 1.31).

Once *Benchmark Mismatch* is greater than 60%, the average performance boost is consistently economically large and statistically significant. Funds with a *Benchmark Mismatch* between 60% and 80% have an average performance boost of 1.37% per year (*t*-stat = 2.25), and the performance boost for that group is at least 1% per year in each of the different ranges of *Active Gap*. There is some evidence of a performance boost for funds with a *Benchmark Mismatch* between 40% and 60%, but on average, it is economically much smaller (0.53%) and statistically weaker (*t*-stat = 1.70). The performance boost for that group also varies without an obvious trend depending on *Active Gap*. Overall, after controlling for *Benchmark Mismatch*, *Active Gap* appears to matter little.[12]

If we group funds based on whether *Benchmark Mismatch* is above and below 60%, as in Panel B, the results are similar. The average performance boost when *Benchmark Mismatch* is greater than 60% is 1.50% per year (*t*-stat = 3.20), compared to 0.18% (*t*-stat = 0.94) when *Benchmark Mismatch* is less than 60%. *Active Gap* has negligible impact within those groups. Each *Active Gap* range for funds with *Benchmark Mismatch* greater than 60% shows an economically large and statistically significant average performance boost. Further, within that group, there is no difference in average performance boost between funds with low and high *Active Gap*. Conversely, funds with *Benchmark Mismatch* less than 60% have average performance boosts that are economically small and statistically insignificant regardless of *Active Gap*. As a

[12] High *Active Gap* appears to be related to a lower performance boost for funds with a *Benchmark Mismatch* less than 20%. However, there are very few funds with an *Active Gap* greater than 3.75% and a *Benchmark Mismatch* less than 20% (< 1% of the tested sample), so we believe caution should be exercised in making any inferences concerning those funds.

result, we conclude that the prospectus benchmark on average sets a lower bar for funds to clear than the AS benchmark only when *Benchmark Mismatch* is large.[13]

We consider the determinants of the performance boost more robustly using the following model:

$$R_{AS,i,t} - R_{P,i,t} = \alpha + \beta * Active\ Share_{i,t} + \delta * Mismatch_{i,t} + \gamma * Chars_{i,t} + FE + \varepsilon_{i,t} \qquad (6)$$

where $R_{AS,i,t}$ is the annualized return on fund *i*'s AS benchmark in month *t* and $R_{P,i,t}$ is the annualized return on fund *i*'s prospectus benchmark in month *t*. $Active\ Share_{i,t}$ is a vector of information about fund *i*'s active share at the start of month *t* that includes the fund's prospectus active share and a dummy variable equal to one if the prospectus active share is among the top 20% at the start of the month. $Mismatch_{i,t}$ is a vector of information about fund *i*'s mismatch status at the start of month *t*. It includes *Benchmark Mismatch*, *Active Gap*, and a dummy variable equal to one if *Benchmark Mismatch* is among the top 20% at the start of the month. $Chars_{i,t}$ is the same vector of the characteristics used in Eq. (5) measured for fund *i* as of the start of month *t*. $FE$ represents style and year-month fixed effects. We estimate the model using the sample of funds with a non-zero *Benchmark Mismatch*.

Table 6 presents the results of these performance boost regressions. Isolated from each other, active share, *Benchmark Mismatch*, and *Active Gap* each predict the performance boost. However, when considered simultaneously, only *Benchmark Mismatch* and active share continue to have predictive power. A 1% increase in *Active Gap* is associated with an increase in the performance boost of 0.095% per year (*t*-stat = 2.62) in column 3, but an increase of only 0.040%

[13] Elton, Gruber, and Blake (2014) primarily study separate accounts, but they also briefly consider mutual funds and, similar to our results here, show a difference in return between mutual funds' prospectus benchmarks and their full sample maximum correlation benchmarks (0.74% per year). However, they report neither separate results for the mutual funds whose benchmarks actually differ nor the number of mutual funds whose benchmarks actually differ, and it is uncertain whether the average difference in performance they document is statistically significant.

per year ($t$-stat = 1.23) in column 4, wherein *Benchmark Mismatch* and active share are included in the model as well. In comparison, a 1% increase in *Benchmark Mismatch* is associated with an increase in the performance boost of 0.033% per year ($t$-stat = 3.21) in column 2 and of 0.023% per year ($t$-stat = 2.54) in column 4.[14] The relation between active share and the performance boost is consistently strong, but non-linear. A fund in the highest quintile of active share has a performance boost 0.46% per year higher ($t$-stat = 2.07) than that implied by the linear coefficient, as shown in column 7.

[Table 6 about here]

Overall, these results show that a substantial number of funds have a prospectus benchmark that on average is easier to outperform compared to the benchmark implied by fund holdings. Funds with this performance boost can be identified using *Benchmark Mismatch* and active share.

### *6.2. Comparing funds' benchmark-adjusted returns*

The previous section shows how the prospectus benchmark can set a lower bar for a fund to clear than the AS benchmark if fund performance is evaluated through a simple comparison with the performance of the fund's benchmark. We now consider how different performance evaluation methods lead to different conclusions about fund performance, depending on whether the fund has a benchmark discrepancy (i.e., *Benchmark Mismatch* above 60%).

In this analysis, we independently double sort all funds (including those with a *Benchmark Mismatch* of zero) into groups based on prospectus active share and *Benchmark Mismatch*. Using active share, we sort funds into quintiles, and using *Benchmark Mismatch*, we sort funds into two groups based on the 60% cut-off. It is rare for funds to have a benchmark discrepancy (i.e., a high

[14] Note that a 1% increase in *Active Gap* is a much bigger change than a 1% increase in *Benchmark Mismatch*. In this sample, the standard deviation of *Active Gap* is 2.7%, compared to 24.7% for *Benchmark Mismatch*.

*Benchmark Mismatch*) and a low active share; therefore, to avoid reporting results for groups with very few funds, we collapse the four lowest quintiles of active share into a single group.

We evaluate average net performance within each group using three performance evaluation models: the prospectus-benchmark-adjusted return, the BM-benchmark-adjusted return, and the Cremers, Petajisto, and Zitzewitz (2012) seven-factor model (henceforth, the CPZ7 model). The BM-benchmark-adjusted return is the fund return minus the prospectus benchmark (AS benchmark) return if *Benchmark Mismatch* is less (greater) than 60%. It represents how the prospectus-benchmark-adjusted returns would appear if funds with a benchmark discrepancy were no longer evaluated relative to their prospectus benchmark but were instead evaluated relative to their AS benchmark. The alpha from the CPZ7 model represents the abnormal performance after accounting for funds' exposures to size, value, and momentum factors. It does not rely on the assignment of a singular benchmark, and it helps confirm whether our inferences using the benchmark-adjusted returns are valid.[15]

Table 7 shows the results from this analysis. When we do not sort on active share (see the top part of Table 7 that considers 'All Funds'), funds with a large *Benchmark Mismatch* on average outperform funds with a small *Benchmark Mismatch* by 1.04% per year (*t*-stat = 3.20) using the prospectus benchmark. However, using the BM benchmark, there is no statistically or economically significant difference in performance. The CPZ7 model indicates some outperformance by funds with a large *Benchmark Mismatch* relative to those with a small *Benchmark Mismatch*, but the economic size of the difference is smaller than it was with the prospectus benchmark (only 0.54% per year) and the difference is not statistically significant at conventional levels (*t*-stat = 1.46). Importantly, the positive relation between *Benchmark*

[15] The results using the alternative Cremers, Petajisto, and Zitzewitz (2012) four-factor model are similar to those from their seven-factor model.

*Mismatch* and active share prevents us from drawing any strong conclusions. Funds with a *Benchmark Mismatch* greater than 60% tend to have higher active share and, as shown in Cremers and Petajisto (2009), higher active share predicts better performance.

[Table 7 about here]

Therefore, we turn next to results conditional on active share. Within the high active share quintile and when using prospectus-benchmark-adjusted returns, funds perform about the same whether they have or do not have a benchmark discrepancy. Both groups show marginal evidence of outperformance (about 0.7% per year) that is marginally insignificant statistically at conventional levels (*t*-stats of about 1.6). If we compare fund performance using BM-benchmark-adjusted returns instead, the performance evaluation changes substantially. Among high active share funds with a benchmark discrepancy (i.e., the group where 'BM > 60%'), average prospectus-benchmark-adjusted performance is 0.72% per year (*t*-stat = 1.55), while average BM-benchmark-adjusted performance is −0.92% per year (*t*-stat = −2.32). In other words, while high active share funds with a benchmark discrepancy on average outperform their prospectus benchmark (albeit without strong statistical significance), they clearly underperform their AS benchmark (with strong statistical significance).

Using BM-benchmark-adjusted returns, high active share funds without a benchmark discrepancy outperform high active share funds with a benchmark discrepancy by 1.67% per year (*t*-stat = 3.32). Turning to the factor models, the CPZ7 alpha of high active share funds with a benchmark discrepancy is 0.07% per year (*t*-stat = 0.14). In comparison, high active share funds without a benchmark discrepancy tend to outperform, with a CPZ7 alpha of 1.28% per year (*t*-stat

= 2.14).[16] The difference of 1.21% per year is economically large and statistically significant ($t$-stat = −2.09), indicating that high active share funds without a benchmark discrepancy outperform high active share funds with a benchmark discrepancy even outside the context of benchmark-adjusted returns.

Results for the other quintiles of active share are similar to the full sample results, though, like the full sample results, they should be considered cautiously because of the strong positive correlation between *Benchmark Mismatch* and active share. In particular, the funds in the bottom four quintiles of active share that have a benchmark discrepancy tend to have a much higher average active share compared to the funds in the bottom four quintiles of active share that do not have a benchmark discrepancy.

### *6.3. Accounting for active share, past performance, and benchmark discrepancies*

Notably, the outperformance of high active share funds is concentrated among those funds without a benchmark discrepancy. If that group is further limited to those funds that are also in the top 20% of CPZ7 alpha during the prior year, performance again improves. In untabulated tests, we find that the funds in that top performing subgroup have a CPZ7 alpha of 2.18% per year ($t$-stat = 2.51).

In this section, we expand upon that finding by reexamining a key result from Cremers and Petajisto (2009): funds in the highest quintiles of both active share and benchmark-adjusted return over the prior year significantly outperform in the future. We test the impact of accounting for benchmark discrepancies on that result by sorting funds by prospectus active share, prospectus-

[16] Using the Fama-French four-factor model, this group of funds has a positive alpha, but the statistical significance varies depending on the dependent variable. Using prospectus-benchmark-adjusted returns, the four-factor alpha is 1.01% per year ($t$-stat = 2.10); however, using AS-benchmark-adjusted returns and excess returns, the four-factor alphas are 0.46% per year ($t$-stat = 1.03) and 0.63% per year ($t$-stat = 0.78). The decrease in alpha when using this model on these funds is consistent with the biases in the model documented in Cremers, Petajisto, and Zitzewitz (2012).

benchmark-adjusted return over the previous year, and whether the fund has a benchmark discrepancy (conditionally and in that order). Using the resulting groups, we then form equal weight portfolios and estimate annualized alphas using the CPZ7 model. We use active share and past performance relative to the prospectus benchmark to capture the groups investors would form if taking the prospectus benchmark at face value, but we evaluate portfolio performance using the CPZ7 model to generate a more accurate, less benchmark-dependent measure of subsequent alpha.

[Table 8 about here]

Table 8 shows the performance of the portfolios formed at each level of sorting. If investors focus on the prospectus benchmark and choose to buy only the funds in the highest quintiles of past performance and active share, they obtain an alpha of 2.31% per year (*t*-stat = 3.01). However, if those same investors are aware of the biases of prospectus benchmarks and drop the funds with a benchmark discrepancy from that group, the alpha increases to 3.21% per year (t-stat = 2.37).[17] In comparison, the funds with a benchmark discrepancy within the highest quintiles of past performance and active share have an average alpha of only 1.72% per year (*t*-stat = 1.97). While the difference in alpha between those two groups is economically large, it is not statistically significant at conventional levels (*t*-stat = 1.13), which is attributable, at least in part, to the relatively small number of funds within each group after sorting on three dimensions.

