Credit Where It's Not Due: Misbenchmarking by Active Bond Funds

Caitlin Dannhauser Michele Dathan Michael Young Qifei Zhu*

March 2026

Abstract

Misspecified benchmarks are widespread among active bond mutual funds, driven in part by the industry's heavy dependence on the Bloomberg U.S. Aggregate Index. Active funds outperform their self-declared benchmarks by 3% over a 36-month horizon. However, this outperformance is fully explained by their choice of benchmarks, which tend to have lower systematic risk than their actual portfolios. When evaluated against more appropriate benchmarks identified from return correlations, these funds underperform. Benchmark-adjusted returns strongly predict future flows, giving managers incentives to misbenchmark. Although benchmark changes are infrequent, they appear strategic: funds switch to easier-to-beat benchmarks following periods of underperformance and subsequently experience a temporary increase in flows.

JEL Classification: D22, G11, G18, G23
Keywords: Mutual funds; Bond investing; Benchmarking; Obfuscation

*Caitlin Dannhauser (caitlin.dannhauser@villanova.edu) is at the Villanova School of Business; Michele Dathan (michele.h.dathan@frb.gov) is at the Board of Governors of the Federal Reserve; Michael Young (myoung@missouri.edu) is at the University of Missouri, and Qifei Zhu (qifeizhu@nus.edu.sg) is at the National University of Singapore. The analysis and conclusions are those of the authors and do not reflect the views of the Board of Governors of the Federal Reserve.

Benchmarks are central to the mutual fund industry, serving as the standard against which fund performance is evaluated and managerial compensation is often determined (Cremers and Petajisto, 2009; Ivković and Weisbenner, 2009; Sensoy, 2009). However, benchmark-adjusted performance paints strikingly different pictures in the equity and fixed-income markets. For active equity funds, Chen, Evans, and Sun (2025) document an average three-year benchmark-adjusted return of -3.22%, with at least half of funds failing to beat their benchmarks. By contrast, in a large sample of active fixed income funds, we find substantial benchmark outperformance: the average three-year benchmark-adjusted return is 2.97%, and more than three-quarters of funds deliver positive benchmark-adjusted performance. These contrasting performance patterns are also reflected in investor flows. Over the past decade, active equity funds have experienced persistent outflows amid the rise of passive investing, whereas active bond funds have continued to attract cumulative inflows (Morningstar, 2024).

The strong performance of actively managed bond funds lends support to some practitioners' claim that active management is particularly effective in fixed-income markets and can justify investors' preference for active over passive bond funds (Financial Times, 2026).[1] A natural explanation lies in the size and heterogeneity of fixed-income markets (Bessembinder, Spatt, and Venkataraman, 2020). Since benchmark indices often contain thousands of constituent securities, active managers have considerable discretion to generate excess returns through both detailed security selection and macroeconomic factor timing (Huang and Wang, 2014; Chen, Ferson, and Peters, 2010).

A fundamentally different explanation, however, is that active bond funds' benchmark outperformance is largely a mirage driven by strategic benchmark selection. Because funds have discretion in selecting their prospectus benchmarks, they may choose benchmarks that do not accurately capture the risk characteristics of their portfolios, a practice we refer to as *misbenchmarking*. If funds systematically adopt benchmarks with lower risk exposure than

[1]Such arguments in the financial press echo asset-manager marketing materials, including Fidelity's "Why active management may be better for bond funds," JP Morgan's "Why active fixed income consistently outperforms the Agg," and Wellington Management's "Five reasons to be active in fixed income."

their actual holdings—and are thus easier to outperform—then benchmark-adjusted returns will be artificially inflated. Under this view, the widespread and persistent outperformance of active bond funds reflects not superior investment ability, but rather a misalignment between fund portfolios and their stated benchmarks.

In this paper, we investigate whether active bond funds' self-designated benchmarks are appropriate measures of portfolio risk and quantify the extent to which benchmark selection contributes to their measured outperformance. This is an economically important issue because benchmark-relative performance is central to investor capital allocation decisions.[2] Empirically, we show that flows respond strongly to benchmark-adjusted returns controlling for raw fund returns, even when comparing funds within the same Lipper category and with the same Morningstar rating. The evidence suggests that investors place primary focus on performance measured relative to funds' self-declared benchmarks when directing capital. Managers thus have a clear incentive to select benchmarks against which outperformance is easier to demonstrate.

Given the importance investors attach to benchmark-adjusted performance, what indices do bond funds in the U.S. benchmark to? In a large sample of actively managed bond funds, we document that approximately 56% of them select the Bloomberg U.S. Aggregate Bond Index (the "Agg") as their benchmark.[3] The Agg is designed to provide broad exposure to U.S. dollar denominated investment grade government, agency, and corporate bonds. However, Agg-benchmarked funds in our sample span more than 30 distinct Lipper investment categories, suggesting substantial heterogeneity in portfolio objectives and risk exposures. One important dimension of this heterogeneity is a persistent overweight in corporate bonds, the riskiest issuer sector represented in the Agg. Consequently, many funds that benchmark to the Agg have investment profiles that are materially riskier than the benchmark they claim to track.

[2] Bogfan and Schrass (2025) surveys mutual fund investors. 86% of mutual fund investors view benchmark-adjusted returns as either "somewhat important" or "very important" when allocating capital.

[3] The use of a singular benchmark mirrors the dependence on the S&P 500 by active equity mutual funds in early 2000s, as documented by Sensoy (2009).

Furthermore, benchmark selection can explain the full amount of average benchmark-adjusted outperformance in our sample. Over a 36-month horizon, the average active bond fund outperforms its self-declared benchmark by 2.97%. Remarkably, this is exactly the same amount by which the average return in the fund's Lipper category exceeds the return on that benchmark. This finding suggests that the measured outperformance of active bond funds is largely attributable to benchmark choice, rather than to fund-specific value added. Given the widespread use of the Agg, such outperformance is consistent with many funds adopting benchmarks whose risk profiles are materially more conservative than those of their actual portfolios, thereby making benchmark outperformance easier to achieve.

To quantify the extent of benchmark mismatch in fixed-income asset management, we examine better-matched benchmarks for each fund. Following Sensoy (2009), we define a fund's best-matched benchmark as the benchmark whose returns have the highest correlation with the fund's gross returns. We consider two alternative benchmark universes. The first consists of the 30 bond indices most commonly used as benchmarks in our sample. Within this universe, a fund is classified as properly benchmarked if its stated benchmark is the one most highly correlated with its returns. The second universe consists of 67 pseudo-benchmarks constructed from different portfolio weights on (i) the Agg, (ii) the Bloomberg U.S. Corporate Index, and (iii) the ICE BofA U.S. High Yield Index. This expanded benchmark set allows for intermediate levels of credit risk and high-yield exposure beyond the investment-grade profile of the Agg. Within this alternative universe, a fund is considered properly benchmarked if both the fund and its stated benchmark map to the same pseudo-benchmark.

Only one-quarter of the funds in our sample select the best-matched benchmark as their prospectus benchmarks. For the remaining three-quarters, a better alternative is available: replacing the stated benchmark with the best-matched benchmark would raise the return correlation by 0.12 on average, indicating that a closer representation of the fund's investment portfolio is readily available. To quantify the effects of misbenchmarking, we conservatively

classify a fund as *meaningfully misbenchmarked* if the best-matched alternative improves the benchmark correlation by at least 0.10 relative to the self-declared benchmark. Using this definition, 33% of funds are meaningfully misbenchmarked. These funds earn benchmark-adjusted returns that are 2.5% higher over a 36-month period than those of control funds. Yet on an unadjusted basis, they underperform the control group by more than 1.2%, indicating that the appearance of their benchmark outperformance is entirely driven by benchmark selection. Meaningfully misbenchmarked funds also charge significantly higher expense ratios. As a result, investors misled by inflated benchmark-adjusted return pay higher fees without receiving better performance.

Greater exposure to credit risk appears to be the primary source of inflated benchmark-adjusted performance. To show this, we compare the sector weights implied by each fund's self-declared benchmark with those of the pseudo-benchmark that best matches the fund's returns. For stated benchmarks, the average best-matched pseudo-benchmark assigns weights of 82.1% to the Agg, 7.8% to investment-grade corporate, and 10.2% to high yield. By contrast, the average pseudo-benchmark that best matches fund returns assigns only 41.2% to the Agg, but 23.9% to investment-grade corporate and 34.9% to high yield. This gap in composition reflects the industry's heavy reliance on the Agg despite its inadequate representation of the credit risk embedded in many funds' portfolios. As a result, active bond funds mechanically boost reported benchmark-adjusted performance by holding bonds with substantially greater exposure to systematic credit risk than their stated benchmarks imply.

We further study the incidence, determinants, and effects of benchmark switching. Using the history of benchmarks reported in fund prospectuses, we find that 20% of funds in our sample switch benchmarks at least once, and that 12% undertake material benchmark changes that alter the credit- or term-risk exposure of the benchmark. On average, these material switches move funds toward benchmarks with lower risk and lower historical returns, making past benchmark-adjusted performance appear stronger. Consistent with strategic benchmark choice, we show that funds are more likely to implement material benchmark

changes after periods of underperformance. We also find that switching to a lower-risk benchmark or one with weaker past returns generates a temporary increase in investor flows.

We conclude by discussing why benchmark misalignment remains pervasive among active bond funds even as it has diminished among active equity funds. First, we test whether the flexibility of broad bond benchmarks reflects the need to accommodate dynamic portfolio management. The evidence does not support this view. Instead, we find that key fund characteristics —such as weights in corporate bonds and U.S. Treasury securities, as well as duration —are highly persistent over time. Second, we examine whether institutional oversight and competition from passive funds, factors that Chen et al. (2025) argue have improved benchmarking in equity funds, mitigate incentives to misbenchmark in fixed income. We find that neither the presence of institutional investors nor passive competition weakens the relation between fund flows and benchmark-adjusted returns. Moreover, the likelihood of switching benchmarks, which we show is typically toward easier-to-beat benchmarks, is higher in categories with greater passive competition. We argue that the industry's heavy reliance on a single index, the Agg, may be an important force behind the persistence of misbenchmarking in fixed income.

Our study extends the strategic benchmark mismatch phenomenon documented in equity mutual funds (e.g., Sensoy, 2009; Chen et al., 2025) and pension funds (Augustin, Binfarè, and Fermand, 2024) to fixed income markets. Our findings have substantial implications for investor decision-making and fund selection in active fixed-income management. In equity mutual funds, the widespread misuse of the S&P 500 first documented by Sensoy (2009) has improved in recent years (Chen et al., 2025). Nevertheless, Mullally and Rossi (2025) find that equity managers use benchmark changes to give the appearance of materially improved performance to attract investor flows. Our evidence suggests that active fixed income funds continue to be misbenchmarked at levels similar to those documented in the early equity literature and that benchmark changes observed provide no evidence of a trend

towards better aligned benchmarks over time.[4] Instead, the evidence suggests that bond fund managers strategically select low-risk, easier-to-beat benchmarks that are poorly matched to the risk characteristics of their portfolios. The industry's heavy dependence on the Agg exemplifies this pattern, as funds with substantial allocations to higher-risk corporate and high yield bonds systematically default to an investment-grade benchmark that is overweight U.S. Treasury securities, and understates their true credit risk exposure. Taken together, our findings have substantial implications for investor decision-making and fund selection in active fixed-income management.

We also contribute to the literature on bond fund performance manipulation. Cici, Gibson, and Merrick Jr (2011) show that bond funds mark the valuation of hard-to-value bonds to their advantage and smooth returns. While Chen, Cohen, and Gurun (2021) find that bond funds misrepresent the true risk profile of their holdings in an effort to be assigned to an easier-to-beat Morningstar category. We document another distinct manner in which bond funds obfuscate the presentation of fund returns to attract larger flows and charge higher fees: by selecting easier-to-beat benchmarks.

I. Data

To construct our sample of active bond funds, we use the CRSP Survivorship-Bias-Free Mutual Funds database to identify all funds that hold at least part of their portfolio in U.S. corporate bonds. We exclude all ETFs and passive bond funds, identified using the CRSP index fund flag and fund names. After matching to benchmark returns and characteristics, our sample includes 1,184 active fixed income mutual funds between 2010 and 2024.

We collect fund-level data from CRSP and Bloomberg, including returns, fund age, fund assets, expense ratio, turnover, and tracking error. We compute the growth of assets to fund i in month m as

[4] 57% of all switches post 2019 are to benchmarks that improve a funds' benchmark-adjusted returns compared to only 35% before 2019.

$$Flow_{i,m} = \frac{TNA_{i,m} - TNA_{i,m-1} \cdot (1 + R_{i,m})}{TNA_{i,m-1}}. \tag{1}$$

Next, we identify each fund's primary benchmark index using funds' semi-annual shareholder reports filed with the SEC and the Morningstar database. For additional tests, we also collect secondary benchmarks where provided. We collect index-level characteristics from index providers, including monthly time series of the number of bonds, number of corporate bonds, percentage of the index allocated to each of four bond sectors (Treasury, government-related, corporate, and securitized), yield to maturity, duration, coupon, and credit rating. We compute each fund's performance relative to its stated primary prospectus benchmark (PPB) as

$$PPB\ Adj\ Return_{i,m} = Fund\ Return_{i,m} - PPB\ Return_{m}, \tag{2}$$

where *Fund Return* is its gross or net reported monthly return.

In order to analyze fund holding characteristics, we obtain bond-level characteristics primarily from ICE Pricing and Reference Data, supplemented by Mergent's FISD, ICE Indexes, TRACE, and name searches to identify bond sectors (Treasury, government-related, corporate, and securitized). We calculate portfolio-weighted characteristics of bond funds across the following dimensions: percentage of holdings in corporate bonds, percentage of holdings in Treasury securities, percentage of holdings in Securitized and Government related bonds, average duration of portfolio holdings, average yield-to-maturity of portfolio holdings, average coupon of portfolio holdings, and average credit rating of portfolio holdings. To estimate these fund-level characteristics, we require non-missing data for at least 90% of each fund's ex-cash dollar holdings, except for credit ratings, where we require coverage of at least 70% of ex-cash dollar holdings. To validate our methodology, we compare our calculated statistics to Morningstar fund-level characteristics, which are available beginning in 2017.

Over this overlapping period, our estimates for fund-level average duration and credit rating have correlation with Morningstar's reported values of 0.92 and 0.98, respectively, confirming the accuracy of our approach. Summary statistics are presented in Table I.

[Insert Table I here]

II. Stylized Facts about Bond Fund Benchmarking

To understand benchmarking practices by fixed-income mutual funds, we first examine variation in benchmark selection across all funds. Figure 1A displays the frequency with which each index is designated as a fund's benchmark. We find that 662 of 1,184 funds (56%) in our sample select the Bloomberg U.S. Aggregate Index as one of their their benchmarks. The prevalence of the Agg as a benchmark for fixed-income funds is even higher than the widespread use of the S&P 500 among equity funds documented by Sensoy (2009). During his 1994–2004 sample period, 44% of equity funds were benchmarked to the S&P 500.

[Insert Figure 1 here]

The fact that a large number of funds select the Agg as their primary benchmark is not inherently indicative of misbenchmarking. The Agg is a "broad-based flagship benchmark that measures the investment grade, US dollar-denominated, fixed-rate taxable bond market" (Bloomberg, 2025); funds providing similar exposures could reasonably select the Agg as their benchmark. To examine whether such benchmarking is appropriate, we tabulate the Lipper categories of funds that declare the Agg as their benchmark. Figure 1B shows that the majority of Agg-benchmarked funds are categorized as core or core-plus, categories with issuer types and risk exposures similar to those of the Agg. However, a significant number funds with other Lipper categories, including corporate bond and high yield funds, also declare the Agg as their benchmark. These Lipper categories represent market segments not included in the Agg and carry materially different risk-return profiles. This provides initial

evidence that reliance on the Agg as a benchmark may be inappropriate for a substantial subset of funds.

In Figure 2 we further examine the variation in investment strategies among funds benchmarked to the Agg. Here, we take all funds benchmarked to the Agg in December 2023 and plot each fund's percentage allocation to corporate bonds (y-axis) against the average duration of its holdings (x-axis). Each circle represents an individual fund benchmarked to the Agg and the "x" mark represents the Agg itself.

Focusing on corporate bond holdings, the Agg allocates approximately 25% to corporate bonds. However, among active bond funds using the Agg as their benchmark, there is wide variation in corporate bond allocations. Some funds hold less than 20% of their portfolios in corporate bonds, while others allocate more than 80%. This dispersion highlights the heterogeneity in portfolio composition across Agg-benchmarked funds, suggesting that some may be more appropriately benchmarked to alternative indices. We observe similar, though less pronounced, variation in portfolio duration relative to the Agg's duration.

[Insert Figure 2 here]

We systematically examine the characteristics of funds benchmarked to the Agg, including percentage holdings of corporate bonds, percentage holdings of Treasuries, average credit rating, and average duration, and tabulate their year-by-year distributions relative to the characteristics of the Agg. Figure 3 shows that the portfolio characteristics of Agg-benchmarked funds tend to differ substantially from those of the Agg, with wide dispersion across funds. For example, Panel A shows that funds benchmarked to the Agg hold, on average, a higher percentage of corporate bonds than the Agg itself. The average over-allocation to corporate bonds has declined over time, from 16% in 2010 to 8% in 2023. However, the distribution of corporate bond over-allocation remains wide across funds, with an interquartile range of more than 21%.

[Insert Figure 3 here]

Panel B of Figure 3 shows Agg-benchmarked funds' allocations to Treasury securities relative to the Agg itself. On average, funds benchmarked to the Agg under-allocate to Treasuries by approximately 20 percentage points relative to the index. This under-allocation has remained stable throughout our sample period, with wide cross-fund variation. Panel C shows that Agg-benchmarked funds tend to hold securities with lower credit ratings than the index, although this difference has narrowed toward the end of our sample period. Finally, Panel D shows that the average duration of holdings in Agg-benchmarked funds is lower, on average, than that of the Agg. Duration also exhibits wide dispersion across Agg-benchmarked funds.

As an illustrative example, consider the J.P. Morgan Corporate Bond Fund. Panel A of Figure 4 shows the fund's legal prospectus. The J.P. Morgan Corporate Bond Fund primarily invests in corporate bonds; its investment strategies stipulate that at least 80% of the assets should be in corporate bonds. The prospectus states that "[t]he Fund is managed relative to the Bloomberg U.S. Corporate Index". However, when the fund presents its annual total return on the right hand side of the prospectus, it lists both Bloomberg U.S. Aggregate Index and Bloomberg U.S. Corporate Index as comparisons. Moreover, in the fund's marketing material ("Fact Sheet"), shown in Panel B of Figure 4, the relative performance graph only compares the fund return to the Agg, while excluding the return of the stated corporate benchmark altogether. Notably, the fund's performance lags the Bloomberg U.S. Corporate Index over the 1-, 5-, and 10-year horizons, while comfortably outperforming the Agg over the same periods.

[Insert Figure 4 here]

III. Main Results

A. Strategic incentives for misbenchmarking

We first investigate whether fund managers have incentives to select mismatched benchmarks. We posit that mutual fund flows are positively influenced by benchmark-adjusted

returns, generating incentives for funds to select poorly matched benchmarks to make their performance appear more attractive to investors. To estimate the relationship between investor flows and PPB-adjusted performance, we estimate the following regression specification:

$$Flow_{i,m} = \alpha_m + \beta_1 PPBAdj.Ret_{i,m-1} + \beta_2 Unadj.Ret_{i,m-1} + \gamma X_{i,m-1} + \epsilon_{i,m}, \quad (3)$$

where $PPBAdj.Ret$ is a fund's net return adjusted by the return of its primary prospectus benchmark and $Unadj.Ret$ is the unadjusted fund net return, both measured over a 36-month window.[5] All performance metrics are measured over a trailing 36-month window. The vector X includes fund characteristics: the natural logarithm of fund assets, the natural logarithm of fund age, the expense ratio, and the turnover ratio. Standard errors are double-clustered by fund and time.

[Insert Table II here]

Table II displays the results from the flow-performance regressions. In Columns (1) and (2), we show that a fund's unadjusted return and benchmark-adjusted return are separately positively related to fund flows after controlling for Lipper category-by-date, and Morningstar star rating-by-date fixed effects. For example, in Column (2), a one standard deviation increase in 36-month $PPBAdj.Ret$ is associated with a 0.31% increase in monthly fund flow (calculated as 0.056 coefficient multiplied by 5.64% standard devation in $PPBAdj.Ret$). In the next few specifications, we include both $PPBAdj.Ret$ and fund unadjusted return in horse-race regressions. The results reported in column (3) shows that, when both variables are included as regressors, $PPBAdj.Ret$ is positively associated with future fund flows, while the coefficient on fund unadjusted return is insignificant. This suggests that investors

[5]Note that the primary prospectus is current as of July 2025, when the data was downloaded from Morningstar. This may cause mismeasurement of $PPBAdj.Ret$ for funds that have changed their benchmark over time. To confirm that these funds are not driving our results, we re-run the tests in this section after dropping the sample of funds that have switched benchmarks, as discussed below in Section III.D. The results in this section are robust to this change.

rely heavily on (self-reported) benchmark-adjusted return of bond funds in allocating their capital. In column (4), while Morningstar ratings explain some variation in fund flows, consistent with Ben-David, Li, Rossi, and Song (2022) and Evans and Sun (2021), the positive association between benchmark-adjusted returns and fund flows remains statistically significant, though with slightly attenuated magnitude. The coefficient on unadjusted fund return is also positive, but meaningfully smaller than the coefficient on $PPBAdj.Ret$. Finally, in column (5) we include fund fixed effects and in column (6) we use a Fama-MacBeth specification. Our results remain robust to these specifications. The coefficient on tracking error is either statistically indistinguishable from zero, or negatively associated with fund flows.

Given that bond fund flows respond positively to PPB-adjusted returns, we decompose PPB-adjusted returns into two components for fund i in Lipper category l in month m:

$$PPBAdj.Ret_{i,l,m} = (FundRet_{i,m} - \overline{CatRet_{l,m}}) + (\overline{CatRet_{l,m}} - PPBRet_{i,m}). \quad (4)$$

The first term on the right-hand side represents fund performance relative to its Lipper category peers, capturing returns from active management strategies such as security selection or reaching for yield that deliver outperformance above the category average. The second term represents category-level outperformance relative to the fund's declared benchmark, capturing returns that arise when a fund selects a benchmark with mismatched characteristics (e.g., lower risk exposure) relative to the fund's Lipper category.

In Table III, we examine whether investors can differentiate the sources of a fund's benchmark-adjusted returns, and whether fund flows respond differently to within-category fund performance versus category-benchmark mismatch. If fund investors can "see through" funds' strategic choice of benchmarks and compare fund returns against category peers, we should expect the term $(\overline{CatRet_{l,m}} - PPBRet_{i,m})$ to have no association with with fund flows.

[Insert Table III here]

The results in Table III show that both $(FundRet_{i,m} - \overline{CatRet_{l,m}})$ and $(\overline{CatRet_{l,m}} - PPBRet_{i,m})$ are significantly positively associated with fund flows. This suggests that fund investors respond positively to benchmark-adjusted performance generated by category-benchmark mismatch, above and beyond flows due to fund outperformance. The significant positive coefficient on $(\overline{CatRet_{l,m}} - PPBRet_{i,m})$ provides evidence that bond funds have incentives to strategically game benchmark selection to attract capital flows.

B. Mismatched and corrected benchmarks

In this section we examine the extent to which active bond funds select benchmarks that do not accurately reflect their risk profiles and whether such misbenchmarking generates favorable benchmark-adjusted performance. Following Sensoy (2009), we calculate the correlation between a fund's gross monthly returns and its chosen benchmark's monthly returns over rolling 36-month periods, and compute the average correlation over the full time series.[6]

We then similarly compute the correlation between the fund's gross monthly returns and the returns for two universes of potential alternative benchmarks. First, we compile 30 fixed income benchmarks which are used by at least five funds in our sample.[7] The selection of a subsample of frequently used benchmarks mitigates concerns of multiple comparisons (Sensoy, 2009), and prevents us from spuriously identifying a fund as misbenchmarked because there is a closely related index with marginal correlation improvement. The list of benchmarks is found in the Appendix. For each of the funds in our sample, we denote the best-matched matched benchmark as the alternative with the highest correlation to a fund's gross returns. In Table IV, we show statistics for funds whose correlation is highest with

[6]The majority of funds report a single benchmark. However, some funds report two or more benchmarks for unclear reasons (Mullally and Rossi, 2025). To be conservative in this analysis, when a fund reports multiple benchmarks, we designate the benchmark with the highest correlation with fund returns as the fund's "primary" benchmark.

[7]Note that we eliminate some benchmarks that are used by more than five funds if there is a close substitute; for example, the Bloomberg US HY 2% Issuer Cap Index is used by 37 funds while the Bloomberg US Corporate HY Index is used by 25 funds. We include the former and exclude the latter.

its chosen benchmark (properly benchmarked) in Panel A, and for funds whose correlation is highest with an alternative benchmark (misbenchmarked) in Panel B. Finally, in Panel C, we compute statistics for funds with different levels of misbenchmarking, as measured by the correlation improvement between their best matched and chosen benchmarks.

[Insert Table IV here]

Of the 1,184 analyzed funds, only 325 (27%) exhibit the highest correlation with their declared primary prospectus benchmark. Panel A of Table IV shows that, for these funds, their return has an average correlation of 0.95 with their declared benchmarks. In Panel B, we examine the remaining funds, whose highest-correlated benchmark is different from their declared benchmark. The correlation between fund returns and the self-declared benchmark is 0.73, while the correlation with the best-matched alternative benchmark is 0.88; switching to the best-matched benchmark would improve the average correlation by 0.15. Despite being used by 662 funds in our sample at either the primary or secondary chosen benchmark, the Bloomberg US Aggregate Index is the best-matched alternative for only 38 funds.

Moreover, funds appear to strategically select benchmarks they can outperform more easily. For misbenchmarked funds, Panel B of Table IV shows that best-matched benchmark returns exceed declared benchmark returns by an average of 9.62 basis points (bps) per month. When adjusted by their declared benchmarks, active bond funds generate average gross monthly returns of 9.79 bps and net monthly returns of 4.25 bps. However, when adjusted by their best-matched benchmarks, active bond funds generate gross benchmark-adjusted returns of -0.16 bps, and net benchmark-adjusted returns of -5.71 bps per month.