## 7. Comparison with the procedure in Sensoy (2009)

In sections 1 and 2, we discussed the details of the procedure used in Sensoy (2009) to identify benchmark discrepancies, as well as the main differences with respect to our procedure.

[17] Using the AS benchmark instead of the prospectus benchmark in the sorting process does not meaningfully change the alpha for this group (3.15% per year, *t*-stat = 2.70). Further, after switching the sorting benchmark, the alpha for this group is still the largest among all of the tested groups.

Here, we compare the two procedures.[18] We first look at the overlap between the procedures in terms of which funds have a benchmark discrepancy according to each procedure. We then consider the extent to which the choice of procedure affects performance evaluation, particularly for cases in which the two procedures disagree.

Sensoy finds that 31.2% of his sample has a benchmark discrepancy over the period 1994 to 2004. Among those funds, he also finds that the average $R^2$ of fund returns regressed on the returns of Sensoy's procedurally-determined benchmark is 82.6%, compared to 70.6% on the returns of the prospectus benchmark. Replicating the Sensoy procedure in our sample, which covers the period 1991 to 2015, we find that 23.1% of funds have a benchmark discrepancy with an average $R^2$ of 87.6% with respect to Sensoy's benchmark versus 80.2% with respect to the prospectus benchmark.

Figure 4 shows a broad comparison of the Sensoy procedure to the *Benchmark Mismatch* procedure we proposed in Section 4. The pie chart shows the commonality in fund-month observations identified as having a benchmark discrepancy using each procedure. In this analysis and subsequent comparisons, we only consider a benchmark discrepancy to exist by our procedure if *Benchmark Mismatch* is greater than 60%. The observations are at the fund-month level because whether a fund has a benchmark discrepancy is time-varying in our procedure, although it is time-invariant in Sensoy's.

[Figure 4 about here]

About 60% of fund-month observations do not have a benchmark discrepancy using either procedure. Among the remaining 40%, about 17% of observations only have a benchmark

[18] While do not exactly replicate Sensoy's results, we aim to follow his procedure very closely. The only difference between Sensoy's procedure and our replication is the set of benchmarks. Sensoy uses 12 benchmarks, but we use a larger set of 21 benchmarks as motivated in Section 3.3.

discrepancy according to our procedure, and another 14% only have a benchmark discrepancy according to Sensoy's procedure. Just 2% of observations have a benchmark discrepancy by both procedures that results in the same alternative benchmark, with another 7% having both procedures identify a benchmark discrepancy but assign different alternative benchmarks. All considered, the two procedures generate notably different conclusions about which funds have a benchmark discrepancy.

Given those differences, we next consider fund performance relative to the prospectus, AS, and Sensoy benchmarks, and the extent to which these differ. Table 9 shows the average benchmark-adjusted performance of funds conditional on whether each procedure identifies a benchmark discrepancy. Considered separately, both procedures find evidence of a performance boost for funds with a benchmark discrepancy, but our procedure finds a significantly larger performance boost compared to Sensoy's. Funds with a benchmark discrepancy according to our procedure have an average performance boost (i.e., AS benchmark return – prospectus benchmark return) of 1.52% per year (*t*-stat = 3.14), while funds with a benchmark discrepancy according to Sensoy's procedure have an average performance boost (i.e., Sensoy benchmark return – prospectus benchmark return) of 0.76% per year (*t*-stat = 1.64).

[Table 9 about here]

Furthermore, when our procedure identifies a benchmark discrepancy and Sensoy's does not, there is a large performance boost, but the reverse is not true. Funds identified by our procedure, but not by Sensoy's, have an average performance boost of 1.33% per year (*t*-stat = 2.25), while funds identified by Sensoy's procedure, but not by ours, have a performance boost of only 0.34% per year (*t*-stat = 0.82).[19] When both procedures agree there is a benchmark discrepancy, the funds have a large average performance boost relative to both alternative benchmarks. These results indicate that the benchmark discrepancies that our procedure identifies have a larger impact on fund performance evaluation.[20]

## 8. Differences in the systematic exposures of the prospectus benchmark and AS benchmark

The expected return on a passively managed index is driven solely by systematic exposures, as passive indices by construction have no alpha (arguably, see Cremers, Petajisto and Zitzewitz, 2012). Therefore, the significant difference in the average returns between prospectus and AS benchmarks among funds with a benchmark discrepancy (documented above) must logically arise out of differences in factor exposures. And since the average return on AS benchmarks is greater than the average return on prospectus benchmarks, the AS benchmarks should have greater net systematic factor exposures than the prospectus benchmarks. In this section, we first analyze differences in exposures between the two benchmarks for funds with a benchmark discrepancy. We then test whether the AS benchmark or the prospectus benchmark more accurately reflects the actual exposures of those funds.

We model the difference between the AS benchmark returns and the prospectus benchmark returns of funds with a benchmark discrepancy as:

[19] Elton, Gruber, and Blake (2014) use full sample fund-benchmark correlations alone to identify benchmark discrepancies for mutual funds. If we replicate their procedure, the results are similar to those presented in this section. Funds identified by our procedure, but not by theirs, have an average performance boost of 1.29% per year (*t*-stat = 1.78), while funds identified by their procedure, but not by ours, have a performance boost of only 0.38% per year (*t*-stat = 1.37).

[20] A potential alternative explanation for these results is that the AS benchmarks overstate risk for funds with a benchmark discrepancy according to our procedure. However, as we show in Section 8, the AS benchmarks for those funds accurately reflect both traditional and non-traditional factor exposures. In untabulated tests, we also find no evidence that the Sensoy benchmarks for funds with a benchmark discrepancy according to Sensoy's procedure systematically overstate or understate net systematic factor exposure. The difference in the performance results arises from the particular benchmark discrepancies identified by each procedure, not the accuracy of the alternative benchmarks selected.

$$Return_{AS,t} - Return_{pro,t} = \beta * Factor_t + \varepsilon_t \quad (7)$$

where $Return_{AS,t}$ is the annualized return on the AS benchmark averaged across all tested funds in month *t*, and $Return_{pro,t}$ is the annualized return on the prospectus benchmark averaged across all tested funds in month *t*. $Factor_t$ is a vector of factor returns in month *t*. The base model includes all the factors in the Cremers, Petajisto, and Zitzewitz (2012) seven-factor model.[21] We intentionally exclude a constant from the model because the difference in return between two benchmarks should be fully explainable by differences in systematic exposures alone (there should not be any alpha).[22] The model is estimated using all funds with a benchmark discrepancy (i.e., with *Benchmark Mismatch* greater than 60%) and across various investment style subgroups of those funds (as implied by the prospectus benchmark).

Table 10 shows the exposures to the CPZ7 factors. In this test, we consider the extent to which the traditional factors (market, size, value, and momentum) considered in the CPZ7 model explain the average difference in returns between the prospectus and AS benchmarks. As a reference, the first set of rows reports the average annualized benchmark-adjusted returns using both the prospectus and AS benchmarks for each group of funds. This shows the economic magnitude of the performance boost within each group. The second set of rows then reports the estimated coefficients associated with the factors. In the third set of rows, we report the $R^2$ from the regression and the sum of the products of the estimated factor exposures and annualized factor returns. The "Total Factor Return" row can be compared to the "Difference" row to determine how

[21] In a few instances, the CPZ7 model explains all the variation in returns between a fund's AS and prospectus benchmark because the model's index-based factors correspond to those two benchmarks. For example, a fund with an S&P 500 prospectus benchmark and a Russell Midcap AS benchmark will have a difference in returns that is fully explained by the *RMS5* factor. The small number of funds whose two benchmarks correspond to a CPZ7 factor are dropped in these tests.
[22] Although, our results are similar if a constant is included in the model.

much of the average difference in the returns between the benchmarks can be explained by differences in traditional factor exposures.

[Table 10 about here]

Using all funds, the factors explain 0.57% (*t*-stat = 2.65) of the 1.50% per year difference in average returns between the prospectus and AS benchmarks. The primary differences relate to the two size factors. The coefficient associated with *RMS5* (the difference in returns between the Russell Midcap index and the S&P 500 index) is positive, which indicates the prospectus benchmark has a lower exposure to the mid-cap factor than the AS benchmark. However, the reverse is true of the coefficient associated with *R2RM* (the difference in returns between the Russell 2000 index and the Russell Midcap index). Overall, only 38% (=0.57%/1.50%) of the average difference in returns between the benchmarks is explained by the traditional factors included in the CPZ7 model.

Looking across the style subgroups, there is substantial variation. Differences in traditional factor exposures explain 1.60% (*t*-stat = 1.64) of the 2.38% per year difference in returns for large-cap funds. Most of this difference comes from the prospectus benchmark having lower small- and mid-cap exposure compared to the AS benchmark. In other words, large-cap funds with a benchmark discrepancy tend to own smaller cap stocks than their prospectus benchmarks indicate.

In comparison, funds with a small- or mid-cap style have a smaller return difference of 1.12% per year to explain, and differences in traditional factor exposures explain only 0.41% (*t*-stat = 1.10) of that return difference. The AS benchmarks of small- and mid-cap funds have lower small-cap exposures than their prospectus benchmarks. Interestingly, the AS benchmarks of both large-cap and small-/mid-cap funds lean more towards the center of the size distribution than their prospectus benchmarks, which increases the average difference in returns for large-cap funds and decreases it for small-/mid-cap funds. A similar lean towards the center of the distribution can be seen among growth and value funds (e.g.., the coefficients on *S5VS5G*, *RMVRMG*, and *R2VR2G*). As a result, there is no statistically significant difference in returns between the prospectus and AS benchmarks for value funds with a benchmark discrepancy.

The CPZ7 model explains only a small portion of the difference in returns between the benchmarks. This indicates the prospectus and AS benchmarks should vary along dimensions that are unrelated to the traditional factors. Hence, we next consider some “non-traditional” factors. Cochrane (2011) remarks that there is now “a zoo of new factors” in the academic literature that purport to explain the cross-section of returns. Rather than attempt to test all potential factors, we consider the explanatory power of a subset of non-traditional factors which have either received a particularly large amount of attention or been shown to be particularly robust (in, e.g., Feng, Giglio, and Xiu, 2017).

The non-traditional factors we consider are: the Fama and French (2015) profitability (*RMW*) and investment (*CMA*) factors; the Stambaugh and Yuan (2017) management (*MGMT*) and performance (*PERF*) factors; the Frazzini and Pedersen (2014) betting against beta (*BAB*) factor; the Asness, Frazzini, and Pedersen (2017) quality-minus-junk (*QMJ*) factor; and the Pastor and Stambaugh (2003) traded liquidity (*LIQ*) factor. These factors are added as a group to our base CPZ7 model, and the previous analysis is repeated.

Table 11 shows that once these non-traditional factors are included in the model, the difference in prospectus and AS benchmark returns that can be explained considerably increases (and is statistically significant for each group of funds considered). The total factor returns now captures most of the average difference in returns between the prospectus and AS benchmarks. Looking at the ‘All’ funds group, 1.31% (*t*-stat = 4.48) of the 1.50% difference in return is explained by the expanded model. The $R^2$ almost doubles from 27.7% using the CPZ7 factors alone to 52.4% in the expanded model.

[Table 11 about here]

The factor that most consistently adds new explanatory power is the profitability factor *RMW*. In all tested groups, the prospectus benchmark has a lower *RMW* exposure than the AS benchmark. This result indicates that, on average, funds with a benchmark discrepancy tend to have an AS benchmark that invests in more profitable companies compared to the prospectus benchmark. The economic impact of this difference is large. The difference in *RMW* exposure alone adds 0.57% per year to the total factor return within the 'All' funds group. Considering subgroups, the quality-minus-junk factor, *QMJ*, and management factor, *MGMT,* also have some economically large and statistically significant explanatory power.