There is considerable variation in the degree of misbenchmarking. In Panel C of Table IV, we examine different set of funds for which the best-matched alternative benchmark would improve the return correlation relative to the self-declared benchmark by various thresholds. A larger threshold indicates more severe misbenchmarking. For 500 funds with correlation improvements of at least 0.05, declared benchmarks underperform best-matched benchmarks by an average of 14.49 bps per month, substantially enhancing reported benchmark-adjusted

returns and potentially influencing investor flows. While gross returns relative to declared benchmarks appear economically meaningful for this group (12.59 bps per month), adjusting fund returns by best-matched benchmark returns yields negative benchmark-adjusted performance: average gross adjusted returns of -2.44 bps and net adjusted returns of -8.08 bps per month.

For subsets of funds where best-matched benchmarks improve return correlation by at least 0.10, 0.25, and 0.50, we find increasingly strong evidence that benchmark mismatch inflates reported benchmark-adjusted returns. For example, among funds with potential correlation improvements of at least 0.50, the average fund reports gross outperformance of 20.73 bps per month relative to its declared benchmark while underperforming its best-matched benchmark by 6.97 bps per month on a gross basis. The underperformance is more pronounced on a net basis, averaging -12.69 bps per month. This analysis provides robust evidence of widespread strategic benchmark selection in fixed-income funds.

While using the complete set of primary benchmarks declared by sample funds is an intuitive choice for alternative benchmarks, we also recognize that funds may invest a significant portion of their holdings in asset classes not covered by their best-matched benchmark from this universe. In such cases, selecting the best-matched benchmark from this limited set may not fully capture a fund's risk profile and invoke regulatory compliance risk. As a second universe of alternative benchmarks, we construct a suite of pseudo-benchmarks that allow for hybrid exposures across credit quality segments:

$$Pseudo_{b,t} = \omega_{b.AGG} R_{AGG,t} + \omega_{b,IG} R_{IG,t} + \omega_{b,HY} R_{HY,t}. \tag{5}$$

For each pseudo-benchmark b, the weights $\omega_{j,b}$ vary across three market indices: the Bloomberg U.S. Aggregate Index (AGG), the Bloomberg U.S. Corporate Index (IG), which includes only investment-grade corporate bonds, and the ICE BofA U.S. High Yield Index (HY), which includes only non-investment-grade bonds. We vary the weights in increments of 10% across the three indices, subject to the constraint that weights sum to one. In

addition, we include an equally weighted index. This generates 67 pseudo-benchmarks in total. This universe of alternative benchmarks allows for hybrid levels of credit risk and introduces high-yield exposure beyond the investment-grade universe of the Agg.

We calculate the correlation between each fund's gross returns and all 67 pseudo-benchmark returns. We also calculate the correlation between each fund's primary prospectus benchmark returns and all pseudo-benchmark returns. If the same pseudo-benchmark exhibits the highest correlation with both the fund's gross returns and its primary prospectus benchmark returns, we classify the fund as properly benchmarked. By this criterion, 292 of the 1,184 sample funds (25%) are properly benchmarked, shown in Panel A of Table V. These funds have relatively high correlation with their chosen primary prospectus benchmark (0.93).

[Insert Table V here]

For the remaining 75% of the funds, we classify them as misbenchmarked and tabulate three metrics in Table V: (i) the correlation between fund returns and primary prospectus benchmark returns, (ii) the correlation between fund returns and best-matched pseudo-benchmark returns, and (iii) the correlation improvement, defined as the difference between (ii) and (i). On average, switching to the best-matched pseudo-benchmark increases the correlation with fund returns by 0.12. The median improvement is 0.05.

Examining the returns of primary prospectus benchmarks and best-matched pseudo-benchmarks, we again find that best-matched pseudo-benchmarks outperform declared benchmarks by an average of 9.17 bps per month. By selecting their declared benchmarks instead of best-matched pseudo-benchmarks, active bond funds thus generate substantially inflated benchmark-adjusted returns. For example, as shown in Table V, when adjusted by their declared benchmarks, sample funds' average gross and net benchmark-adjusted returns are 10.33 bps and 4.95 bps per month, respectively. However, when adjusted by best-matched pseudo-benchmark returns, average gross adjusted returns fall to -2.31 bps and average net adjusted returns fall to -7.69 bps per month.

Finally, by comparing the weights ω_{AGG}, ω_{IG}, and ω_{HY} of the pseudo-benchmark best matched to a fund's returns against the weights of the pseudo-benchmark best matched to the fund's primary prospectus benchmark returns, we demonstrate how credit risk exposure drives the return differential between declared and best-matched benchmarks. Table V shows that the pseudo-benchmark best matched to primary prospectus benchmark returns has mean (median) weights of 82.1% (100%) on the Agg, 7.8% (0.0%) on the investment grade corporate index, and 10.2% (0.0%) on the high yield index. In contrast, the pseudo-benchmark best matched to fund returns has mean (median) weights of 41.2% (40%) on the Agg, 23.9% (20%) on investment grade, and 34.9% (25%) on high yield. Therefore, we take this as evidence that active bond funds systematically select benchmarks that understate their corporate bond holdings and credit risk exposure, and as a result, are able to claim positive benchmark-adjusted performance.

C. Misbenchmarked funds

Misbenchmarked funds differ systematically from properly benchmarked funds across several dimensions. In Table VI, we present a comparison of fund characteristics between "meaningfully misbenchmarked" and properly benchmarked funds. We define a fund as meaningfully misbenchmarked if switching to its best-matched alternative benchmark would improve the correlation with fund returns by at least 0.10. The control group consists of funds that either exhibit the highest correlation with their primary prospectus benchmark or have correlation improvements of less than 0.10 from their best-matched benchmark.[8]

[Insert Table VI here]

Table VI shows that meaningfully misbenchmarked funds generate significantly higher benchmark-adjusted gross returns (4.65% over 36 months) compared to properly benchmarked funds (2.15%), with this difference driven almost entirely by benchmark selec-

[8] In the Appendix, we repeat the following analyses using the pseudo-benchmark criteria to determine misbenchmarked funds; the interpretation is unchanged.

tion rather than managerial skill. Specifically, for meaningfully misbenchmarked funds, best-matched alternative benchmarks outperform their declared benchmarks by 5.56% over 36 months. If these funds were benchmarked to their best-matched alternatives, their benchmark-adjusted gross returns would fall to -0.91%. In contrast, if the control funds were benchmarked to their best-matched alternatives, their benchmark-adjusted returns would decline only modestly to 1.29%.

When we examine the portfolio characteristics of meaningfully misbenchmarked and properly benchmarked funds, we find that meaningfully misbenchmarked funds on average have very slightly higher yields to maturity (3.96% vs 3.82%), and lower allocations to corporate bonds as compared unconditionally to properly benchmarked funds. This suggests that meaningfully misbenchmarked funds are not in general "reaching for yield" relative to other funds in general; rather, they strategically select benchmarks that understate the risk (and hence expected return) of their portfolios relative to their chosen benchmarks.

To further distinguish bond fund misbenchmarking from reaching-for-yield behavior (for example, as documented by Choi and Kronlund, 2018), we regress funds' PPB-adjusted gross return on both an indicator for meaningfully misbenchmarked ($Misbenchmarked$) and the funds' yield to maturity:

$$\begin{aligned} PPB\text{-}adj\ Gross\ Ret_{i,m} = \alpha_{l,m} + \beta_1 Misbenchmarked_i + \beta_2 YTM_{i,m-1} \\ + \beta_3 (Misbenchmarked_i * YTM_{i,m-1}) + \gamma X_{i,m-1} + \epsilon_{i,m} \end{aligned} \tag{6}$$

where $Misbenchmarked$ is equal to one for funds where the correlation improvement of the best matched benchmark improves the returns correlation by more than 0.10 relative to the primary prospectus benchmark, and zero otherwise.[9] We include Lipper cateogry-by-date fixed effects, $\alpha_{l,m}$ in regressions, effectively comparing two funds from the same Lipper

[9]In unreported tests, we find that the results of VII are robust to using different cutoffs for severe misbenchmarked funds. Using a correlation improvement cut-offs of .05, .15., and .25 we find that the results are similar, both economically and statistically.

class as determined by their quantitative methodology in the same month. The control variables $X_{i,m-1}$ include fund total net assets, fund age, expense ratio, turnover and fund flow. Standard errors are double-clustered at the fund and month level.

[Insert Table VII here]

In Panel A, Columns (1) and (2) of Table VII show that funds with mismatched benchmarks generate significantly higher benchmark-adjusted gross returns compared to other funds in the same Lipper category and those with the same best-matched benchmark, outperforming by approximately 7 bps per month. In columns (3) and (4), we include fund yield-to-maturity, and meaningfully misbenchmarked funds continue to generate significantly higher benchmark-adjusted gross returns even after controlling for these risk measures. Importantly, the coefficient on the interaction between the risk measure and the meaningfully misbenchmarking indicator is statistically insignificant, suggesting that meaningfully misbenchmarked funds generate strong benchmark-adjusted performance independent of reaching-for-yield behavior, consistent with strategic benchmark selection. In the final two columns of Panel A of Table VII, we examine whether meaningfully misbenchmarked funds generate higher returns on an unadjusted basis. We replace the dependent variable in Equation (6) with unadjusted monthly gross fund returns. As reported in columns (5) and (6), there is no evidence that meaningfully misbenchmarked funds generate statistically different unadjusted gross returns relative to other funds in the same Lipper category. This is important as the results in Columns (1) to (4) do not rule out the possibility that these funds beat their mismatched benchmark, while also beating their peer funds. Columns (5) and (6) provide evidence that these funds do not beat their peer funds, and that the superior benchmark adjusted performance is entirely due to the use of a poorly match benchmark.

In Panel B of Table VII, we examine the yield to maturity of meaningfully misbenchmarked and control funds relative to their chosen benchmarks (column (1)) and best-matched benchmarks (column (2)). While meaningfully misbenchmarked funds have significantly higher YTM than control funds for chosen benchmarks, these differences are eliminated

completely when adjusting for corrected benchmarks. Consider for example two funds that benchmarks to the Agg, one of which is correctly benchmarked and one that was better benchmarked to a corporate or high yield index. The meaningfully misbenchmarked fund would be tilted to riskier corporate bonds than the Agg, leading to a higher YTM and the appearance of reaching for yield. However, this meaningfully misbenchmarked fund would have a YTM similar to funds with the same best matched benchmark.

While meaningfully misbenchmarked funds generate similar gross return as compared to other funds, they also charge a higher expense ratio to investors. In Table VIII, we examine the expense ratio of meaningfully misbenchmarked funds relative to other funds. Controlling for fund characteristics and including Lipper category-by-year fixed effects, we find that funds that we classify as meaningfully misbenchmarked —either relative to the universe of observed benchmarks (columns (1) to (3)) or the universe of pseudo benchmarks (columns (4) to (6))—tend to charge a higher expense ratio by the amount of 7.5 to 11.4 basis points per annum as compared to other funds. Therefore, misbenchmarking appears attractive for fund managers: they are able to market better benchmark-adjusted performance, attract higher investor flows, and charge higher fees for their assets under management.

[Insert Table VIII here]

D. Bond fund benchmark changes

In this section, we focus on a subset of bond funds that have changed their benchmark within our sample period. To this end, we manually collect from SEC EDGAR the full time series of each bond fund's self-declared benchmark(s) from their summary prospectuses. Through our sample period, 225 funds (19% of our sample) make at least one change to their benchmark, and there are 396 fund-years with changes.[10] In Figure 5, we show that the number of changes is relatively constant over time, though there is a large spike of funds adding indices

[10]For the purposes of this analysis, we do not distinguish the 'order' in which benchmarks are presented, so a fund that keeps the same set of benchmarks but swaps the primary and secondary benchmark is not counted as a change.

in the final sample year. Related to our previous findings that many misbenchmarked funds use Bloomberg U.S. Aggregate Index as their benchmark, we find that 76 benchmark changes involve a fund adding the Agg as its benchmark, including 39 funds in 2024 alone.

[Insert Figure 5 here]

We next analyze *how* funds change their benchmarks. In particular, we classify all benchmarks into seven categories based on their risk exposure. We then examine whether an added (dropped) benchmark has the same or different risk or term as the existing (remaining) benchmarks. From the least risky to the most risky, these categories include

1. Risk-free: indexes linked to LIBOR or Treasury securities;
2. GovCredit: investment grade indexes that hold U.S. Treasuries, government-related bonds (encompassing non-U.S. sovereigns, government agencies, local authorities, and supranationals), and corporate bonds;
3. Agg: investment grade indexes that hold Treasuries, government-related bonds, corporate bonds and securitized bonds;
4. Credit: investment grade indexes that hold government-related bonds and corporate bonds;
5. Universal: indexes that hold securities in the Aggregate indexes plus high yield corporate bonds;
6. IGCorp: indexes that hold investment grade corporate bonds; and
7. HY/Equity: indexes that hold high yield bonds or equities.

In Table IX, for all fund-years with benchmark changes, we compare the fund's benchmark risk category in year t (displayed in rows) to its benchmark risk category in year $t+1$

(displayed in columns).[11] For example, if a fund has a Credit benchmark in year t and adds an Agg benchmark in year $t+1$, we count it as a "Credit-to-Agg" change. If a fund uses both a Credit and a Universal benchmarks in year t and drops the Universal benchmark in year $t+1$, we count it as a "Universal-to-Credit" change.

[Insert Table IX here]

We first observe that there are a large number of observations on the diagonal, indicating that funds change their benchmark indexes within the same risk category. As an example, consider a fund that drops the ICE BofA US High Yield Constrained Index, but retains Bloomberg's U.S. High Yield - 2% Issuer Cap. In later analysis, we term these benchmark changes as "negligible " changes (assuming there is also no change to the time to maturity focus of the benchmarks, e.g., from intermediate to short). On the off-diagonal elements, we show fund benchmark changes in which the risk category of the benchmarks changes. We identify these changes to the risk or term of the benchmark as "significant" changes. Of note, here we show that across almost all categories, the most common switch is a change to an Aggregate benchmark (which is dominated by, but not limited to, the Bloomberg U.S. Aggregate Bond Index). By ranking the categories according to increasing risk, we also show that there are slightly more funds that move their benchmarks toward lower risk categories (77 moves below the diagonal line) than toward higher risk categories (69 moves above the diagonal line). The largest off-diagonal category is high yield funds moving toward an Aggregate benchmark. This again is further evidence of industry-wide reliance on a single benchmark even when the systematic risk of the benchmark does not match the funds holdings.

Why do funds change their benchmarks? The most straightforward explanation for a benchmark change is a concurrent change to the funds' investment strategy. Here, a significant shift in strategy would require a change in the fund benchmark to match the new risk

[11] There are many funds that perform benchmark changes in two year increments, with a typical pattern of benchmarks of index A in year t-1, indexes A and B in year t, and index B in year t+1. For purposes of these transition matrices, these two-year changes are counted as one change.

characteristics of the portfolio. However, when we compare the Lipper investment category of benchmark-switching funds, we only find 22 instances (out of 225 changes) where the category is different before and after the benchmark change. As a result, we do not think strategy switches are driving our results.

Another possible reason to change benchmarks is to improve the optics of the fund's performance. Simply put, by adding a benchmark with lower returns than current benchmarks or dropping a benchmark with higher returns than remaining benchmarks, a fund can improve the reported benchmark-adjusted return (Mullally and Rossi, 2025). In order to examine this possibility, we define $\Delta BenchmarkRet$ as follows. For benchmark additions:

$$\Delta BenchmarkRet_{p,t} = Return\_Added_{p,t} - Return\_Existing_{p,t}. \tag{7}$$

For benchmark deletions:

$$\Delta BenchmarkRet_{p,t} = Return\_Remaining_{p,t-1} - Return\_Dropped_{p,t-1}. \tag{8}$$

Note that there is a slight difference in timing for measuring return differences. For benchmark additions, we can observe the returns of added and existing benchmarks in the year of benchmark addition. However, for benchmark deletions, we can only observe the return of both dropped and remaining benchmarks as reported by the fund in the year prior to the deletion. Across benchmark additions and deletions, the interpretation on the sign of $\Delta BenchmarkRet$ is the same: a negative number indicates that the fund has switched benchmarks in a manner that improves their self-reported benchmark-adjusted return.

In Figure 6, we display the $\Delta BenchmarkRet$ for all benchmark additions (top panel) and benchmark deletions (bottom panel) in our sample. Consistent with how benchmark returns and fund returns are displayed in fund prospectuses, we plot 1-year, 5-year, and 10-year return separately.

[Insert Figure 6 here]

Across all time periods, mean $\Delta BenchmarkRet$ are significantly negative, indicating that on average funds switch benchmarks in a way that improves their fund's perceived performance. For example, the top left sub-figure shows that, when a fund adds an additional benchmark, the 1-year return of the added benchmark is on average 1.1% lower than the fund's existing benchmark. The bottom left figure suggests that, when a fund drops one of its (multiple) existing benchmarks, the return of the remaining benchmark(s) is on average 0.5% lower than the dropped benchmark.

We further split the sample of benchmark changes into material changes and negligible changes. We deem a benchmark change as significant if a fund adds or drops a benchmark that moves into either a different risk or term category. In Figure 7, we plot the 1-year $\Delta BenchmarkRet$ for material changes in the left column, and negligible changes in the right column. Funds with a significant benchmark addition have a mean $\Delta BenchmarkRet$ of -1.7% (statistically significant at the 1% level), compared to a non-significant mean of 0.1% for negligible benchmark additions. For benchmark deletions, the average $\Delta BenchmarkRet$ is similar across significant and negligible changes.

[Insert Figure 7 here]

If funds strategically add or drop benchmarks to improve their perceived performance, we hypothesize that the funds most likely to do this would be those that have underperformed in the past. We examine this possibility in Table X. To test the relationship between past performance and the likelihood of switching we use the following linear probability model on a fund's decision to switch based on historical performance:

$$D(Switch_{i,t}) = alpha_{c,t} + \beta * 3yr\_return_{i,t} + controls_{i,t} + \epsilon_{i,t}, \tag{9}$$

where $D(Switch_{i,c,t})$ is a dummy that takes on a value of 1 if fund i in Lipper category c changes benchmarks in month t. We measure historical returns on both a gross basis, and adjusted for the performance of the fund's best-matched pseudo benchmark (as described in

section III.B). For the sample of control funds, we include only those funds that never make a benchmark change that are in the same Lipper category as the switching fund; we include Lipper category-by-month fixed effects to compare within category fund performance. Fund-level controls include expense ratio, turnover ratio, the natural logarithm of fund age and the natural logarithm of lagged fund assets. Standard errors are clustered at the year level.

[Insert Table X here]

The results reported in Table X show that, consistent with our hypothesis, funds are more likely to change their benchmark if they have underperformed their Lipper category peers over the previous three years. For all fund switches in columns (1) and (4), this relationship is stronger using the pseudo benchmark adjusted returns, compared to using actual fund returns. However, as we noted above, not all benchmark switches are equal. Next, we directly examine the relationship between past performance and the nature of the benchmark change. In columns (2) and (5), we look at *Material* switches, which take a value of one for material switches, and zero for negligible changes and control funds that never change their benchmark. In columns (3) and (6), *negligible* switches takes the value of one for negligible switches, and zero for material switches and control funds. For negligible changes, we find no significant relationship between past performance and the likelihood of switching benchmarks. However, for the material switches, we find a strong negative coefficient on both past three year raw and risk-adjusted fund returns. In other words, poorly performing funds are more likely to switch benchmarks in a significant manner, consistent with a strategic motive that may obfuscate poor past performance.

We next examine if these strategic changes are successful in increasing fund flows. To do this, we examine the dynamics of monthly fund flows around the changes of bond fund benchmarks. For each benchmark change event in month t, we consider month $t-36$ to month $t-13$ as the pre-period and month $t+1$ to $t+24$ as the post-period.[12] The control

[12] As we do not know the precise month in which a fund changes its benchmark within a year, we discard the observations between month $t-12$ and $t-1$.

sample includes funds that never make change to their benchmarks in our sample period. We then estimate the effect of benchmark changes on fund flows using a difference-in-difference (diff-in-diff) framework:

$$Flow_{i,t} = \alpha_i + \gamma_t + \beta Post_t \times BmrkChg_i + Controls + \epsilon_{i,t}. \tag{10}$$

Since we include both fund fixed effects and event-by-time fixed effects in the regression, coefficients of the base level of $Post$ and $BmrkChg$ are absorbed. In some specifications, we interact our coefficient of interest with variables that separate the benchmark switches into sub-categories, including whether the fund switches to a lower risk benchmark, or whether the $\Delta BenchmarkRet$ of the new index is negative (e.g., making benchmark returns easier to beat).

[Insert Table XI here]

In column (1) of Table XI, we find that, on average, fund flows are relatively unchanged for a fund that changes its benchmark. As we documented earlier, we may not expect to see a flow response to all switches, as about 40% of our switches are negligible and do not materially change the risk or return of the benchmark. However, as column (2) indicates, if a fund switches to an alternative benchmark with a lower risk exposure profile, investor flows in the ensuing 24 months are significantly relatively higher than the flows prior to the change of benchmarks. This is consistent with the notion that investors are attracted to the fund's seemingly higher benchmark-adjusted return after the fund switches to an easier-to-beat index. Similarly, in column (3), funds that add or drop an index in a way that makes performance appear better experience significantly higher relative flows after the switch. In column (4), we include both lower risk and lower return, as well as an interaction term for changes that include both of these terms; here we see that the significant increase in flows is due to the switch to lower risk benchmarks. This represents another significant difference between misbenchmarking in fixed income and equities. Equity fund managers

cannot reliably predict when a particular style or size tilt will be easier to beat, so they rely on switching to benchmarks with lower past returns (Mullally and Rossi (2025)). Due to the clear delineation in bond markets of risk categories, and more predictable return differences, fund managers are able to shift to lower risk benchmarks that they will also be more likely to beat in the future, regardless of past returns.

We plot these fund flows in Figure 8. Noticeably, funds that make meaningful switches to a lower risk benchmark (blue long dash line) have much higher post-switch flows, compared to all switches (gray solid line).

[Insert Figure 8 here]

In this section we document multiple novel facts about active bond funds that make benchmark switches. First, while equity funds have moved toward more accurate benchmarks (Chen et al., 2025), active bond funds have continued to switch to lower-return, easier-to-beat benchmarks, even in the later part of our sample. Second, our results are driven by significant benchmark switches, where a fund moves to a benchmark in a different risk or term category, as opposed to negligible benchmark switches. It is the material switches that move funds' benchmark returns lower on average, and these switches are done by funds with poor historical performance. And finally, it is only funds that make material changes who experience higher fund flows.

IV. Discussion: The Persistence of Misbenchmarking

The pervasive benchmark misalignment we document among active bond funds raises important questions about its underlying drivers and the forces that might constrain it. In this section, we discuss several reasons why misbenchmarking persists in fixed-income mutual funds in contrast to equity funds and examine the limited effectiveness of market-based oversight in disciplining this behavior.

A primary justification for the use of broad-based benchmarks, such as the Bloomberg U.S. Aggregate Bond Index, is that they give managers flexibility to rotate across subsectors and adjust portfolio risk dynamically. Our evidence, however, suggests that such flexibility is not central to the behavior of the funds in our sample. Tables A5 and A6 show that fund characteristics, such as the weight in corporate bonds, the weight in Treasuries, and duration, are persistent over time. For instance, an Agg-benchmarked fund that overweights corporate bonds in one period is highly likely to continue overweighting corporate bonds in the next period. The evidence therefore indicates that deviations from stated benchmarks do not primarily reflect temporary tactical reallocations. Rather, they capture stable and persistent exposure to higher-risk securities relative to benchmarks that are consistently easier to outperform.

Misbenchmarking also appears to be more persistent among active bond funds than in the equity mutual fund sector. For active equity funds, Chen et al. (2025) show that institutional oversight and passive competition mitigate misbenchmarking over time by debiasing investors and disciplining managers' benchmark choices. We test whether these same forces operate in fixed income and find no evidence that they do. In our sample, the average misbenchmarked bond fund holds 60.6% of assets in institutional share classes, compared with 57.8% for the average properly benchmarked fund. Contrary to the sophistication one might expect from professional investors, mismatched bond funds therefore exhibit higher institutional ownership than properly benchmarked funds, suggesting that benchmark selection receives limited scrutiny even from institutional capital.

Similarly, we find no evidence that institutional investors weaken the sensitivity of fund flows to benchmark-adjusted returns, the key channel we study in Table II. In Panel A of Table XII, we test whether the flow-performance relation changes with the level of *Institutional* by interacting benchmark-adjusted returns with institutional share ownership. If greater institutional presence dampens investors' reliance on benchmark-adjusted performance and, in turn, reduces managers' incentives to misbenchmark, we should expect the coefficient on

the interaction between benchmark-adjusted return and institutional share to be negative.

[Insert Table XII here]

Instead, our results suggest that the level of institutional holdings does not attenuate the relationship between flows and benchmark-adjusted returns. Under some specifications, we even find that greater institutional ownership exacerbates the sensitivity.

We also find no evidence that competition from passive investment vehicles mitigates misbenchmarking among active bond funds. In Panel B of Table XII, the interaction between *Passive %* and benchmark-adjusted return is positive. This result indicates that the sensitivity of investor flows to benchmark-adjusted performance does not weaken when more passive alternatives are available within the same investment category. If anything, it becomes stronger. This pattern contrasts with the evidence from the equity market, where Chen et al. (2025) argue that the rise of passive funds makes active managers' attempts to obfuscate performance less effective.

The absence of the usual disciplining forces from institutional investors and passive competition suggests that misbenchmarking toward low-risk benchmarks, especially the Agg, is sustained by entrenched norms within the fixed-income fund sector. The recent rise in funds adopting the Agg as a benchmark points to a systemic preference for a single, highly visible performance hurdle. Whether this preference reflects the perceived complexity of fixed-income markets or investor demand for a universal point of comparison, the industry's reliance on the Agg appears to institutionalize benchmark misalignment.

V. Conclusion

This paper provides systematic evidence of widespread misbenchmarking among active bond mutual funds. We document that these funds choose benchmarks that do not accurately reflect their risk exposures or investment strategies. By strategically selecting lower-risk, lower-return benchmarks, these funds generate inflated benchmark-adjusted returns that

attract significant investor flows despite delivering no genuine risk adjusted performance. While benchmark switches are relatively rare, underperforming funds exhibit a clear pattern of switching to benchmarks with lower risk profiles following poor performance, suggesting intentional gaming rather than benign misspecification. Despite passive and institutional investors driving better matched benchmarks in equity markets, a large percentage of fixed income funds remain misbenchmarked.