While these results show that the AS benchmarks of funds with a benchmark discrepancy have greater average net systematic factor exposures compared to those funds' prospectus benchmarks, it is possible that the AS benchmarks overstate funds' net exposures, instead of the prospectus benchmarks understating net exposures. To evaluate this possibility, we note that if a benchmark effectively captures a fund's net systematic factor exposure, then the benchmark-adjusted return should be the same as the estimated alpha that results from regressing that return against various factors. Conversely, if a benchmark understates (overstates) net exposure, then the benchmark-adjusted return should be greater (less) than the estimated alpha. We consider whether the prospectus benchmarks or AS benchmarks better reflect funds' net systematic factor exposures using the following model:

$$Return_{fund,t} - Return_{bench,t} = \alpha + \beta * Factor_t + \varepsilon_t \tag{8}$$

where $Return_{fund,t}$ is the average annualized net return across all tested funds in month $t$ and $Return_{bench,t}$ is the average annualized return on those funds' benchmarks in month $t$. We consider both the prospectus and AS benchmarks in our analysis. $Factor_t$ is a vector of factor returns in month *t*. Depending on the specification, it includes either no factors, the CPZ7 factors, or the CPZ7 factors along with the non-traditional factors from Table 11. We include a constant in this model since the difference between a fund's return and its benchmark's return should reflect the fund's alpha. The model is estimated using all funds with a benchmark discrepancy (i.e., *Benchmark Mismatch* greater than 60%) and for the subgroup of those funds with a large-cap style. That subgroup is of particular interest as our previous tests indicate that those funds have the largest difference in average returns between their prospectus and AS benchmarks.

[Table 12 about here]

The results are presented in Table 12. Using all funds with a benchmark discrepancy in Panel A, the prospectus-benchmark-adjusted returns are unaffected by the traditional factors (CPZ7). Adjusting for those factor exposures reduces the abnormal return by only 0.15% per year (*t*-stat = −0.33). However, when considering non-traditional factors, the prospectus benchmark significantly understates net systematic factor exposure. The abnormal return based on prospectus-benchmark-adjusted returns decreases from 0.65% per year using no factors to −0.29% per year after adding both the traditional and non-traditional factors (CPZ7+) to the model. That change in performance of −0.94% per year is statistically significant (*t*-stat = −2.37) and indicates that the prospectus benchmark sets a lower bar for the fund than is appropriate given the fund's net systematic exposure.

In comparison, the factors matter less when using the AS-benchmark-adjusted returns of funds with a benchmark discrepancy. The abnormal return does not change significantly even after including the non-traditional factors in the model. The change when switching from no factors to all factors is 0.57% per year, which is statistically insignificant (*t*-stat = 1.44) at conventional levels. This result indicates that if fund performance is adjusted using the AS benchmark, only relatively minor net systematic exposure remains.

The above outcomes are magnified if we focus on just large-cap funds with a benchmark discrepancy. The prospectus-benchmark-adjusted return decreases by 1.18% per year (*t*-stat = −1.92) after adjusting for the CPZ7 factors, which shows those funds' prospectus benchmarks significantly understate net systematic exposures to traditional factors. After including the non-traditional factors in the model that decrease becomes 1.94% per year (*t*-stat = −3.05), so net exposures to the non-traditional factors is also understated by those funds' prospectus benchmarks. In comparison, the AS-benchmark-adjusted returns are unaffected by both the traditional and non-traditional factors. Among funds with a benchmark discrepancy and a prospectus benchmark implying a large-cap style, adjusting performance using the AS benchmark leaves little net systematic exposure, regardless of whether an investor considers non-traditional factors.

In the big picture, funds with benchmark discrepancies perform better relative to their prospectus benchmarks because those benchmarks tend to understate those funds' net systematic factor exposures. Given that the AS benchmarks of those funds on average neither over nor understate those funds' net exposures, using the AS benchmark whenever a fund has a benchmark discrepancy can be expected to result in a more accurate measure of fund performance compared to using the prospectus benchmark.

## 9. Funds flows and the prospectus benchmark

The economic importance of benchmark discrepancies, in large part, depends on the extent to which investors actually rely on the fund's performance relative to prospectus benchmark when evaluating performance. On the one hand, if investors can identify which funds have benchmark discrepancies and ignore the prospectus benchmarks of those funds, then the performance boost from the benchmark discrepancy should have no impact on the competition between funds for capital. On the other hand, if investors cannot identify benchmark discrepancies or fail to fully discount the prospectus benchmark when they do identify a discrepancy, then the competition for capital between funds will be affected.

In this section, we examine investors' net flows to funds to determine the degree to which performance relative to the prospectus benchmark and benchmark discrepancies affect investor decisions. We model the relation between a fund's net flows and the past performance of a fund as follows:

$$Flow_{i,t} = \theta + \beta * Performance_{i,t} + \gamma * Mismatch_{i,t} + \delta * Chars_{i,t} + FE + \varepsilon_{i,t} \qquad (9)$$

where $Flow_{i,t}$ is the percentage implied net flow for fund *i* in month *t*.[23] $Performance_{i,t}$ is a vector of information about fund *i*'s performance over the year ending at the start of month *t*. It includes the difference between fund *i*'s return and the return on fund *i*'s AS benchmark, the difference between the return on fund *i*'s AS benchmark and prospectus benchmark (i.e., the performance boost), and fund *i*'s annualized CAPM alpha.[24] In some instances, we use actual fund returns. In other instances, the returns are ranked at the start of each month and scaled from zero to one. Using the ranked returns allows for a more natural test of a potential non-linear relation between measures of fund performance and flows (see, e.g., Sirri and Tufano, 1998).

[23] The calculation of implied net flows assumes that all inflows and outflows occur at the end of the month. That assumption is obviously incorrect. However, Clifford, Fulkerson, Jordan, and Waldman (2013) find that the implied net flows have a correlation of 0.996 with the actual net flows calculated from funds' filings of the SEC's Form N-SAR.

[24] Our conclusions are the same if alternative measures of alpha (e.g., Fama-French four-factor or CPZ7) are used.

$Mismatch_{i,t}$ is the *Benchmark Mismatch* for fund *i* as of the start of month *t*, and $Chars_{i,t}$ is a vector of characteristics for fund *i* available as of the start of month *t*. It contains the same characteristics as in Eq. (5). $FE$ represents style and year-month fixed effects. The model is estimated using the sample of fund-months that have different prospectus and AS benchmarks (i.e., *Benchmark Mismatch* > 0).[25]

Table 13 shows estimates of this model. In the first three columns, we consider whether fund flows depend on performance relative to the prospectus benchmark. If they do, then an increase in the performance boost should increase net flows. After controlling for performance relative to the AS benchmark, a 1% increase in the performance boost (i.e., the difference in return between the AS benchmark and the prospectus benchmark) increases net flows by 0.07% per month (0.84% annualized, *t*-stat = 14.03). That effect is about half the effect of a 1% increase in performance relative to the AS benchmark, and thus seems economically meaningful. Controlling for a fund's CAPM alpha further lessens, but does not eliminate, the effect of the performance boost on flows. While investors are influenced by other measures of performance, these results indicate that performance relative to the prospectus benchmark is an important determinant of investors' capital allocation choices.[26]

[Table 13 about here]

As the level of *Benchmark Mismatch* increases, we would expect the importance of the prospectus benchmark to decrease. In the fourth column, we find that as *Benchmark Mismatch* increases fund flows are less sensitive to the performance boost. Every 10% increase in *Benchmark Mismatch* reduces the impact of a 1% performance boost on net flows by 0.01% (*t*-stat = −5.64).

---

[25] Results are similar using the full sample of fund-months.

[26] While Del Guercio and Reuter (2014) and Barber, Huang, and Odean (2016) both suggest that the net flows of more sophisticated investors are less responsive to simple measures of performance, we find similar results regardless of a fund's level of institutional ownership or whether a fund is likely to be sold through a broker.

Changes in *Benchmark Mismatch* have no impact on the weight investors give to performance relative to the AS benchmark. While these results indicate some level of sophistication on the part of investors, even at the maximum *Benchmark Mismatch* of 100%, the performance boost still has an economically large and statistically significant impact on flows.

The final three columns consider non-linearity in the relationship between fund flows and performance. We expect benchmark discrepancies to be particularly salient to investors when the performance boost is relatively large, and expect that a fund that beats its prospectus benchmark by 10% should receive more scrutiny than a fund that beats its prospectus benchmark by 1%. While net flows are convex with respect to performance relative to the AS benchmark, we find they are concave with respect to the performance boost. Further, as indicated by Clifford, Jordan, and Riley (2014), the flow-performance relation is linear for performance relative to the AS benchmark for large funds (i.e., top 20% in total net assets), but it remains concave regardless of fund size with respect to the performance boost. These results are consistent with our hypothesis.

## 10. Conclusion

Risk-adjustment is central to performance evaluation. To facilitate that process, mutual funds are legally required to provide a benchmark to investors in the fund prospectus. Given that funds rarely change their prospectus benchmark and market themselves in a competitive environment to investors that often have limited sophistication, we might expect funds to respond strategically when constructing their portfolios. While most funds appear to have a risk-appropriate prospectus benchmark, we find that a substantial portion of funds have a prospectus benchmark that understates risk and, consequently, overstates relative performance. Funds benefit from that overstatement, as investor flows respond to performance relative to the prospectus benchmark even when a fund has a benchmark discrepancy. In general, researchers and investors should exercise significant caution when using prospectus benchmarks to evaluate fund performance, although using the prospectus benchmarks to diagnose benchmark discrepancies can help in the process of identifying skilled fund managers.

Our results contribute to several topics in the literature. First, they suggest researchers should be careful when choosing benchmarks for the analysis of performance. Benchmark-adjusted returns are common in studies of mutual funds (e.g., Pastor, Stambaugh, and Taylor, 2017; Cremers, Ferreira, Matos, and Starks, 2016; Berk and van Binsbergen, 2015; Angelidis, Giamouridis, and Tessaromatis, 2013). If such studies use the prospectus benchmark, there can be substantial noise and biases in the results. In the majority of academic studies, researchers using benchmark-adjusted returns assign their own benchmark or rely on benchmark providers such as Morningstar, so we do not expect the current bias in the literature to be large.[27]

Second, despite often failing to match the fund's portfolio, our paper demonstrates the importance of prospectus benchmarks for funds. Barber, Huang, and Odean (2016) and Berk and van Binsbergen (2016) both indicate that performance relative to the Sharpe (1964) and Lintner (1965) capital asset pricing model (CAPM) best explains the flow-performance relation, but neither study considers (prospectus or AS) benchmark-adjusted returns. We show that the impact on flows of benchmark-adjusted returns in general, and the prospectus-benchmark-adjusted returns specifically, are economically meaningful even after accounting for CAPM alpha.

Third, our fund flow results add to a growing number of studies that show how fund investors respond to information that appears to be of questionable economic value. Cooper, Gulen, and Rau (2005) find that funds that change their name to align with popular investment

[27] Angelidis, Giamouridis, and Tessaromatis (2013) conclude that skill can be misstated when funds' self-declared benchmarks are ignored; however, they use inferred self-declared benchmarks, rather than the actual prospectus benchmarks, in their analysis. Hence, their conclusions are more about the value of accounting for self-declared style than the biases of prospectus benchmarks.

styles receive larger flows, regardless of whether the name change reflects a change in the fund's portfolio. Jain and Wu (2000) show that funds that advertise their strong past performance in *Barron's* or *Money* magazine receive larger flows than comparable funds that do not advertise, even though the two groups have the same subsequent performance. Kaniel and Parham (2017) demonstrate that funds that just make the cut-off to qualify for *Wall Street Journal* lists receive substantially larger flows in periods when those lists are labeled as "Category Kings." Solomon, Soltes, and Sosyura (2014) find media coverage of the stocks held by funds has a strong influence on subsequent fund flows despite that coverage having no relation with subsequent performance. In a similar fashion, we show that a fund that overstates its performance by using an inaccurate prospectus benchmark will receive substantially larger flows compared to an equivalent fund with an accurate prospectus benchmark, even when the magnitude of the inaccuracy is large.