Our findings have important implications for investors, asset managers, and regulators. For investors, our results highlight the critical need to look beyond benchmark-adjusted returns when evaluating bond fund performance. The positive relationship between benchmark-adjusted returns and fund flows, even when generated purely through benchmark mismatch, suggests that many investors fail to account for the appropriateness of funds' self-declared benchmarks. For the asset management industry, our evidence that meaningfully misbenchmarked funds charge significantly higher expense ratios while delivering inferior risk-adjusted performance raises concerns about value destruction.

References

Augustin, Niklas, Matteo Binfarè, and Elyas Fermand, 2024, Benchmarking private equity portfolios: Evidence from pension funds, *Available at SSRN 4590271* .

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2022, What do mutual fund investors really care about?, *The Review of Financial Studies* 35, 1723–1774.

Bessembinder, Hendrik, Chester Spatt, and Kumar Venkataraman, 2020, A survey of the microstructure of fixed-income markets, *Journal of Financial and Quantitative Analysis* 55, 1–45.

Bloomberg, 2025, Bloomberg US Agg Total Return Value Unhedged USD.

Bogfan, Michael, and Daniel Schrass, 2025, What US household consider when they select mutual funds, 2024, *ICI Research Perspective* 31.

Chen, Huaizhi, Lauren Cohen, and Umit G Gurun, 2021, Don't take their word for it: the misclassification of bond mutual funds, *The Journal of Finance* 76, 1699–1730.

Chen, Huaizhi, Richard Evans, and Yang Sun, 2025, Self-declared benchmarks and fund manager intent:"cheating" or competing?, *Journal of Financial Economics* 165, 103975.

Chen, Yong, Wayne Ferson, and Helen Peters, 2010, Measuring the timing ability and performance of bond mutual funds, *Journal of Financial Economics* 98, 72–89.

Choi, Jaewon, and Mathias Kronlund, 2018, Reaching for yield in corporate bond mutual funds, *Review of Financial Studies* 31, 1930–1965.

Cici, Gjergji, Scott Gibson, and John J Merrick Jr, 2011, Missing the marks? Dispersion in corporate bond valuations across mutual funds, *Journal of Financial Economics* 101, 206–226.

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

Evans, Richard B, and Yang Sun, 2021, Models or stars: The role of asset pricing models and heuristics in investor risk adjustment, *The Review of Financial Studies* 34, 67–107.

Financial Times, 2026, "When it comes to bond funds, which is better: Passive or active?",

https://www.ft.com/content/b4257989-182a-4696-8db7-191f1f23e80a .

Huang, Jing-Zhi, and Ying Wang, 2014, Timing ability of government bond fund managers: Evidence from portfolio holdings, *Management Science* 60, 2091–2109.

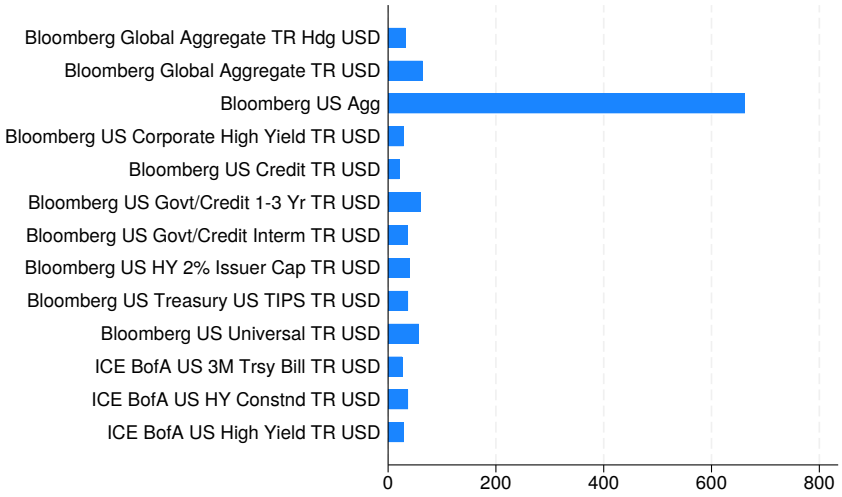
Ivković, Zoran, and Scott Weisbenner, 2009, Individual investor mutual fund flows, *Journal of Financial Economics* 92, 223–237.

Morningstar, 2024, The strange case of the active bond fund, *https://www.morningstar.com/business/insights/blog/funds/active-bond-funds* .

Mullally, Kevin, and Andrea Rossi, 2025, Moving the goalposts? Mutual fund benchmark changes and relative performance manipulation, *Review of Financial Studies* 38, 1067–1119.

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

Panel A. Common Bond Fund Benchmarks



Panel B. Lipper Categories of AGG-benchmarked funds

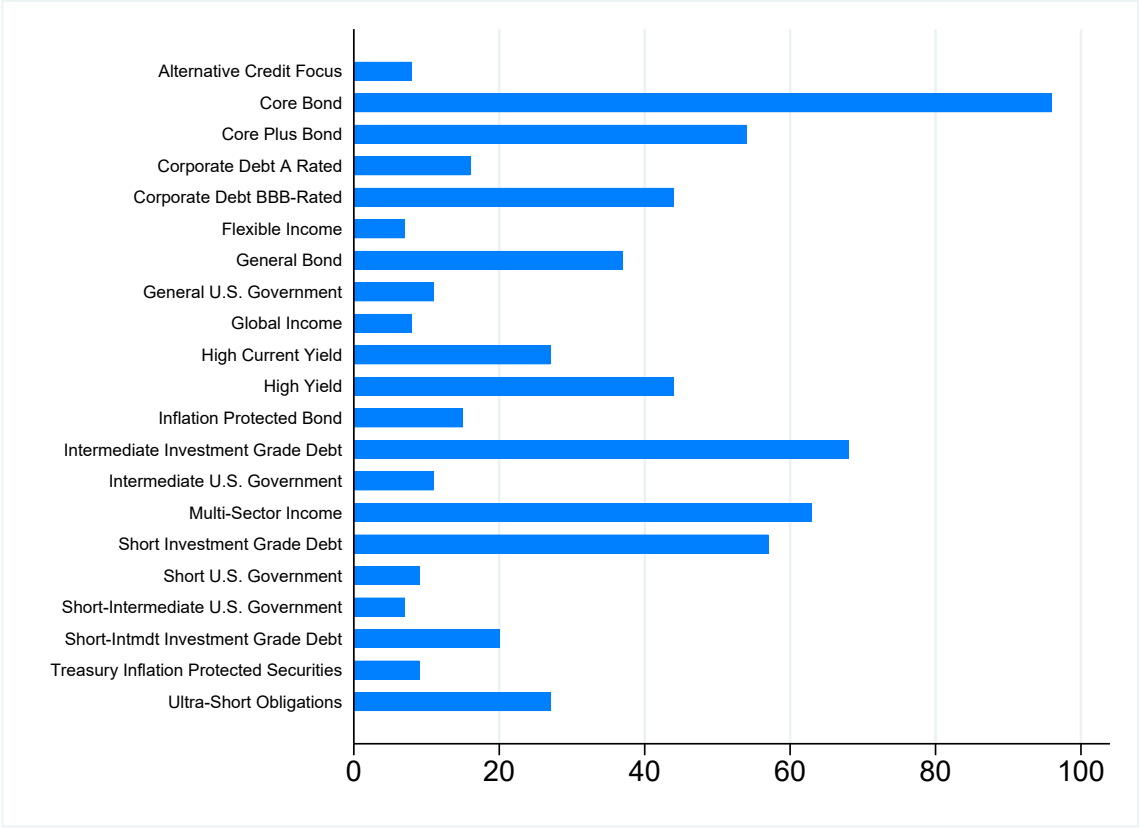


Figure 1. Fund Benchmarks and Lipper Categories
Panel A reports benchmarks used by at least 20 funds in our sample (includes both primary and secondary prospectus benchmarks from Morningstar). Panel B reports the number of funds in each Lipper category that use the Bloomberg U.S. Aggregate Index as a benchmark. A category is only listed if more than 5 funds use the Agg as a benchmark.

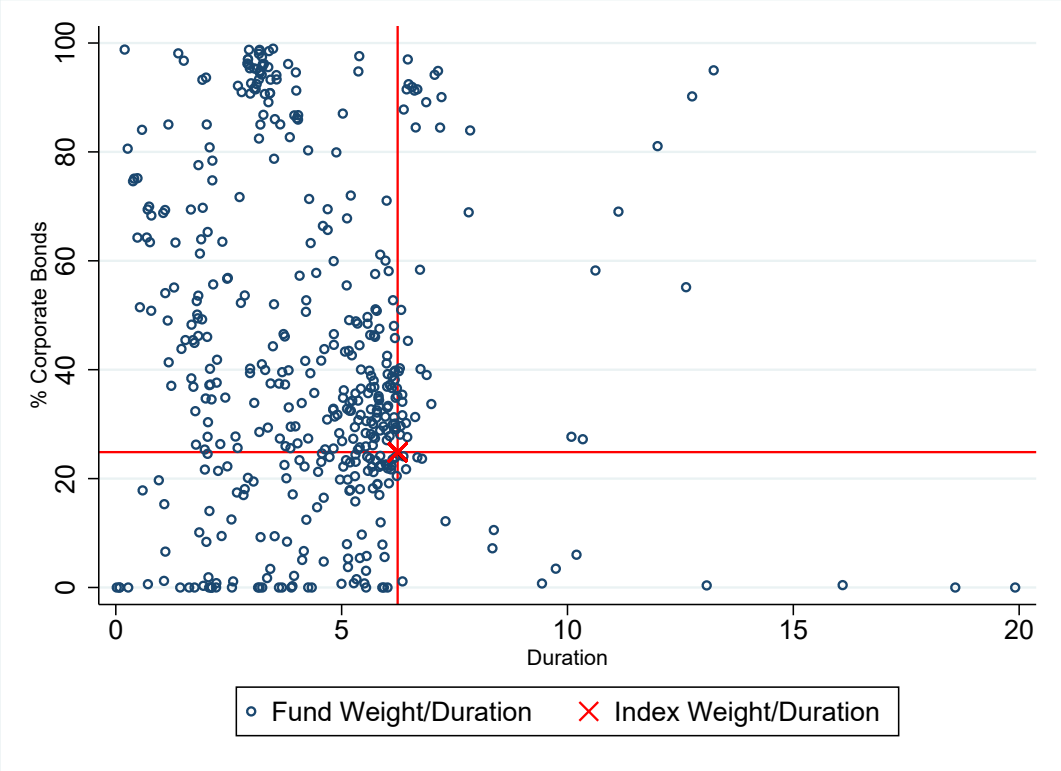


Figure 2. Cross-Sectional Fund Characteristics by % Corporate Bond Holdings and Duration

The sample for this figure is all active bond funds that list the Bloomberg U.S. Aggregate Index as a benchmark in December 2023. Blue circles represent the fund observations, and the red x represents the Agg itself. The y-axis is the % of the porfolio in corporate bonds, and the x-axis is the portfolio duration.

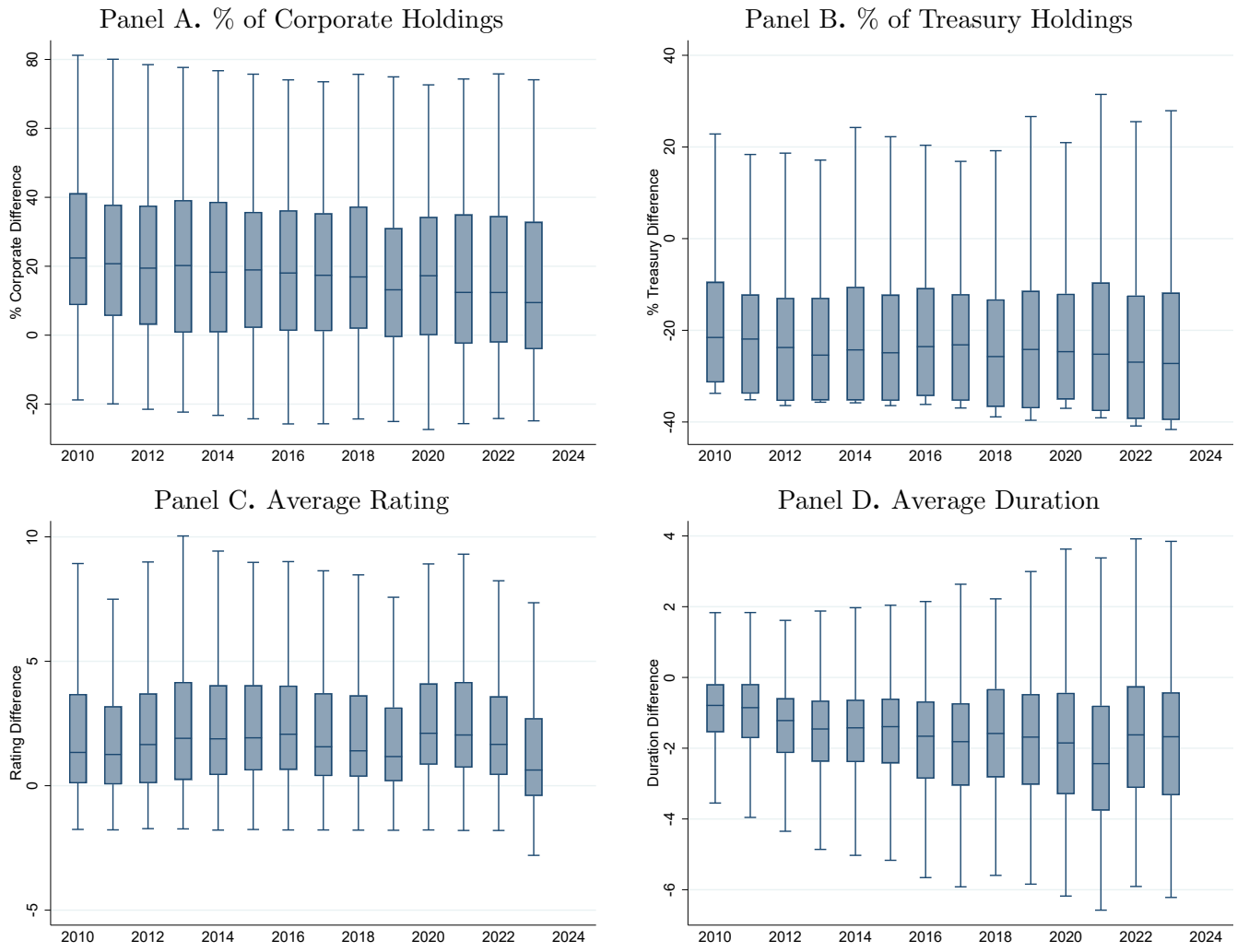


Figure 3. Difference between Fund Characteristics and Agg Characteristics: By Year
This figure reports the inter-quartile range of the deviation of fund characteristics from benchmark characteristics for Agg-benchmarked funds for each year in the sample. Panel A and B measure the deviation in corporate and treasury bond holdings, respectively. Panel C and D measure the deviation in portfolio rating and duration, respectively.