Fourth, our comparison of different benchmark-adjusted returns adds to the debate on the appropriate factor structure for evaluating fund performance. It is common to control for market risk and the size and value factors (e.g., the Fama and French (1993) model and Cremers, Petajisto, and Zitzewitz (2012) model), but Harvey, Liu, and Zhu (2016) and Hou, Xue, and Zhang (2017) find that there are hundreds of apparent pricing anomalies that could be used to form pricing factors. Whether the use of any or all of those other factors is appropriate in the evaluation of mutual fund performance remains unclear. If these non-traditional factors are not directly investable, then providing exposure to them through active management could represent a value-added activity.[28] Our results make evident that mutual funds often have exposures to non-traditional factors that are not indicated by their prospectus benchmark and do impact the fund's

[28] While "Smart Beta" strategies designed to give investors exposure to non-traditional factors are popular today, they did not exist during much of our period of study.

performance. However, given that the AS benchmarks, which are directly investable at low cost, generally have similar exposures to the non-traditional factors as the funds, the additional performance arising from those exposures should not represent alpha from an investor's prospective.

Finally, our results contribute to the debate on mutual fund manager skill. On the one hand, studies including Carhart (1997) and Fama and French (2010) find little evidence of skill. On the other hand, studies such as Kosowski, Timmermann, Wermers, and White (2006), Barras, Scaillet, and Wermers (2010), and Berk and van Binsbergen (2015) find material evidence of skill. Our results support the existence of significant investment skill for at least a subset of funds. Like Cremers and Petajisto (2009), we show that funds with high active share have a positive alpha, but like Petajisto (2013) and Cremers and Pareek (2016), we also show that the outperformance of high active share funds is concentrated within a sub-group—high active share funds without a benchmark discrepancy.

## References


Agarwal, Vikas, Clifton Green, and Honglin Ren. 2018. Alpha or beta in the eye of the beholder: What drives hedge fund flows?. *Journal of Financial Economics* 127, 417-434.

Angelidis, Timotheos, Daniel Giamouridis, and Nikolaos Tessaromatis. 2013. Revisiting mutual fund performance evaluation. *Journal of Banking & Finance* 37, 1759-1776.

Asness, Clifford, Andrea Frazzini, and Lasse Pedersen. 2017. Quality minus junk. Working paper.

Bams, Dennis, Roger Otten, and Ehsan Ramezanifar. 2017. Investment style misclassification and mutual fund performance. Working paper.

Barber, Brad, Xing Huang, and Terrance Odean. 2016. Which factors matter to investors? Evidence from mutual fund flows. *Review of Financial Studies* 29, 2600-2642.

Barras, Laurent, Olivier Scaillet, and Russ Wermers. 2010. False discoveries in mutual fund performance: Measuring luck in estimated alphas. *Journal of Finance* 65, 179-216.

Berk, Jonathan, and Jules van Binsbergen. 2015. Measuring skill in the mutual fund industry. *Journal of Financial Economics* 118, 1-20.

Berk, Jonathan, and Jules van Binsbergen. 2016. Assessing asset pricing models using revealed preference. *Journal of Financial Economics* 119, 1-23.

Brown, Keith, Van Harlow, and Hanjiang Zhang. 2009. Staying the course: The role of investment style consistency in the performance of mutual funds. Working paper.

Brown, Keith, Van Harlow, and Laura Starks. 1996. Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. *Journal of Finance* 51, 85-110.

Cao, Charles, Peter Iliev, and Raisi Velthuis. 2017. Style drift: Evidence from small-cap mutual funds. *Journal of Banking & Finance* 78, 42-57.

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

Chan, Louis, Hsiu-Lang Chen, and Josef Lakonishok. 2002. On mutual fund investment styles. *Review of Financial Studies* 15, 1407-1437.

Clifford, Christopher, Bradford Jordan, and Timothy Riley. 2014. Average fund versus average dollars: Implications for mutual fund research. *Journal of Empirical Finance* 28, 249-260.

Clifford, Christopher, Jon Fulkerson, Bradford Jordan, and Steve Waldman. 2013. Risk and fund flows. Working paper.

Cochrane, John. 2011. Presidential address: Discount rates. *Journal of Finance* 66, 1047-1108.

Cooper, Michael, Huseyin Gulen, and Raghavendra Rau. 2005. Changing names with style: Mutual fund name changes and their effects on fund flows. *Journal of Finance* 60, 2825-2858.

Cremers, Martijn. 2017. Active share and the three pillars of active management: Skill, conviction, and opportunity. *Financial Analysts Journal* 73, 61-79.

Cremers, Martijn, and Ankur Pareek. 2016. Patient capital outperformance: The investment skill of high active share managers who trade infrequently. *Journal of Financial Economics* 122, 288-306.

Cremers, Martijn, and Antti Petajisto. 2009. How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies* 22, 3329-3365.

Cremers, Martijn, and Quinn Curtis. 2016. Do mutual fund investors get what they pay for? The legal consequences of closet index funds. *Virginia Law & Business Review* 11, 31-93.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz. 2012. Should benchmark indices have alpha? Revisiting performance evaluation. *Critical Finance Review* 2, 1-48.

Cremers, Martijn, Miguel Ferreira, Pedro Matos, and Laura Starks. 2016. Indexing and active fund management: International evidence. *Journal of Financial Economics* 120, 539-560.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52, 1035-1058.

Del Guercio, Diane, and Jonathan Reuter. 2014. Mutual fund performance and the incentive to generate alpha. *Journal of Finance* 69, 1673-1704.

DiBartolomeo, Dan, and Erik Witkowski. 1997. Mutual fund misclassification: Evidence based on style analysis. *Financial Analysts Journal* 53, 32-43.

Elton, Edwin, Martin Gruber, and Christopher Blake. 2003. Incentive fees and mutual funds. *Journal of Finance* 58, 779-804.

Elton, Edwin, Martin Gruber, and Christopher Blake. 2014. The performance of separate accounts and collective investment trusts. *Review of Finance* 18, 1717-1742.

Evans, Richard. 2010. Mutual fund incubation. *Journal of Finance* 65, 1581-1611.

Fama, Eugene, and Kenneth French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3-56.

Fama, Eugene, and Kenneth French. 2010. Luck versus skill in the cross-section of mutual fund returns. *Journal of Finance* 65, 1915-1947.

Fama, Eugene, and Kenneth French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116, 1-22.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu. 2017. Taming the factor zoo. Working paper.

Frazzini, Andrea, and Lasse Pedersen. 2014. Betting against beta. *Journal of Financial Economics* 111, 1-25.

Harvey, Campbell, Yan Liu, and Heqing Zhu. 2016. …And the cross-section of expected returns. *Review of Financial Studies* 29, 5-68.

Hirt, Joshua, Ravi Tolani, and Christopher Philips. 2015. Global equity benchmarks: Are prospectus benchmarks the correct barometer?. Vanguard Research.

Hou, Kewei, Chen Xue, and Lu Zhang. 2017. Replicating anomalies. Working paper.

Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang. 2011. Risk shifting and mutual fund performance. *Review of Financial Studies* 24, 2575-2616.

Jain, Prem, and Joanna Wu. 2000. Truth in mutual fund advertising: Evidence on future performance and fund flows. *Journal of Finance* 55, 937-958.

Kaniel, Ron, and Robert Parham. 2017. WSJ Category Kings – The impact of media attention on consumer and mutual fund investment decisions. *Journal of Financial Economics* 123, 337-356.

Kim, Moon, Ravi Shukla, and Michael Tomas. 2000. Mutual fund objective misclassification. *Journal of Economics and Business* 52, 309-323.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White. 2006. Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis. *Journal of Finance* 61, 2551-2596.

Lee, Jung, Charles Trzcinka, and Shyam Venkatesan. 2017. Mutual fund risk-shifting and management contracts. Working paper.

Lintner, John. 1965. The valuation of risk assets and the selection of risk investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47, 13-37.

Mateus, Irina, Cesario Mateus, and Natasa Todorovic. 2017. The impact of benchmark choice on US mutual fund benchmark-adjusted performance and ranking. Working paper.

Morningstar. 2004. Fact Sheet: The Morningstar Style Box. White paper available at https://corporate.morningstar.com/US/documents/MethodologyDocuments/FactSheets/MorningstarStyleBox_FactSheet_.pdf

Pastor, Lubos, and Robert Stambaugh. 2003. Liquidity risk and expected stock returns. *Journal of Political Economy* 111, 642-685.

Pastor, Lubos, Robert Stambaugh, and Lucian Taylor. 2017. Do funds make more when they trade more?. *Journal of Finance* 72, 1483-1528.

Petajisto, Antti. 2013. Active share and mutual fund performance. *Financial Analysts Journal* 69, 73-93.

Schwarz, Christopher, and Mark Potter. 2016. Revisiting mutual fund portfolio disclosure. *Review of Financial Studies* 29, 3519-3544.

Sensoy, Berk. 2009. Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics* 92, 25-39.

Sharpe, William. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19, 425-442.

Sharpe, William. 1992. Asset allocation: Management style and performance measurement. *Journal of Portfolio Management* 18, 7-19.

Sirri, Erik, and Peter Tufano. 1998. Costly search and mutual fund flows. *Journal of Finance* 53, 1589-1622.

Solomon, David, Eugene Soltes, and Denis Sosyura. 2014. Winners in the spotlight: Media coverage of fund holdings as a driver of flows. *Journal of Financial Economics* 113, 53-72.

Stambaugh, Robert, and Yu Yuan. 2017. Mispricing factors. *Review of Financial Studies* 30, 1270-1315.

**Figure 1: Percentage of funds with different prospectus and AS benchmarks by quarter**

This figure shows (1) the percentage of funds each quarter with a prospectus benchmark different from their AS benchmark and (2) the percentage of funds each quarter with a prospectus benchmark different from their AS benchmark and a *Benchmark Mismatch* of greater than 60%.

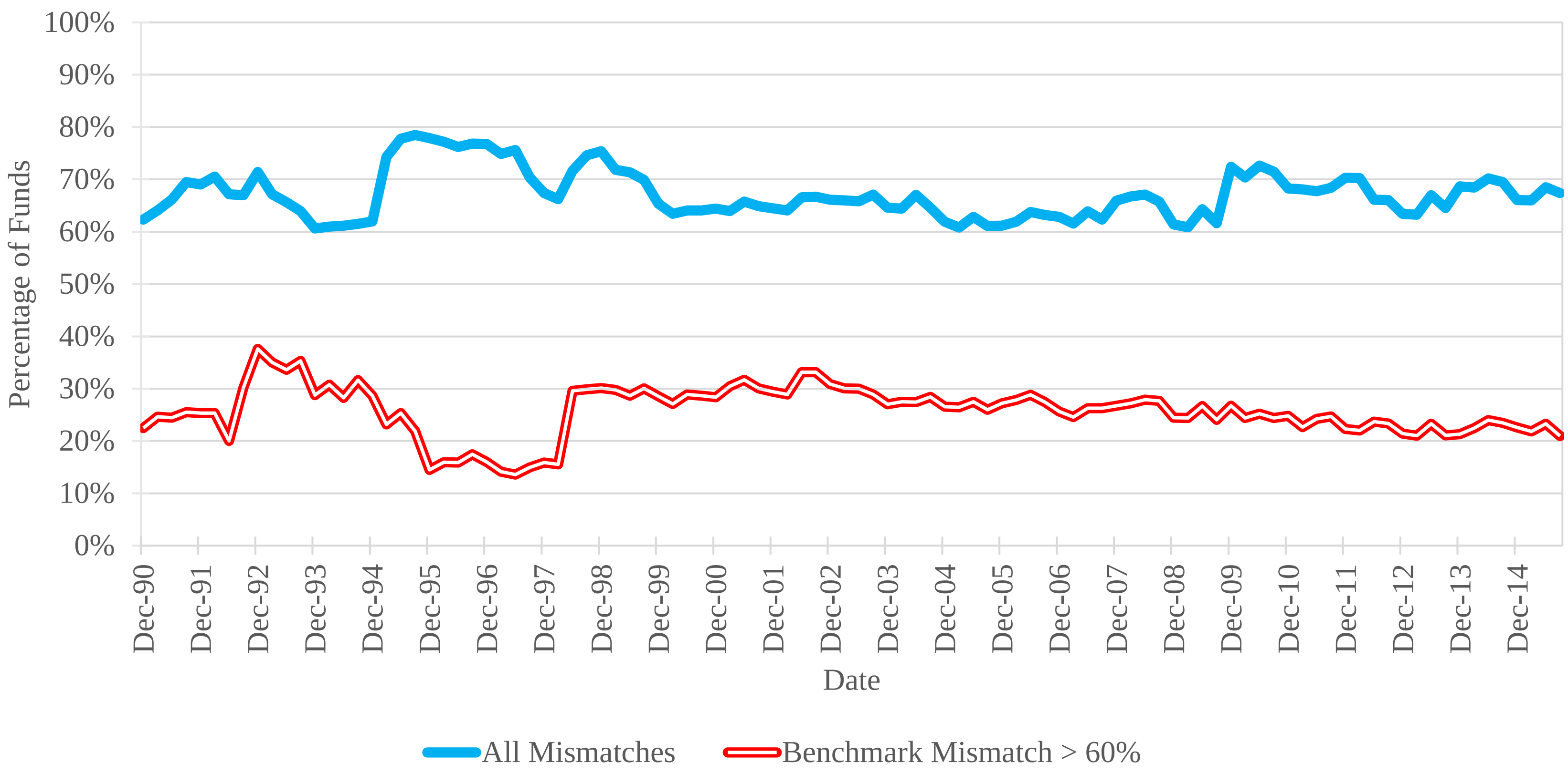

**Figure 2: CDF of *Active Gap***

This figure shows a cumulative density function of *Active Gap* for all fund-month observations where the prospectus benchmark does not match the AS benchmark.