JPMorgan Corporate Bond Fund

Class/Ticker: A/CBRAX; C/CBRCX; I/CBFSX

What is the goal of the Fund?

The Fund seeks to provide total return.

What are the Fund's main investment strategies?

The Fund mainly invests in corporate bonds that are rated investment grade by a nationally recognized statistical rating organization or in securities that are unrated but are deemed by the Fund's adviser, J.P. Morgan Investment Management Inc. (JPMIM or the adviser) to be of comparable quality. Under normal circumstances, the Fund invests at least 80% of its assets in corporate bonds. "Assets" means net assets plus the amount of borrowings for investment purposes. A "corporate bond" is defined as a debt security issued by a corporation or non-governmental entity with a maturity of 90 days or more at the time of its issuance. As part of its principal strategy, the Fund invests in corporate bonds structured as corporate debt securities, debt securities of real estate investment trusts (REITs) and master limited partnerships (MLPs), public or private placements, restricted securities and other unregistered securities.

The Fund is managed relative to the Bloomberg U.S. Corporate Index (the benchmark). Under normal circumstances, the Fund's duration is the duration of the benchmark, plus or minus one year. Duration is a measure of price sensitivity of a debt security or a portfolio of debt securities to relative changes in interest rates. For instance, a duration of "five years" means that a security's or portfolio's price would be expected to decrease by approximately 5% with a 1% increase in interest rates (assuming a parallel shift in yield curve). As of May 31, 2024, the duration of the benchmark was 6.85 years. The Fund

AVERAGE ANNUAL TOTAL RETURNS
(For periods ended December 31, 2023)

	Past 1 Year	Past 5 Years	Past 10 Years
CLASS I SHARES			
Return Before Taxes	8.94%	2.84%	3.18%
Return After Taxes on Distributions	7.03	1.00	1.39
Return After Taxes on Distributions and Sale of Fund Shares	5.23	1.47	1.72
CLASS A SHARES			
Return Before Taxes	4.67	1.81	2.53
CLASS C SHARES			
Return Before Taxes	7.18	2.10	2.52
BLOOMBERG U.S. AGGREGATE INDEX (Reflects No Deduction for Fees, Expenses, or Taxes)	5.53	1.10	1.81
BLOOMBERG U.S. CORPORATE INDEX (Reflects No Deduction for Fees, Expenses, or Taxes)	8.52	2.63	2.95

Figure 4. Panel A: JP Morgan Corporate Bond Fund Prospectus

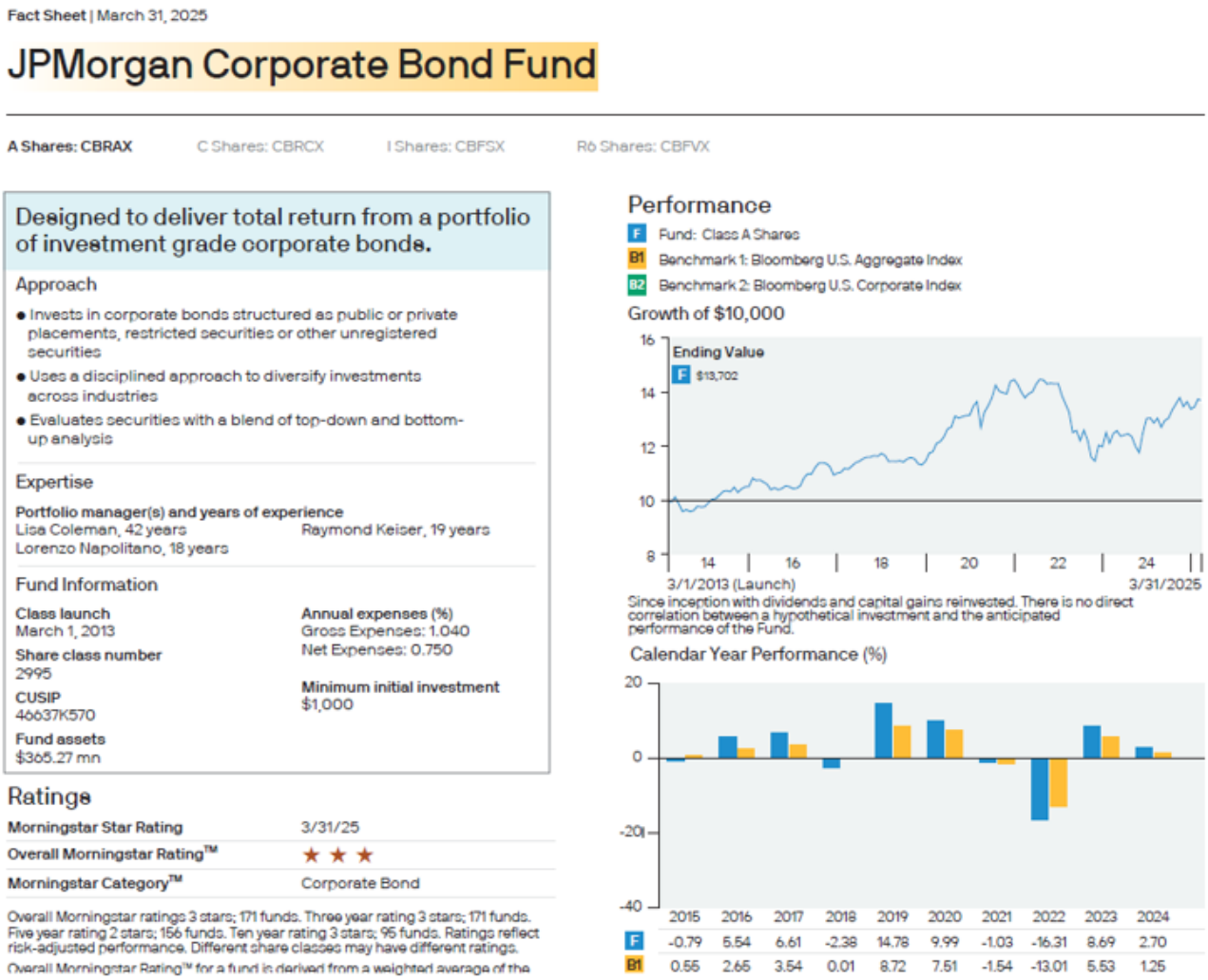


Figure 4. Panel B: JP Morgan Corporate Bond Fund Fact Sheet

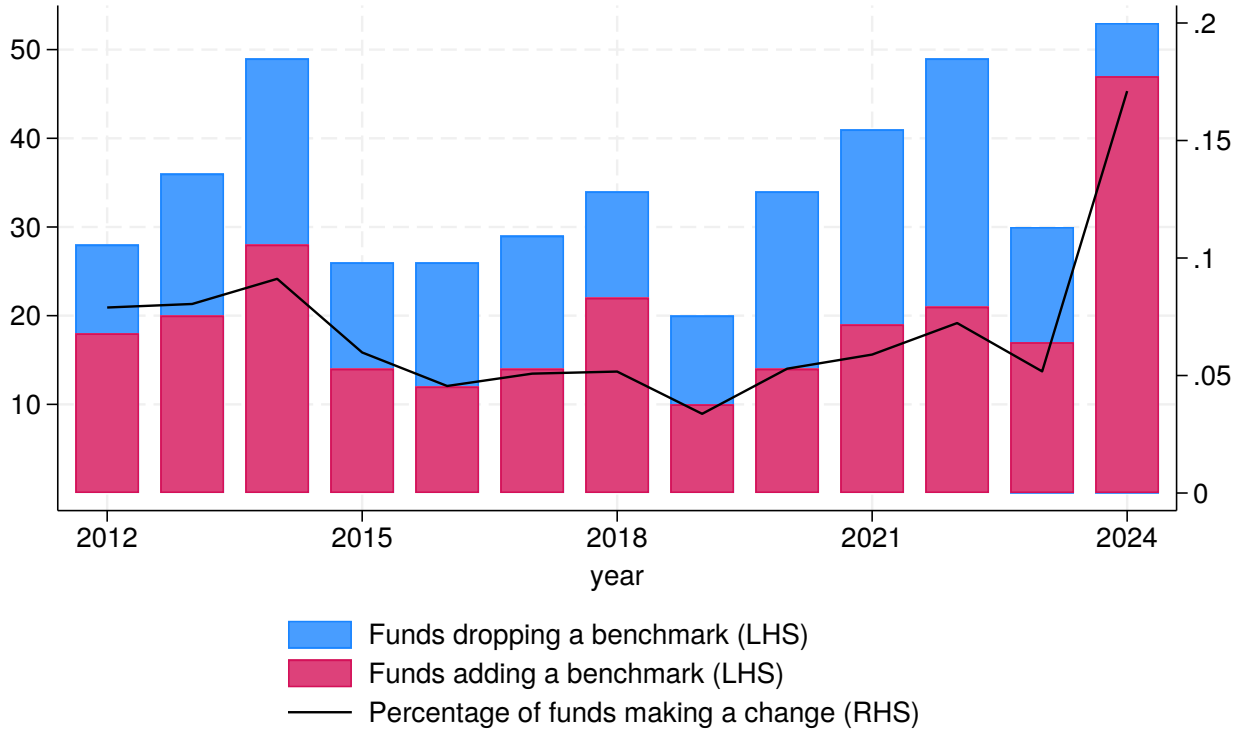


Figure 5. Number of Benchmark Changes by Year
This figure uses a sample of benchmark changes collected directs from funds' summary prospectus. The blue (red) bars represent the number of funds dropping (adding) a benchmark each year. The black line represents the percentage of active bond funds that change their benchmark each year.

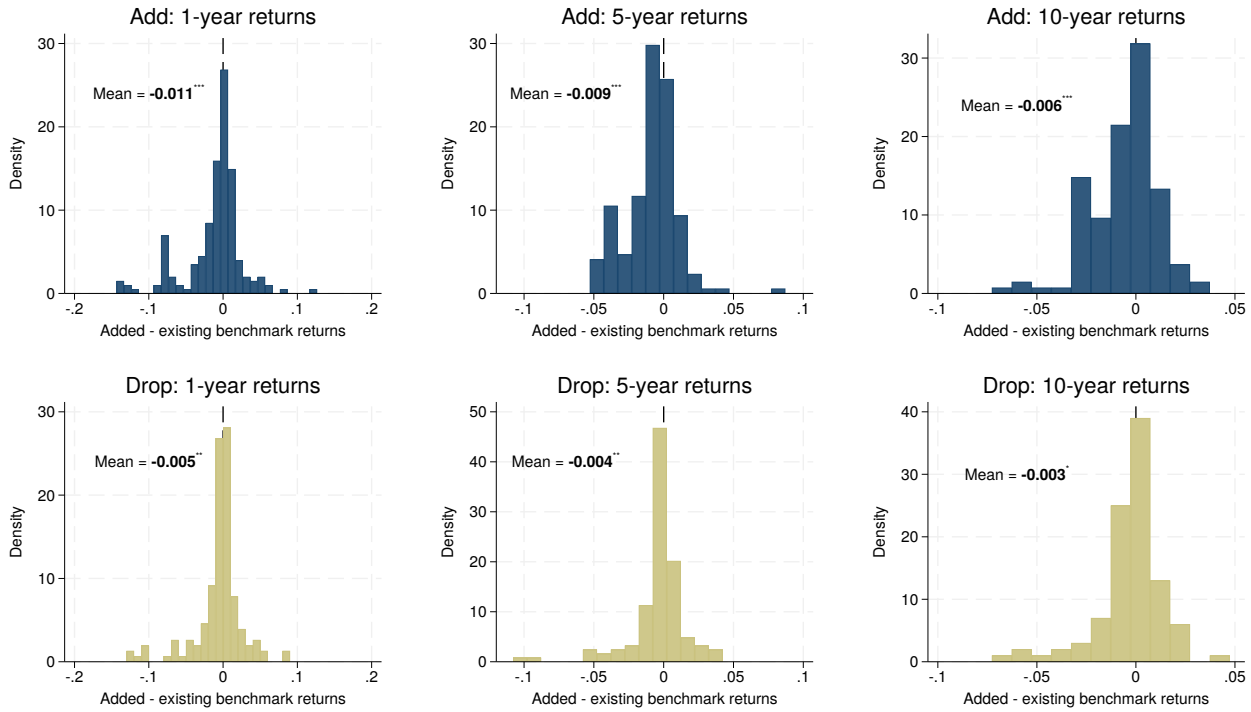


Figure 6. Calculations of $\Delta BenchmarkRet$: All changes
This figure plots the distribution of the change in returns when funds add (top row) or drop (bottom row) a benchmark. For benchmark additions, the return is calculated as the return of the benchmark(s) added minus the return of the existing benchmark(s). For benchmark deletions, the return is calculated as the return of the remaining benchmark(s) minus the return of the dropped benchmark(s).*,**,*** represent significance at the 10%, 5%, and 1% level respectively

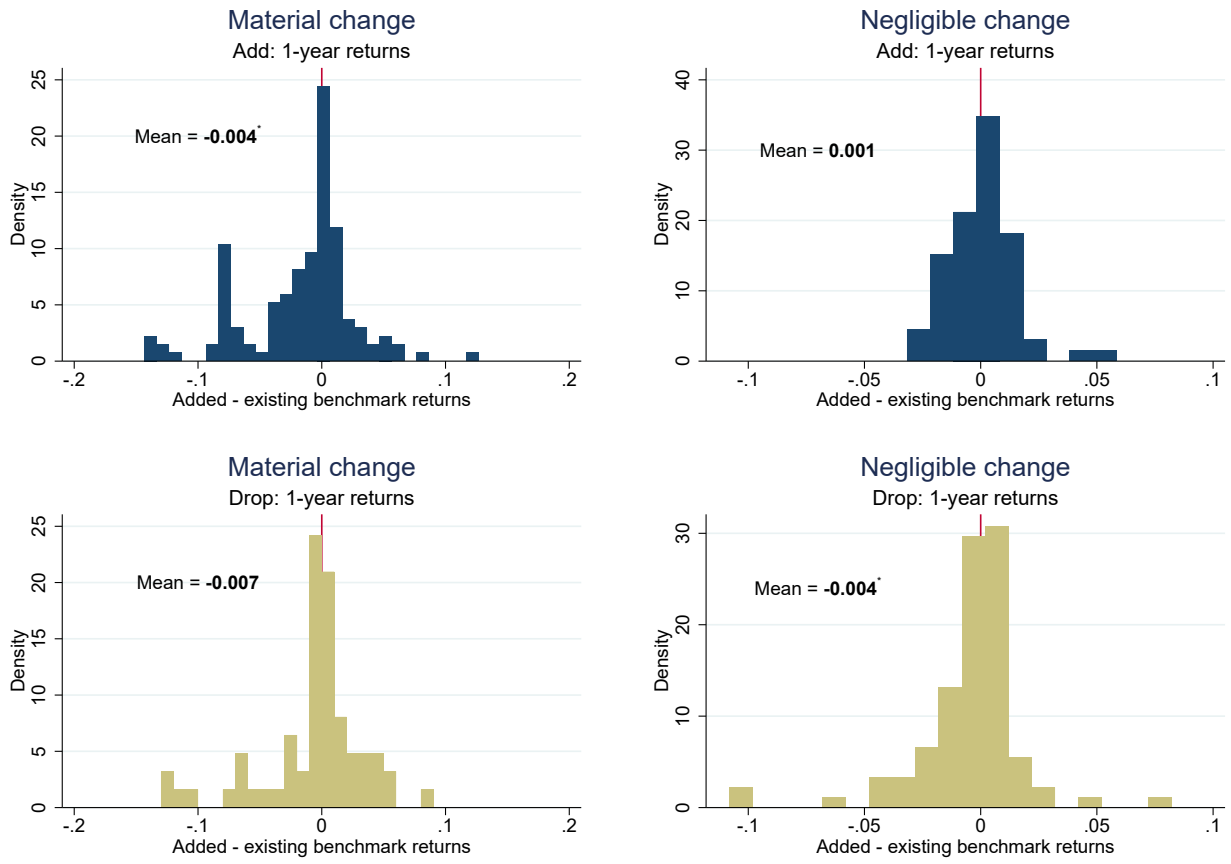


Figure 7. Calculations of $\Delta BenchmarkRet$: Material vs negligible changes
This figure plots the distribution of the change in returns when funds make a material benchmark change (top row) or and negligible benchmark change (bottom row) a benchmark. Material benchmark changes are defined as those benchmark switches where the new benchmark has a different risk categorization. A negligible switch is one where the there is the new benchmark has the same risk categorization.*,**,*** represent significance at the 10%, 5%, and 1% level respectively

Figure 8. Flows Around Benchmark Changes
This figure plots the average 12 month fund flows for funds that make a benchmark change. Flows are adjusted based on the average flows to all funds in the same Lipper category. The solid grey line is the average flows for all funds that switch. The dashed blue line is the average flows for funds that change to a lower risk benchmark. The dashed red line is the average flows for funds that change to a lower return benchmark. The dashed green line is the average flow for those funds that make a meaningful change.

Table I. Summary Statistics

This table presents the summary statistics for the active fixed income funds in our sample. PPB Adj Gross Return (%) is the gross return of the fund, computed as the net return plus one-twelfth of the expense ratio, minus the return of the fund's Primary Prospectus Benchmark (PPB). Category Return is the fund's gross return in excess of the average return of funds in the same Lipper Category and date.

	Mean	Std. Dev.	p10	p25	p50	p75	p90
Net Return 1m (%)	0.26	1.43	-1.28	-0.27	0.23	0.89	1.78
Gross Return 1m (%)	0.31	1.43	-1.22	-0.22	0.28	0.95	1.83
PPB Adj Gross Return 1m (%)	0.09	0.84	-0.56	-0.14	0.07	0.30	0.79
PPB Adj Gross Return 36m (%)	2.97	5.64	-2.08	0.22	2.07	4.97	9.94
Category Return - PPB Return 36m (%)	2.97	4.94	-1.62	0.41	2.44	4.89	8.94
Flow (%)	0.18	5.05	-3.49	-1.40	-0.14	1.31	4.09
Fund Age (Years)	17.40	11.15	4.00	8.04	16.00	24.99	32.48
Fund Assets ($M)	2381.97	8348.94	54.90	148.10	489.90	1637.30	5095.30
Expense (%)	0.65	0.33	0.25	0.46	0.64	0.85	1.05
Turnover (%)	0.37	0.87	0.00	0.00	0.00	0.36	1.14
Rating	5.47	3.76	1.02	2.81	4.43	7.42	12.00
Yield to Maturity	3.86	2.71	1.26	2.15	3.29	5.21	6.96
Duration	4.04	2.34	1.57	2.44	3.88	5.11	6.25
Coupon (%)	3.76	1.68	1.75	2.63	3.49	4.86	6.14
Time to Maturity	9.43	4.78	4.03	5.93	8.80	12.39	15.37
Option Adjusted Spread	145.27	146.25	17.93	46.44	92.73	185.35	372.20
% Corporate	44.61	30.78	0.84	20.40	41.60	68.80	91.83
% Treasury	17.42	23.11	0.00	0.24	8.94	24.28	46.44
% Securitized	17.70	20.42	0.00	0.00	11.31	29.47	44.32
% Equity	1.15	6.84	0.00	0.00	0.00	0.00	1.11
% Muni	1.04	6.58	0.00	0.00	0.00	0.27	1.23

Table II. The Importance of Benchmark Adjusted Returns on Flows

This table examines the relationship between benchmark adjusted returns and fund flows. The dependent variable in all specifications is the fund's monthly flow, as defined in equation I. *Benchmark-adjusted return* is the net return of the fund minus the prospectus benchmark over the past 36 months. *Fund unadjusted return* is the net return of the fund over the previous 36 months. Tracking error is the standard deviation of the residual from a regression of fund returns on benchmark returns over the previous 36 months. We also control for the total net assets of the fund, the age of the fund, expense ratio, turnover. Columns (1)-(5) are OLS regressions, and Column (6) is a Fama-MacBeth regression with Newey-West standard errors lagged by 6 months. For all OLS models, standard errors are clustered by fund and date; *,**,*** represent significance at the 10%, 5%, and 1% level respectively.

Regression model	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	FMB (6)
Benchmark-adjusted return		0.056***	0.068***	0.042***	0.081***	0.071***
		(0.009)	(0.009)	(0.010)	(0.012)	(0.010)
Fund unadjusted return	0.059***		0.008	0.029**	-0.004	0.021**
	(0.011)		(0.007)	(0.012)	(0.010)	(0.008)
Tracking error	-0.023	-0.064	-0.309***	-0.053	-0.646***	-0.195*
	(0.090)	(0.092)	(0.095)	(0.092)	(0.211)	(0.111)
Observations	108,852	108,852	116,615	108,852	116,615	116,615
Adjusted R-squared	0.088	0.088	0.043	0.089	0.097	0.033
Date FE	NO	NO	YES	NO	YES	NO
Category-by-Date FE	YES	YES	NO	YES	NO	NO
Rating-by-Date FE	YES	YES	NO	YES	NO	NO
Fund FE	NO	NO	NO	NO	YES	NO
Controls	YES	YES	YES	YES	YES	YES

Table III. Flow-performance Relationship: Decomposing Returns

This table presents the results of monthly regressions of monthly flow on the lagged components of benchmark-adjusted returns. Benchmark-adjusted returns are the 36-month sum of difference between net monthly fund return, and of the monthly return on the primary prospectus benchmark (PPB). We then decompose the excess return into two components. First, *Fund return - Category return* measures the fund's net returns in excess of the Lipper category average. Second, *Category return - PPB return* is the average net return of all funds in the same Lipper category and date in excess of the fund's PPB return. We control for the total net assets of the fund, the age of the fund, expense ratio, turnover. Columns (1)-(3) are OLS regressions, and Column (4) is a Fama-MacBeth regression with Newey-West standard errors lagged by 6 months. For all OLS models, standard errors are clustered by fund and date; *,**,*** represent significance at the 10%, 5%, and 1% level respectively.

Regression model	OLS (1)	OLS (2)	OLS (3)	FMB (4)
(1) Fund return - category return	0.116***	0.071***	0.115***	0.131***
	(0.012)	(0.012)	(0.015)	(0.022)
(2) Category return - PPB return	0.042***	0.042***	0.047***	0.048***
	(0.009)	(0.010)	(0.013)	(0.008)
Observations	116,615	108,852	116,615	116,615
Adjusted R-squared	0.045	0.089	0.098	0.037
Date FE	YES	NO	YES	NO
Category-by-Date FE	NO	YES	NO	NO
Rating-by-Date FE	NO	YES	NO	NO
Fund FE	NO	NO	YES	NO
Controls	YES	YES	YES	YES

Table IV. Evidence of Misbenchmarking Using the Universe Benchmarks

In this table we present the performance of funds compared to their primary prospectus benchmarks (PPB) and the universe of all benchmarks used within our sample. Panel A presents statistics for funds whose correlation is highest with their PPB. Panel B presents the subsample of funds whose correlation is highest with a benchmark that is not their PPB. Panel C includes the funds in Panel B whose matched benchmark outperforms their PPB and that have increasing levels of correlation improvement (at least 0.05, 0.10, 0.25, and 0.50).

Panel A: Properly benchmarked funds	N	Mean	SD	p10	p25	p50	p75	p90
Fund correlation with PPB	325	0.954	0.052	0.898	0.948	0.971	0.983	0.991
Fund PPB adjusted gross return 1m (bps)	325	4.805	8.163	-3.086	0.879	4.169	8.156	13.410
Fund PPB adjusted net return 1m (bps)	325	-0.480	8.437	-8.637	-4.268	-0.966	2.637	7.807

Panel B: Misbenchmarked funds	N	Mean	SD	p10	p25	p50	p75	p90
Fund correlation with PPB (1)	859	0.727	0.265	0.322	0.584	0.830	0.936	0.968
Fund correlation with best matched benchmark (2)	859	0.878	0.126	0.707	0.829	0.929	0.966	0.979
Best matched benchmark correlation improvement (2) - (1)	859	0.151	0.185	0.002	0.014	0.074	0.224	0.418
Best matched benchmark return minus PPB return 1m (bps)	859	9.622	13.091	-1.255	1.424	5.386	9.728	32.932
Fund PPB adjusted gross return 1m (bps)	859	9.794	13.053	-1.352	3.400	7.948	14.376	27.118
Fund best matched benchmark adjusted gross return 1m (bps)	859	-0.160	12.819	-14.922	-3.994	0.953	5.335	10.912
Fund PPB adjusted net return 1m (bps)	859	4.248	12.777	-7.597	-1.479	3.048	9.170	19.153
Fund best matched benchmark adjusted net return 1m (bps)	859	-5.710	13.095	-21.474	-9.819	-3.948	0.327	5.356

	≥ 0.05		≥ 0.10		≥ 0.25		≥ 0.50	
Panel C: Misbenchmarked funds by correlation improvement	N	Mean	N	Mean	N	Mean	N	Mean
Fund correlation with PPB (1)	500	0.587	392	0.519	196	0.356	55	0.147
Fund correlation with best matched benchmark (2)	500	0.836	392	0.817	196	0.789	55	0.798
Best matched benchmark correlation improvement (2) - (1)	500	0.249	392	0.299	196	0.433	55	0.650
Best matched benchmark return minus PPB return 1m (bps)	500	14.492	392	16.684	196	22.892	55	29.924
Fund PPB adjusted gross return 1m (bps)	500	12.586	392	13.789	196	16.261	55	20.730
Fund best matched benchmark adjusted gross return 1m (bps)	500	-2.442	392	-3.393	196	-6.755	55	-6.970
Fund PPB adjusted net return 1m (bps)	500	6.955	392	8.020	196	10.567	55	15.009
Fund best matched benchmark adjusted net return 1m (bps)	500	-8.077	392	-9.167	196	-12.456	55	-12.692

Table V. Evidence of Misbenchmarking Using the Pseudo Benchmarks

In this table we present the performance of funds compared to a set of pseudo benchmarks, created using 67 weighted combinations of the Bloomberg U.S. Aggregate Index (AGG), the Bloomberg U.S. Corporate Index (IG), and the ICE BofA U.S. High Yield Index (HY). Panel A includes funds whose primary prospectus benchmark (PPB) and fund return both correlate to the same pseudo benchmark. Panel B includes funds whose PPB correlate to a different pseudo benchmark than the fund itself. ω represents the weight in each component for the fund's pseudo benchmark and the PPB's pseudo benchmark.