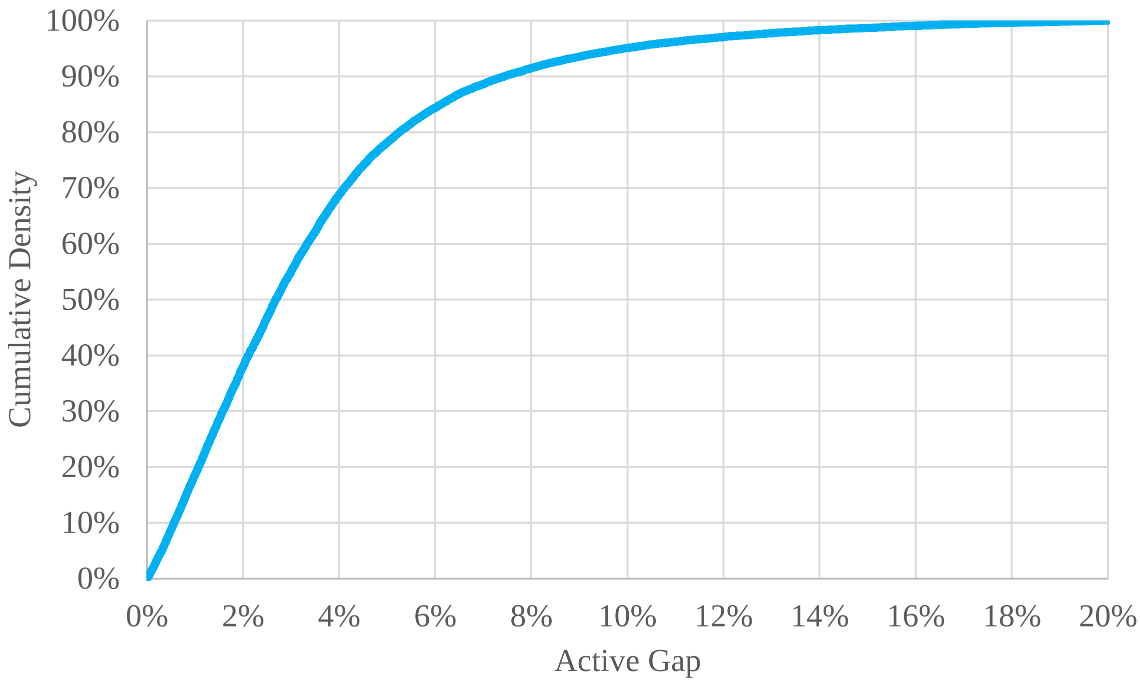

**Figure 3: CDF of *Benchmark Mismatch***

This figure shows a cumulative density function of *Benchmark Mismatch* for all fund-month observations where the prospectus benchmark does not match the AS benchmark.

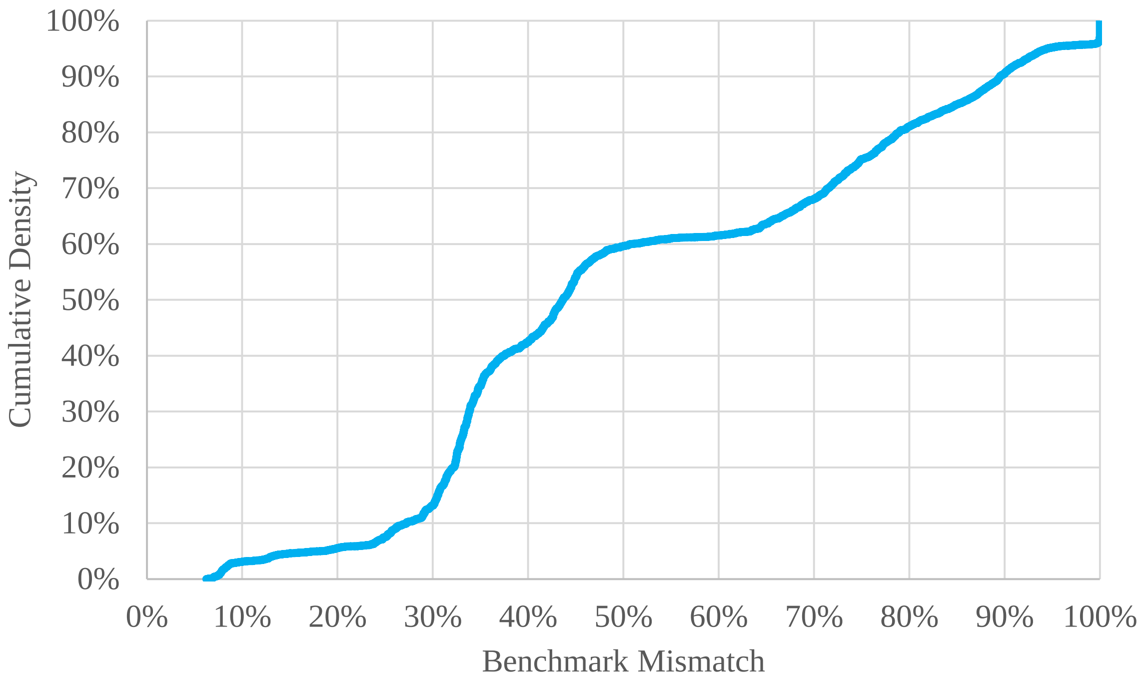

**Figure 4: Overlap between benchmark discrepancy identification procedures**

This figure shows the percentage of the full sample of fund-months identified as having a benchmark discrepancy following two different procedures. In the BM procedure, a fund is considered to have a benchmark discrepancy if our *Benchmark Mismatch* measure is greater than 60%. In the Sensoy (2009) procedure, a fund is considered to have a benchmark discrepancy if the Morningstar style boxes and fund-benchmark correlations indicate a more appropriate benchmark. If the figure legend reads "BM = Yes", then there is mismatch using the BM procedure. If the figure legend reads "Sensoy = Yes", then there is mismatch using the Sensoy procedure When both procedures identify a discrepancy, the benchmark identified as the more appropriate benchmark compared to the prospectus benchmark is sometimes the same across procedures and other times different. We consider those groups separately.

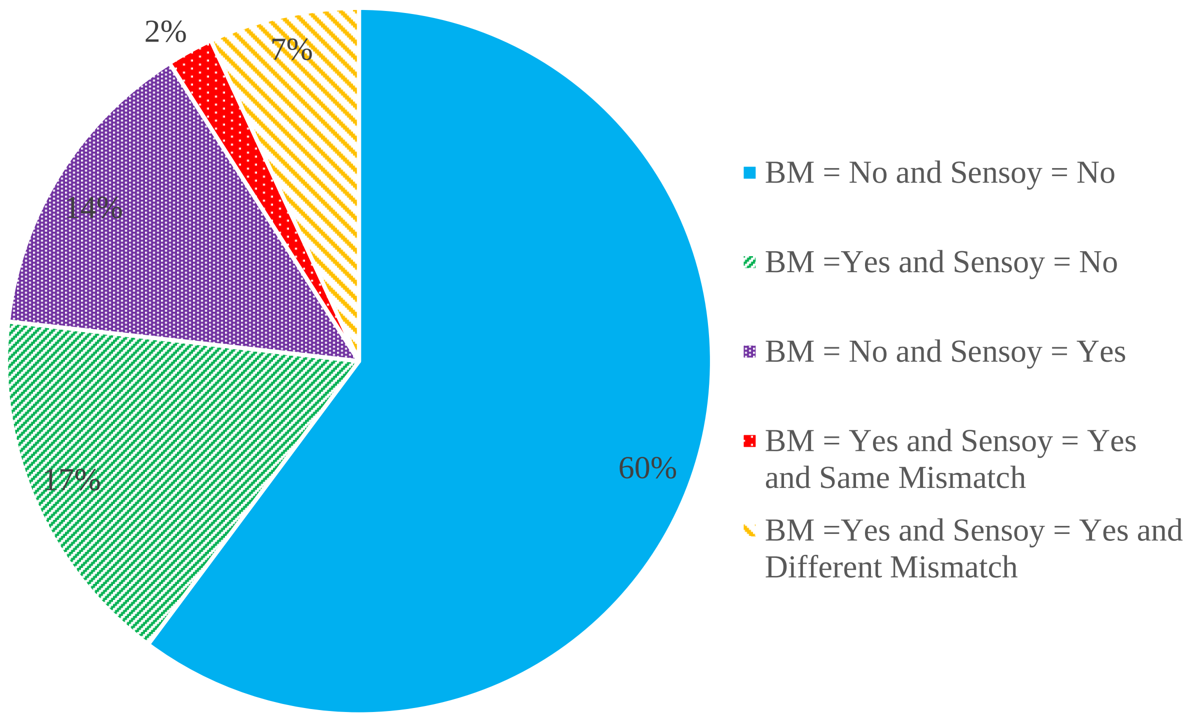

**Table 1: Full sample summary statistics**

This table shows basic summary statistics for the full sample of fund-month observations. Any Mismatch is dummy variable equal to one if the prospectus benchmark and AS benchmark are different. Large Mismatch is a dummy variable equal to one if the prospectus benchmark and AS benchmark are different and *Benchmark Mismatch* is greater than 60%. Prospectus Active Share is active share of the fund relative to the benchmark listed in the fund's prospectus. Minimum Active Share is the lowest active share of the fund across all tested benchmarks. *Benchmark Mismatch* is the active share of the fund's prospectus benchmark relative to its AS benchmark. *Active Gap* is the difference between the fund's prospectus active share and minimum active share. Prospectus Adjusted Return is the fund's annualized monthly return less the annualized monthly return on the fund's prospectus benchmark. Minimum AS Adjusted Return is the fund's annualized monthly return less the annualized monthly return on the fund's AS benchmark. P25, P50, and P75 are the 25th, 50th, and 75th percentiles, respectively. $\rho_{t,t-12}$ is the correlation between the fund's value in month $t$ and month $t-12$.

| | Mean | Standard Deviation | P25 | P50 | P75 | $\rho_{t,t-12}$ |
|---|---|---|---|---|---|---|
| Any Mismatch | 0.67 | 0.47 | 0.00 | 1.00 | 1.00 | 0.59 |
| Large Mismatch | 0.26 | 0.43 | 0.00 | 0.00 | 1.00 | 0.75 |
| Prospectus Active Share | 80.7% | 14.1% | 71.1% | 83.5% | 92.5% | 0.93 |
| Minimum Active Share | 78.4% | 13.8% | 68.9% | 81.0% | 89.7% | 0.92 |
| Benchmark Mismatch | 34.9% | 32.1% | 0.0% | 33.0% | 63.5% | 0.72 |
| Active Gap | 2.3% | 3.1% | 0.0% | 1.3% | 3.5% | 0.75 |
| Prospectus Adjusted Return | -0.33% | 18.90% | -10.50% | -0.57% | 9.52% | 0.02 |
| Minimum AS Adjusted Return | -0.78% | 18.48% | -10.99% | -0.88% | 9.20% | 0.01 |

**Table 2: Most common differences between the prospectus and AS benchmarks**

This table shows the five most common differences between the prospectus benchmark and the AS benchmark. Panel A shows the most common differences for all fund-months with a *Benchmark Mismatch* greater than zero and less than 60%. Panel B shows the most common differences for all fund-months with a *Benchmark Mismatch* greater than 60%. For each difference listed, the percentage of that sample with that difference is reported. The median *Active Gap* for fund-months with that difference and the average *Benchmark Mismatch* for that difference are also provided.

Panel A: 0% < Benchmark Mismatch < 60%

| Prospectus Benchmark | AS Benchmark | Percentage of Differences | Median Active Gap | Benchmark Mismatch |
|---|---|---|---|---|
| S&P 500 | S&P 500 Growth | 19.5% | 3.1% | 33.0% |
| Russell 1000 Growth | S&P 500 Growth | 14.4% | 1.8% | 30.2% |
| Russell 1000 Value | S&P 500 Value | 9.6% | 2.1% | 32.7% |
| S&P 500 | Russell 1000 Growth | 8.0% | 3.4% | 43.4% |
| S&P 500 | S&P 500 Value | 7.3% | 1.9% | 35.8% |

Panel B: Benchmark Mismatch > 60%

| Prospectus Benchmark | AS Benchmark | Percentage of Differences | Median Active Gap | Benchmark Mismatch |
|---|---|---|---|---|
| Russell 2000 | S&P 600 Growth | 11.6% | 3.9% | 77.1% |
| Russell 2000 Value | S&P 600 Value | 9.6% | 2.0% | 68.6% |
| Russell 2000 Growth | S&P 600 Growth | 9.0% | 1.7% | 69.0% |
| Russell 2000 | S&P 600 Value | 5.7% | 2.6% | 75.6% |
| S&P 500 | Russell Midcap Growth | 4.4% | 6.7% | 90.3% |

**Table 3: Characteristics of funds conditional on *Benchmark Mismatch***

This table compares the characteristics of fund-months with a *Benchmark Mismatch* (BM) greater than 60% to funds with a *Benchmark Mismatch* less than 60% (including funds with a *Benchmark Mismatch* of zero). Panel A reports basic fund characteristics. Prospectus Active Share is the active share of the fund relative to the benchmark listed in the fund's prospectus. Assets is the net assets of the fund in billions of dollars. Age is the age of the oldest share class of the fund and is reported in years. Expense Ratio and Turnover Ratio are the annual expense and turnover ratios as reported by the fund. Number of Holdings is the number of common equity positions held by the fund. Institutional is the percentage of the fund's net assets that is held within institutional share classes. Panel B reports the percentage of funds within each group that have a given fund style. The styles are determined based on the prospectus benchmark. Each fund is identified as either large-cap or small-/mid-cap and one of either growth, blend, or value. The *t*-statistics for the differences are calculated using standard errors clustered by fund and year-month.

Panel A: Fund Characteristics

| | Full Sample | BM > 60 | BM ≤ 60 | Difference | *t*-stat |
|---|---|---|---|---|---|
| Prospectus Active Share | 80.7% | 93.5% | 76.8% | 16.7% | 40.84 |
| Assets (billions of $) | 1.16 | 0.88 | 1.26 | -0.38 | -3.71 |
| Age (years) | 16.1 | 13.8 | 16.9 | -3.0 | -5.14 |
| Expense Ratio | 1.20% | 1.29% | 1.16% | 0.13% | 8.91 |
| Turnover Ratio | 77.2% | 77.4% | 77.1% | 0.3% | 0.11 |
| Number of Holdings | 89.4 | 92.3 | 88.3 | 4.0 | 1.31 |
| Institutional (% of assets) | 26.1% | 23.5% | 27.0% | -3.5% | -2.08 |

Panel B: Fund Prospectus Style

| | Full Sample | BM > 60 | BM ≤ 60 | Difference | *t*-stat |
|---|---|---|---|---|---|
| Large Cap | 61.8% | 19.6% | 76.3% | -56.7% | -27.87 |
| Small/Mid Cap | 38.2% | 80.4% | 23.7% | 56.7% | 27.87 |
| Growth | 29.7% | 24.4% | 31.5% | -7.1% | -3.25 |
| Blend | 46.7% | 49.6% | 45.7% | 3.9% | 1.45 |
| Value | 23.6% | 26.0% | 22.8% | 3.1% | 1.34 |

**Table 4: Probability of *Benchmark Mismatch* greater than 60%**

This table shows estimates from the following logit model:

$$BM > 60\%_{i,t} = \alpha + \beta * Active\ Share_{i,t} + \delta * Chars_{i,t} + \gamma * Style_i + FE + \varepsilon_{i,t}$$

where $BM > 60\%_{i,t}$ is a dummy variable equal to one if the *Benchmark Mismatch* for fund *i* based on holdings in quarter *t* is greater than 60%. $Active\ Share_{i,t}$ is a vector of information about fund *i*'s active share in quarter *t*. It includes the active share relative to the prospectus benchmark and a dummy variable equal to one if the prospectus active share is among the top 20% in the quarter. $Chars_{i,t}$ is a vector of characteristics for fund *i* available as of quarter *t*. It includes the natural log of assets, natural log of age, expense ratio, turnover ratio, the number of equity positions, and the percentage of fund assets held within institutional share classes. $Style_i$ is a vector of information about fund *i*'s style based on its prospectus benchmark. It includes a large-cap dummy, a blend dummy, and a growth dummy, which are dummy variables equal to one if a fund's prospectus benchmark aligns with that style. $FE$ represents year-quarter fixed effects and are included only in column (6). The model is estimated using the full sample of fund-quarters. *t*-statistics are reported in brackets below each coefficient and are calculated using standard errors clustered by fund and year-month.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Prospectus Active Share | 0.25 | 0.25 | | | 0.29 | 0.31 |
| | [25.69] | [22.70] | | | [15.71] | [15.73] |
| Top 20% AS Dummy | | 0.07 | | | 0.27 | 0.25 |
| | | [0.63] | | | [2.21] | [1.91] |
| Assets | | | -0.03 | | 0.05 | 0.04 |
| | | | [-0.81] | | [1.33] | [1.12] |
| Age | | | -0.24 | | 0.02 | 0.06 |
| | | | [-3.54] | | [0.22] | [0.63] |
| Expense Ratio | | | 1.09 | | -0.17 | -0.34 |
| | | | [7.11] | | [-0.98] | [-1.93] |
| Turnover Ratio | | | -0.00 | | 0.00 | 0.00 |
| | | | [-2.11] | | [0.64] | [0.05] |
| Number of Holdings | | | 0.00 | | 0.01 | 0.01 |
| | | | [3.92] | | [7.99] | [8.19] |
| Institutional Ownership | | | -0.00 | | 0.00 | -0.00 |
| | | | [-0.60] | | [0.07] | [-0.28] |
| Large Cap Dummy | | | | -2.88 | -1.02 | -0.94 |
| | | | | [-21.31] | [-7.77] | [-6.72] |
| Blend Dummy | | | | 0.40 | -0.48 | -0.52 |
| | | | | [2.84] | [-3.12] | [-3.28] |
| Growth Dummy | | | | -0.81 | -0.30 | -0.25 |
| | | | | [-5.43] | [-1.87] | [-1.53] |
| Fixed Effects | No | No | No | No | No | Yes |
| Observations | 53,316 | 53,316 | 53,316 | 53,316 | 53,316 | 53,316 |

**Table 5: Difference in benchmark returns as a function of *Active Gap* and *Benchmark Mismatch***

This tables show the average differences in annualized return between the AS benchmark and the prospectus benchmark for fund-months in which those benchmarks are different. Fund-months are sorted unconditionally on *Active Gap* (AG) and *Benchmark Mismatch* (BM) based on pre-set cut-offs and average differences are reported for each of the resulting groups. Panel A sorts funds into five groups based on *Benchmark Mismatch* and Panel B sorts funds into two groups based on *Benchmark Mismatch*. The "High – Low" column reports the difference in the results between the "0 < BM ≤ 20" and "BM > 80" groups in Panel A and difference in results between the "BM ≤ 60" and "BM > 60" in Panel B. The "High – Low" row reports the difference in results between the "0 < AG ≤ 1.25" and "AG > 5" groups in both panels. *t*-statistics are reported in brackets below each coefficient and are calculated using standard errors clustered by fund and year-month.

Panel A: Five Ranges for Benchmark Mismatch

| | All | 0 < BM ≤ 20 | 20 < BM ≤ 40 | 40 < BM ≤ 60 | 60 < BM ≤ 80 | BM > 80 | High − Low |
|---|---|---|---|---|---|---|---|
| All | 0.68% | -0.12% | 0.04% | 0.53% | 1.37% | 1.64% | 1.75% |
| | [2.72] | [-0.75] | [0.18] | [1.70] | [2.25] | [2.97] | [3.04] |
| 0 < AG ≤ 1.25 | 0.43% | 0.02% | 0.15% | 0.24% | 1.00% | 1.26% | 1.24% |
| | [2.15] | [0.12] | [0.71] | [0.76] | [1.64] | [1.99] | [2.05] |
| 1.25 < AG ≤ 2.5 | 0.68% | -0.03% | -0.12% | 0.84% | 1.61% | 1.87% | 1.90% |
| | [2.81] | [-0.18] | [-0.47] | [2.19] | [2.52] | [2.84] | [2.92] |
| 2.5 < AG ≤ 3.75 | 0.75% | -0.29% | -0.03% | 0.71% | 1.62% | 1.67% | 1.96% |
| | [2.53] | [-1.40] | [-0.09] | [1.56] | [2.33] | [2.57] | [2.85] |
| 3.75 < AG ≤ 5 | 0.73% | -0.45% | 0.07% | 0.75% | 1.12% | 1.69% | 2.14% |
| | [2.19] | [-1.60] | [0.19] | [1.68] | [1.58] | [2.26] | [2.62] |
| AG > 5 | 0.84% | -0.75% | 0.18% | 0.26% | 1.45% | 1.62% | 2.37% |
| | [2.31] | [-1.78] | [0.46] | [0.67] | [2.10] | [2.54] | [2.68] |
| High − Low | 0.42% | -0.77% | 0.03% | 0.02% | 0.45% | 0.36% | 1.13% |
| | [1.31] | [-2.02] | [0.10] | [0.05] | [0.86] | [0.55] | [1.34] |

Panel B: Two Ranges for Benchmark Mismatch

| | 0 < BM ≤ 60 | BM > 60 | High − Low |
|---|---|---|---|
| All | 0.18% | 1.50% | 1.32% |
| | [0.94] | [3.20] | [3.07] |
| 0 < AG ≤ 1.25 | 0.15% | 1.09% | 0.94% |
| | [1.02] | [2.13] | [1.85] |
| 1.25 < AG ≤ 2.5 | 0.14% | 1.72% | 1.58% |
| | [0.78] | [3.32] | [3.15] |
| 2.5 < AG ≤ 3.75 | 0.18% | 1.64% | 1.46% |
| | [0.79] | [3.03] | [2.92] |
| 3.75 < AG ≤ 5 | 0.26% | 1.40% | 1.14% |
| | [0.97] | [2.39] | [2.04] |
| AG > 5 | 0.20% | 1.56% | 1.36% |
| | [0.61] | [2.90] | [2.78] |
| High − Low | 0.05% | 0.47% | 0.42% |
| | [0.17] | [0.87] | [0.83] |

**Table 6: Model of differences in prospectus and AS benchmark returns**

This table shows results from the following model:

$$R_{AS,i,t} - R_{p,i,t} = \alpha + \beta * Active\ Share_{i,t} + \delta * Mismatch_{i,t} + \gamma * Chars_{i,t} + FE + \varepsilon_{i,t}$$

where $R_{AS,i,t}$ is the annualized return on fund *i*'s AS benchmark in month *t* and $R_{P,i,t}$ is the annualized return on fund *i*'s prospectus benchmark in month *t*. $Active\ Share_{i,t}$ is a vector of information about fund *i*'s active share at the start of month *t*. It includes the fund's prospectus active share and a dummy variable equal to one if the prospectus active share is among the top 20% at the start of the month. $Mismatch_{i,t}$ is a vector of information about fund *i*'s mismatch status at the start of month *t*. It includes *Benchmark Mismatch*, *Active Gap*, and a dummy variable equal to one if *Benchmark Mismatch* is among the top 20% at the start of the month. $Chars_{i,t}$ is a vector of characteristics for fund *i* available as of the start of month *t*. It includes the natural log of assets, natural log of age, expense ratio, turnover ratio, the number of equity positions, and the percentage of fund assets held within institutional share classes. The characteristics are included in all presented models, but the coefficients associated with the variables are suppressed in the table. $FE$ represents style and year-month fixed effects, which are included in all presented models. The model is estimated using the sample of funds with different prospectus and AS benchmarks. *t*-statistics are reported in brackets below each coefficient and are calculated using standard errors clustered by fund and year-month.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Prospectus Active Share | 0.062 | | | 0.039 | 0.052 | | 0.033 |
| | [3.21] | | | [2.72] | [2.95] | | [2.38] |
| Benchmark Mismatch | | 0.033 | | 0.023 | | 0.039 | 0.029 |
| | | [3.21] | | [2.54] | | [3.00] | [2.68] |
| Active Gap | | | 0.095 | 0.040 | | | |
| | | | [2.62] | [1.23] | | | |
| Top 20% AS Dummy | | | | | 0.773 | | 0.461 |
| | | | | | [3.19] | | [2.07] |
| Top 20% BM Dummy | | | | | | -0.432 | -0.456 |
| | | | | | | [-0.65] | [-0.67] |
| Characteristic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 125,352 | 125,352 | 125,352 | 125,352 | 125,352 | 125,352 | 125,352 |

**Table 7: Performance of funds as a function of *Benchmark Mismatch* and active share**

This table shows returns for different groups of funds using multiple models. To form the groups, the full sample of fund-months (including funds with the same prospectus and minimum active benchmark) are sorted independently on prospectus active share and *Benchmark Mismatch* (BM). With respect to active share, funds are sorted into quintiles at the beginning of each month. Those funds in fifth quintile (i.e., those with highest active share) are tested separately from those in the other four quintiles, and the difference in results between those groups is considered in the "Q5 – Q1234" portion of the table. With respect to *Benchmark Mismatch*, funds are sorted based on whether *Benchmark Mismatch* is greater than or less than 60%. The difference in results between those groups is considered in the "Diff" column. To adjust the returns, three different models are used. The prospectus method reports the average of the monthly average differences between the fund return and the prospectus benchmark return. The BM method reports the average of the monthly average differences between the fund return and the prospectus return if *Benchmark Mismatch* is less than 60%. If *Benchmark Mismatch* is greater than 60%, then the AS benchmark is used instead. The "Difference" row reports the difference in the values resulting from the prospectus and BM methods. The CPZ7 method regresses the time-series of the monthly average excess fund returns against the Cremers, Petajisto, and Zitzewitz (2012) seven factors and reports the intercept from that regression. The results from each model are annualized. *t*-statistics are reported in brackets below each measurement.

| | Method | All | BM > 60% | BM ≤ 60% | Diff |
|---|---|---|---|---|---|
| All Funds | Prospectus | -0.25% | 0.53% | -0.51% | 1.04% |
| | | [-0.78] | [1.23] | [-1.67] | [3.20] |
| | BM | -0.59% | -0.86% | -0.51% | -0.35% |
| | | [-2.07] | [-2.38] | [-1.67] | [-1.11] |
| | Difference | 0.35% | 1.39% | 0.00% | 1.39% |
| | | [3.28] | [3.30] | - | [3.30] |
| | CPZ7 | -0.53% | -0.11% | -0.65% | 0.54% |
| | | [-1.59] | [-0.23] | [-2.06] | [1.46] |
| Prospectus Active Share Quintile 5 | Prospectus | 0.75% | 0.72% | 0.75% | -0.02% |
| | | [1.83] | [1.55] | [1.60] | [-0.04] |
| | BM | -0.42% | -0.92% | 0.75% | -1.67% |
| | | [-1.21] | [-2.32] | [1.60] | [-3.32] |
| | Difference | 1.16% | 1.65% | 0.00% | 1.65% |
| | | [3.62] | [3.64] | - | [3.64] |
| | CPZ7 | 0.50% | 0.07% | 1.28% | -1.21% |
| | | [1.05] | [0.14] | [2.14] | [-2.09] |
| Prospectus Active Share Quintiles 1, 2, 3, and 4 | Prospectus | -0.49% | 0.27% | -0.62% | 0.89% |
| | | [-1.55] | [0.59] | [-1.97] | [2.50] |
| | BM | -0.64% | -0.72% | -0.62% | -0.10% |
| | | [-2.11] | [-1.82] | [-1.97] | [-0.30] |
| | Difference | 0.14% | 0.99% | 0.00% | 0.99% |
| | | [2.36] | [2.14] | - | [2.14] |
| | CPZ7 | -0.79% | -0.31% | -0.82% | 0.51% |
| | | [-2.47] | [-0.68] | [-2.67] | [1.12] |
| Q5 - Q1234 | Prospectus | 1.24% | 0.46% | 1.37% | -0.91% |
| | | [4.18] | [1.36] | [3.09] | [-1.69] |
| | BM | 0.22% | -0.20% | 1.37% | -1.57% |
| | | [0.72] | [-0.60] | [3.09] | [-2.98] |
| | Difference | 1.02% | 0.66% | 0.00% | 0.66% |
| | | [3.66] | [1.99] | - | [1.99] |
| | CPZ7 | 1.29% | 0.38% | 2.10% | -1.72% |
| | | [4.18] | [0.95] | [4.26] | [-2.94] |

**Table 8: Performance of funds as a function of active share, past performance, and benchmark discrepancy**

This table shows the average Cremers, Petajisto, and Zitzewitz (2012) seven-factor alpha for different groups of funds. The alpha for a given group is estimated using the monthly returns on an equal weight portfolio of the funds in the group. The reported alpha is annualized. To form the groups, the full sample of fund-months (including funds with the same prospectus and minimum active benchmark) are first sorted based on prospectus active share. The “Bottom 80%” group contains the funds within the lowest 80% of active share at the beginning of each month. The “Top 20%” group contains the funds within the highest 20% of active share at the beginning of each month. Next, funds within each of those active share groups are sorted each month based on their prospectus-benchmark-adjusted return over the previous year. The “Bottom 80%” group contains the funds within the lowest 80% of benchmark-adjusted return. The “Top 20%” group contains the funds within the highest 20% of benchmark-adjusted return. Finally, funds within each active share and past performance group are sorted based on whether the fund has a benchmark discrepancy. If a fund’s *Benchmark Mismatch* is greater than 60% at the beginning of the month, then it is placed in the “Yes” group, otherwise it is placed in the “No” group. *t*-statistics are reported in brackets below each measurement.

| Active Share | Bottom 80% | | | | Top 20% | | | |
|---|---|---|---|---|---|---|---|---|
| CPZ7 Alpha | -0.64% | | | | 0.71% | | | |
| | [-2.25] | | | | [1.37] | | | |
| Past Performance | Bottom 80% | | Top 20% | | Bottom 80% | | Top 20% | |
| CPZ7 Alpha | -0.94% | | 0.53% | | 0.32% | | 2.31% | |
| | [-2.65] | | [0.85] | | [0.57] | | [3.01] | |
| Benchmark Discrepancy | Yes | No | Yes | No | Yes | No | Yes | No |
| CPZ7 Alpha | -0.82% | -0.93% | 0.72% | 0.42% | -0.03% | 1.03% | 1.72% | 3.21% |
| | [-1.44] | [-2.69] | [1.00] | [0.65] | [-0.05] | [1.41] | [1.97] | [2.37] |

**Table 9: Comparison of benchmark discrepancy identification procedures**

This table shows the average return for fund-months identified as having a benchmark discrepancy following two different procedures. In the BM procedure, a fund is considered to have a benchmark discrepancy if our *Benchmark Mismatch* measure is greater than 60%. In the Sensoy procedure, a fund is considered to have a benchmark discrepancy if the Morningstar style boxes and fund-benchmark correlations indicate a more appropriate benchmark. The reported returns are adjusted using various benchmarks. The "Prospectus" row reports the return less the prospectus benchmark return. The "AS" row reports the return less the AS benchmark return. The "Sensoy" row reports the return less the return on the appropriate benchmark identified using the Sensoy procedure. The "Pro − AS" reports the difference between the prospectus and AS results, and the "Pro – Sensoy" reports the difference between the prospectus and Sensoy results. All returns are annualized. *t*-statistics are reported in brackets below each coefficient and are calculated using standard errors clustered by fund and year-month.

| Mismatch? | BM | Yes | - | Yes | No | Yes |
|---|---|---|---|---|---|---|
| | Sensoy | - | Yes | No | Yes | Yes |
| Benchmark Adjusted Return | Prospectus | 0.37% | -0.16% | 0.27% | -0.52% | 0.52% |
| | | [0.74] | [-0.31] | [0.49] | [-1.04] | [0.82] |
| | AS | -1.16% | | -1.06% | | -1.29% |
| | | [-3.44] | | [-2.92] | | [-3.54] |
| | Sensoy | | -0.91% | | -0.86% | -0.97% |
| | | | [-3.17] | | [-3.05] | [-2.44] |
| Differences | Prospectus − AS | 1.52% | | 1.33% | | 1.81% |
| | | [3.14] | | [2.25] | | [3.20] |
| | Prospectus − Sensoy | | 0.76% | | 0.34% | 1.49% |
| | | | [1.64] | | [0.82] | [2.59] |

**Table 10: Factor differences between the prospectus and AS benchmarks**

This table shows results from the following model:

$$Return_{AS,t} - Return_{pro,t} = \beta * Factor_t + \varepsilon_t$$

where $Return_{AS,t}$ is the average annualized return on the AS benchmark in month *t* for all funds with a *Benchmark Mismatch* greater than 60%. $Return_{pro,t}$ is the average annualized return on the prospectus benchmark in month *t* for all funds with a *Benchmark Mismatch* greater than 60%. $Factor_t$ is a vector of factor returns in month *t*. The factors included are all of those in the seven-factor Cremers, Petajisto, and Zitzewitz (2012) model. The model is estimated using the full sample of funds with a *Benchmark Mismatch* greater than 60% and for subgroups with different prospectus identified styles. The "Prospectus" row reports the average of the monthly average differences between the fund return and prospectus benchmark return. The "AS" row reports the average of the monthly average differences between the fund return and AS benchmark return. Both of those return differences are annualized. The "Difference" row tests the difference between the two rows above it and is equal to the average of $Return_{AS,t}$ less the average of $Return_{pro,t}$. Rows "S5RF" through "UMD" report the $\beta$'s from the above model for each factor. The "Total Factor Return" row reports sum of the products of the estimated factor exposures and annualized factor returns. *t*-statistics associated with tests of whether the values in the table are different from zero are reported in brackets below each measurement.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Style | All | Large | Small/Mid | Growth | Value | Blend |
| Prospectus | 0.66% | 1.05% | 0.18% | 0.97% | -0.61% | 0.70% |
| | [1.16] | [0.78] | [0.31] | [1.51] | [-0.92] | [0.94] |
| AS | -0.84% | -1.33% | -0.94% | -0.80% | -1.01% | -1.08% |
| | [-1.70] | [-2.38] | [-1.72] | [-0.99] | [-1.93] | [-1.75] |
| Difference | 1.50% | 2.38% | 1.12% | 1.77% | 0.40% | 1.78% |
| | [3.67] | [2.13] | [1.94] | [2.19] | [0.71] | [3.21] |
| S5RF | -0.01 | 0.02 | -0.01 | -0.02 | -0.01 | -0.01 |
| | [-1.14] | [1.22] | [-0.77] | [-1.18] | [-0.59] | [-1.46] |
| RMS5 | 0.12 | 0.69 | -0.05 | 0.02 | 0.01 | 0.26 |
| | [4.99] | [11.89] | [-1.49] | [0.42] | [0.49] | [9.25] |
| R2RM | -0.12 | 0.16 | -0.22 | -0.13 | -0.11 | -0.08 |
| | [-6.83] | [3.75] | [-9.01] | [-5.20] | [-3.94] | [-3.54] |
| S5VS5G | -0.04 | -0.08 | -0.03 | -0.04 | -0.08 | -0.05 |
| | [-2.08] | [-2.02] | [-1.09] | [-1.22] | [-2.55] | [-1.98] |
| RMVRMG | -0.02 | 0.14 | -0.08 | -0.02 | 0.03 | -0.06 |
| | [-0.72] | [1.67] | [-2.83] | [-0.69] | [0.76] | [-2.08] |
| R2VR2G | 0.04 | -0.15 | 0.10 | 0.25 | -0.15 | 0.03 |
| | [1.49] | [-2.52] | [3.78] | [7.28] | [-3.87] | [1.03] |
| UMD | 0.02 | 0.01 | 0.03 | 0.06 | -0.02 | 0.02 |
| | [2.39] | [0.95] | [2.08] | [3.50] | [-1.76] | [1.99] |
| $R^2$ | 27.7% | 75.7% | 49.9% | 63.3% | 36.9% | 34.5% |
| Total Factor Return | 0.57% | 1.60% | 0.41% | 0.87% | -0.18% | 0.73% |
| | [2.65] | [1.64] | [1.01] | [1.35] | [-0.52] | [2.24] |

**Table 11: Non-traditional factor differences between the prospectus and AS benchmarks**

This table shows results from the following model:

$$Return_{AS,t} - Return_{pro,t} = \beta * Factor_t + \varepsilon_t$$

where $Return_{AS,t}$ is the average annualized return on the AS benchmark in month *t* for all funds with a *Benchmark Mismatch* greater than 60%. $Return_{pro,t}$ is the average annualized return on the prospectus benchmark in month *t* for all funds with a *Benchmark Mismatch* greater than 60%. $Factor_t$ is a vector of factor returns in month *t*. The factors include all of those in the seven-factor Cremers, Petajisto, and Zitzewitz (2012) model, the Fama and French (2015) profitability (RMW) and investment (CMA) factors, the Stambaugh and Yuan (2017) management (MGMT) and performance (PERF) factors, the Frazzini and Pedersen (2014) betting against beta (BAB) factor, the Asness, Frazzini, and Pedersen (2017) quality minus junk (QMJ) factor, and the Pastor and Stambaugh (2004) traded liquidity factor. The model is estimated using the full sample of funds with a *Benchmark Mismatch* greater than 60% and for subgroups with different prospectus identified styles. Rows "S5RF" through "LIQ" report the $\beta$'s from the above model for each factor. The "Total Factor Return" row reports sum of the products of the estimated factor exposures and annualized factor returns. *t*-statistics associated with tests of whether the values in the table are different from zero are reported in brackets below each measurement.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Prospectus Style | All | Large | Small/Mid | Growth | Value | Blend |
| S5RF | 0.00 | 0.05 | 0.00 | -0.02 | 0.02 | 0.01 |
| | [0.26] | [2.95] | [0.07] | [-1.28] | [2.05] | [0.53] |
| RMS5 | 0.13 | 0.68 | -0.01 | 0.04 | 0.02 | 0.28 |
| | [6.27] | [13.34] | [-0.39] | [0.82] | [0.48] | [10.50] |
| R2RM | -0.08 | 0.21 | -0.19 | -0.10 | -0.05 | -0.03 |
| | [-4.52] | [6.02] | [-7.52] | [-3.86] | [-2.21] | [-1.52] |
| S5VS5G | 0.02 | 0.03 | 0.00 | -0.01 | 0.03 | 0.05 |
| | [1.16] | [0.70] | [0.13] | [-0.43] | [1.22] | [1.73] |
| RMVRMG | -0.10 | 0.07 | -0.13 | -0.08 | -0.06 | -0.15 |
| | [-4.95] | [1.02] | [-4.50] | [-2.13] | [-2.03] | [-6.02] |
| R2VR2G | 0.01 | -0.12 | 0.06 | 0.19 | -0.15 | 0.02 |
| | [0.52] | [-2.04] | [1.76] | [4.66] | [-5.46] | [0.57] |
| UMD | 0.01 | -0.01 | 0.02 | 0.06 | -0.05 | -0.00 |
| | [0.97] | [-0.26] | [1.48] | [3.24] | [-3.05] | [-0.30] |
| RMW | 0.14 | 0.08 | 0.09 | 0.15 | 0.12 | 0.14 |
| | [4.47] | [1.34] | [2.31] | [3.17] | [3.37] | [3.85] |
| CMA | -0.02 | 0.03 | -0.04 | -0.02 | 0.03 | -0.01 |
| | [-0.63] | [0.64] | [-1.41] | [-0.38] | [0.90] | [-0.35] |
| MGMT | -0.00 | -0.10 | 0.06 | 0.05 | -0.06 | -0.02 |
| | [-0.07] | [-2.59] | [2.07] | [1.35] | [-2.28] | [-0.70] |
| PERF | -0.00 | 0.03 | -0.02 | -0.03 | 0.03 | 0.02 |
| | [-0.13] | [0.89] | [-0.95] | [-1.47] | [1.68] | [0.98] |
| QMJ | 0.04 | 0.10 | 0.05 | -0.00 | 0.09 | 0.06 |
| | [1.33] | [2.61] | [1.35] | [-0.09] | [2.52] | [1.59] |
| BAB | 0.02 | 0.00 | 0.01 | 0.02 | 0.01 | 0.02 |
| | [2.46] | [0.23] | [0.55] | [1.48] | [0.83] | [1.76] |
| LIQ | -0.00 | 0.02 | 0.02 | 0.00 | 0.01 | -0.02 |
| | [-0.15] | [1.63] | [2.31] | [0.27] | [1.37] | [-2.51] |
| $R^2$ | 52.4% | 81.1% | 58.2% | 67.3% | 59.4% | 53.9% |
| Total Factor Return | 1.31% | 2.18% | 1.23% | 1.50% | 0.74% | 1.50% |
| | [4.48] | [2.17] | [2.80] | [2.26] | [1.71] | [3.70] |

**Table 12: Prospectus- and AS-benchmark-adjusted returns evaluated with factors**

This table shows results from the following model:

$$Return_{fund,t} - Return_{bench,t} = \alpha + \beta * Factor_t + \varepsilon_t$$

where $Return_{fund,t}$ is the average annualized return in month *t* for all funds with a *Benchmark Mismatch* greater than 60% and $Return_{bench,t}$ is the average annualized return on those funds' benchmarks in month $t$. In the "Prospectus Adjusted" row, the benchmark listed in the fund prospectus is used. In the "AS Adjusted" row, the AS benchmark is used. $Factor_t$ is a vector of factor returns in month *t*. In the column labeled "Return", no factors are included in the model. In the columns with the heading "CPZ7", all of the factors in the seven-factor Cremers, Petajisto, and Zitzewitz (2012) model are included. In the columns with the heading "CPZ7+", all of the non-traditional factors discussed in Table 11 are also included. Within the "CPZ7" and "CPZ7+" columns, the "Alpha" column reports the $\alpha$ from the above model and the "Change" column reports the difference between that $\alpha$ and the value from the "Return" column. The "Difference" row reports the difference between the values in the first two rows. Panel A shows results using the full sample of funds with a *Benchmark Mismatch* greater than 60%, and Panel B shows results using just those funds within that group that have a large-cap prospectus identified style. *t*-statistics are reported in brackets below each measurement.

Panel A: All Funds

| | | CPZ7 | | CPZ7+ | |
|---|---|---|---|---|---|
| | Return | Alpha | Change | Alpha | Change |
| Prospectus Adjusted | 0.65% | 0.51% | -0.15% | -0.29% | -0.94% |
| | [1.36] | [1.16] | [-0.33] | [-0.72] | [-2.37] |
| AS Adjusted | -0.87% | -0.40% | 0.47% | -0.30% | 0.57% |
| | [-2.23] | [-1.08] | [1.27] | [-0.75] | [1.44] |
| Difference | 1.52% | 0.91% | -0.61% | 0.01% | -1.51% |
| | [3.45] | [2.28] | [-1.54] | [0.03] | [-4.24] |

Panel B: Large Cap

| | | CPZ7 | | CPZ7+ | |
|---|---|---|---|---|---|
| | Return | Alpha | Change | Alpha | Change |
| Prospectus Adjusted | 1.05% | -0.14% | -1.18% | -0.89% | -1.94% |
| | [0.87] | [-0.22] | [-1.92] | [-1.40] | [-3.05] |
| AS Adjusted | -1.36% | -0.95% | 0.41% | -0.91% | 0.45% |
| | [-3.11] | [-2.21] | [0.96] | [-2.05] | [1.00] |
| Difference | 2.41% | 0.81% | -1.60% | 0.02% | -2.38% |
| | [2.10] | [1.52] | [-3.00] | [0.04] | [-4.52] |

**Table 13: Response of investor flows to different measures of performance**

This table shows results from the following model:

$$Flow_{i,t} = \theta + \beta * Performance_{i,t} + \gamma * Mismatch_{i,t} + \delta * Chars_{i,t} + FE + \varepsilon_{i,t}$$

where $Flow_{i,t}$ is the percentage implied net flow for fund $i$ in month $t$. $Performance_{i,t}$ is a vector of information about fund *i*'s performance over the year ending at the start of month *t*. It includes the difference between fund *i*'s return and the return on fund *i*'s AS benchmark, the difference between the return on fund $i$'s AS benchmark and prospectus benchmark, and fund *i*'s annualized CAPM alpha. In columns (1) through (4), the actual returns are used. In columns (5) through (7), each of the return variables is ranked at the start of each month and scaled from zero to one. $Mismatch_{i,t}$ is the *Benchmark Mismatch* (BM) for fund *i* as of the start of month *t*. $Chars_{i,t}$ is a vector of characteristics for fund *i* available as of the start of month *t*. It includes the natural log of assets, natural log of age, expense ratio, turnover ratio, the number of equity positions, and the percentage of fund assets held within institutional share classes. The characteristics are included in all presented models, but the coefficients associated with the variables are suppressed in the table. $FE$ represents style and year-month fixed effects, which are included in all presented models. The model is estimated using the sample of fund-months with different prospectus and AS benchmarks. In column (7), only the funds in the top 20% of assets at the start of month *t* are used to estimate the model. *t*-statistics are reported in brackets below each coefficient and are calculated using standard errors clustered by fund and year-month.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Fund Ret - AS Ret | 0.12 | 0.13 | 0.09 | 0.129 | 1.97 | 1.26 | 2.24 |
| | [23.27] | [23.86] | [15.35] | [13.38] | [18.90] | [5.46] | [4.78] |
| AS Ret - Prospectus Ret | | 0.07 | 0.04 | 0.140 | 0.69 | 1.83 | 1.83 |
| | | [14.03] | [8.76] | [9.98] | [9.16] | [8.30] | [5.50] |
| CAPM Alpha | | | 0.07 | | 1.78 | | |
| | | | [14.05] | | [16.93] | | |
| Benchmark Mismatch | | | | -0.002 | | | |
| | | | | [-1.52] | | | |
| (Fund Ret - AS Ret) * BM | | | | 0.000 | | | |
| | | | | [0.63] | | | |
| (AS Ret - Prospectus Ret) * BM | | | | -0.001 | | | |
| | | | | [-5.64] | | | |
| (Fund Ret - AS Ret)$^2$ | | | | | | 1.83 | 0.57 |
| | | | | | | [7.48] | [1.25] |
| (AS Ret - Prospectus Ret)$^2$ | | | | | | -0.73 | -0.83 |
| | | | | | | [-3.39] | [-2.46] |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Returns | Actual | Actual | Actual | Actual | Ranking | Ranking | Ranking |
| Sample | BM > 0% | BM > 0% | BM > 0% | BM > 0% | BM > 0% | BM > 0% | BM > 0% Size Q5 |
| Characteristic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 122,411 | 122,411 | 122,411 | 122,411 | 122,411 | 122,411 | 24,363 |