Panel A: Properly benchmarked funds	N	Mean	SD	p10	p25	p50	p75	p90
Fund correlation with PPB	292	0.929	0.108	0.842	0.927	0.962	0.979	0.989
Fund PPB adjusted gross return 1m (bps)	292	2.460	11.111	-4.900	-0.574	2.796	6.667	11.737
Fund PPB adjusted net return 1m (bps)	292	-3.307	11.190	-11.961	-6.453	-2.191	0.966	5.020

Panel B: Misbenchmarked funds	N	Mean	SD	p10	p25	p50	p75	p90
Fund correlation with PPB (1)	892	0.744	0.264	0.338	0.599	0.850	0.946	0.976
Fund correlation with best matched benchmark (2)	892	0.866	0.138	0.684	0.821	0.908	0.966	0.983
Best matched benchmark correlation improvement (2) - (1)	892	0.123	0.196	-0.032	0.004	0.050	0.195	0.414
Best matched benchmark return minus PPB return 1m (bps)	892	9.173	11.611	-4.007	0.866	6.990	17.867	25.600
Fund PPB adjusted gross return 1m (bps)	892	10.331	11.785	-0.408	3.869	8.051	14.356	26.050
Fund best matched benchmark adjusted gross return 1m (bps)	892	-2.308	13.272	-19.462	-8.690	0.274	5.022	10.050
Fund PPB adjusted net return 1m (bps)	892	4.949	11.427	-6.223	-0.905	3.261	9.129	19.153
Fund best matched benchmark adjusted net return 1m (bps)	892	-7.690	13.437	-24.430	-14.512	-5.086	0.238	4.972
ωAGG of fund's best matched pseudo benchmark	892	0.412	0.325	0.000	0.100	0.400	0.700	0.800
ωIG of fund's best matched pseudo benchmark	892	0.239	0.228	0.000	0.000	0.200	0.400	0.600
ωHY of fund's best matched pseudo benchmark	892	0.349	0.303	0.000	0.100	0.250	0.500	0.900
ωAGG of PPB's best matched pseudo benchmark	882	0.821	0.332	0.100	0.900	1.000	1.000	1.000
ωIG of PPB's best matched pseudo benchmark	882	0.078	0.225	0.000	0.000	0.000	0.000	0.200
ωHY of PPB's best matched pseudo benchmark	882	0.102	0.253	0.000	0.000	0.000	0.100	0.200

Table VI. Tests of Differences in Control and Misbenchmarked Funds

This table presents the difference in portfolio characteristics between funds with mismatched benchmarks and a control sample. Funds are designated as misbenchmarked if an alternative benchmark improves gross return correlation by at least 0.10, and zero otherwise. Mismatch is identified using the entire benchmark universe. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Control	Meaningfully Misbenchmarked	Difference
Flow (%)	0.073	0.398	-0.325***
Fund age (years)	18.507	15.156	3.351***
Fund assets ($m)	2560.149	2022.869	537.281***
Expense (%)	0.642	0.678	-0.036***
Turnover (%)	0.391	0.336	0.056***
Net return 36m (%)	8.735	7.386	1.349***
Gross return 36m (%)	10.637	9.392	1.245***
PPB adjusted gross return 36m (%)	2.15	4.645	-2.495***
Match minus PPB return 36m (%)	0.855	5.557	-4.702***
Match adjusted gross return 36m (%)	1.294	-0.912	2.206***
Rating	5.634	5.102	0.532***
Yield to maturity	3.822	3.96	-0.138***
Duration	4.583	2.75	1.833***
Coupon	3.854	3.54	0.314***
Time to maturity	9.846	8.56	1.286***
Option adjused spread	153.943	112.945	40.998***
% corporate	46.153	41.449	4.703***
% Treasury	20.183	11.767	8.417***
% securitized	15.355	22.513	-7.158***
% equities	0.979	1.507	-0.528***
% municipal	0.983	1.161	-0.178***

Table VII. Misbenchmarked Funds and Gross Adjusted Performance

This table reports the results of monthly regression of gross primary prospectus benchmark adjusted returns on mismatched benchmark dummy and proxies for reach for yield. The dependent variable is the difference between gross monthly fund return, computed as net returns plus one-twelfth of the expense ratio, net of the monthly return on the primary prospectus benchmark. *Misbenchmarked* represents meaningfully misbenchmarked funds, and is an indicator variable equal to one if an alternative benchmark improves gross return correlation by at least 0.10 and the monthly difference between the matched and prospectus benchmark (PPB) returns is positive, and zero otherwise. YTM and TTM are value-weighted measures of the yield to maturity and time to maturity respectively, of the fund holdings. All regressions include lagged monthly controls for expense ratio, fund turnover, log of fund assets, log of fund age, and flow, but the coefficients are not reported. Standard errors clustered at the fund and date level are presented in parenthesis below the coefficient.

Panel A: YTM and returns

	PPB adjusted return				Gross fund return	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Misbenchmarked	6.917*	7.656*	5.922*	7.271*	-0.015	-0.013
	(4.139)	(4.557)	(3.483)	(3.898)	(0.012)	(0.012)
Misbenchmarked * YTM(m-1)			0.317	0.200		
			(0.593)	(0.620)		
YTM(m-1)			0.011	0.014		
			(0.330)	(0.379)		
Observations	146,270	135,615	105,305	96,280	146,270	135,615
Adjusted R-squared	0.402	0.457	0.363	0.432	0.874	0.894
Category-by-Date FE	YES	NO	YES	NO	YES	NO
Match bmrk-by-Date FE	YES	NO	YES	NO	YES	NO
Category-by-Match bmrk-by-Date FE	NO	YES	NO	YES	NO	YES
Controls	YES	YES	YES	YES	YES	YES

Panel B: YTM by benchmark category

	Yield to maturity	
VARIABLES	(1)	(2)
Misbenchmarked	0.611***	-0.079
	(0.198)	(0.204)
Observations	97,261	105,638
Adjusted R-squared	0.439	0.458
PPB-by-Date FE	YES	NO
Match bmrk-by-Date FE	NO	YES
Controls	YES	YES

Table VIII. Misbenchmarking and Expense Ratios

This table reports the results of annual regression of expense ratios on mismatched benchmark dummy. The dependent variable is the expense ratio. In Columns (1) to (3), *Misbenchmarked* is an indicator variable equal to one if an alternative benchmark improves gross return correlation by at least 0.10, and zero otherwise. In Columns (4) to (6), *Misbenchmarked* is an indicator variable equal to one if a Pseudo benchmark, composed of the Bloomberg U.S. Aggregate Index, the Bloomberg U.S. Corporate Index, and the ICE BofA U.S. High Yield indices, improves gross return correlation by at least 0.10, and zero otherwise. We control for the total net assets of the fund, the age of the fund, expense ratio, turnover. standard errors are clustered by fund and date; *,**,*** represent significance at the 10%, 5%, and 1% level respectively.

	Universe of benchmarks			Pseudo benchmarks		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Misbenchmarked	0.075***	0.087***	0.097***	0.091***	0.102***	0.114***
	(0.022)	(0.021)	(0.021)	(0.021)	(0.020)	(0.020)
Observations	10,284	10,286	10,286	10,284	10,286	10,286
Adjusted R-squared	0.350	0.356	0.336	0.353	0.361	0.341
Category-by-Date FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES

Table IX. Benchmark Changes: Risk Category Transition Matrix

This transition matrix documents the old and new risk categories for funds making a benchmark change. The columns are the risk categories of a fund's benchmark in the year before the benchmark change, while the rows are the risk categories of a fund's new benchmark following a change.

	Post Switch Risk Category						
Pre Switch Risk Category	Risk-free	GovCredit	Agg	Credit	Universal	IGCorp	HY/Equity
Risk-free	40	0	16	6	2	3	1
GovCredit	2	23	29	1	0	2	2
Agg	4	8	56	4	6	4	4
Credit	0	1	12	12	3	4	0
Universal	1	0	10	2	1	0	0
IGCorp	1	0	7	0	1	6	1
HY/Equity	0	1	29	0	7	1	82

Table X. Why Do Funds Change Benchmarks?

The table examines the likelihood of benchmark switches. *Switch* takes the value of 1 for the first switch of a fund. *Material* takes the value of 1 for the first switch of a fund, if the switch involved a change in the risk category of the benchmark, and zero for control funds and switches that did not involve a change in benchmark risk categorization. *Negligible* takes the value of 1 for the first switch of a fund, if the switch did not result in a change in the risk category of the benchmark, and zero for control funds and switches where benchmark risk categorization changed. Control funds are those funds that never change their benchmark that are in the same Lipper Category as the switching fund. All regressions include lagged monthly controls for expense ratio, fund turnover, log of fund assets, log of fund age, and flow, but the coefficients are not reported. Lipper class-by-date fixed effects are included in each column. Standard errors clustered at the year level are presented in parenthesis below the coefficient.

VARIABLES	All switches (1)	Material switches (2)	Negligible switches (3)	All switches (4)	Material switches (5)	Negligible switches (6)
Gross return 36m (%)	-0.0318	-0.0383*	0.00656			
	(0.0198)	(0.0205)	(0.0166)			
Pseudo bench. adj ret 36m (%)				-0.0723**	-0.0624**	-0.00988
				(0.0301)	(0.0260)	(0.0135)
Observations	11,001	11,001	11,001	11,001	11,001	11,001
Adjusted R-squared	0.021	0.030	0.027	0.021	0.030	0.027
Category-by-Date FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES

Table XI. Benchmark Changes and Subsequent Fund Flows and Performance

The table examines the flow to funds that switch their benchmarks and uses a stacked differences-in-differences design. Control funds are all funds that never change their benchmark. The pre-period for treated and control funds is t-36 to t-12 from the switch date. The post period is all observations 24 months after the switch. *Switch* takes the value of 1 for the 24 month period after a change in benchmarks. *Lower risk bmrk* is an indicator variable that takes the value of one for a switch that lowers the risk classification of the funds benchmark. *Lower return bmrk* is an indicator variable that takes the value of one for a switch that results in a negative $\Delta BenchmarkRet$. All regressions include lagged monthly controls for expense ratio, fund turnover, log of fund assets, log of fund age, and flow, but the coefficients are not reported. Fund, Morningstar Rating, and Event-by-Date fixed effects are included in each column. Standard errors clustered by fund and date level are presented in parenthesis below the coefficient.

VARIABLES	Flow (1)	Flow (2)	Flow (3)	Flow (4)
Switch	-0.114	-0.547*	-0.854**	-0.967**
	(0.237)	(0.279)	(0.354)	(0.377)
Switch * Lower risk bmrk		1.382***		1.448***
		(0.478)		(0.507)
Switch * Lower return bmrk			1.140**	0.755
			(0.457)	(0.531)
Switch * Lower risk bmrk * Lower return bmrk				-0.392
				(0.772)
Observations	345,660	345,660	344,425	344,425
Adjusted R-squared	0.151	0.151	0.151	0.151
Fund FE	YES	YES	YES	YES
Rating FE	YES	YES	YES	YES
Event-by-Date FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES

Table XII. Effect of Institutional Investors and Passive Share on Flow Sensitivity to Benchmark-Adjusted Returns

This table examines the effect of institutional investor presence and passive competition on the relationship between flows and returns. The dependent variable in all specifications is the fund's monthly flow, as defined in equation I. *Benchmark-adjusted return* is the net return of the fund minus the prospectus benchmark over the past 36 months. *Fund unadjusted return* is the net return of the fund over the previous 36 months. *Institutional* is the percentage of a fund's assets in institutional share classes, while *Passive %* is the previous month's share of assets under management in a fund's Lipper category that are in passive funds. In all specifications, we control for tracking error, the total net assets of the fund, the age of the fund, expense ratio, and turnover. Standard errors are clustered by fund and date; *,**,*** represent significance at the 10%, 5%, and 1% level respectively.

Panel A: Effect of Institutional Investors on Flow Determinants

	(1)	(2)	(3)	(4)	(5)
Benchmark-adj. ret.		0.045***	0.058***	0.024*	0.078***
		(0.011)	(0.013)	(0.013)	(0.017)
Benchmark-adj. ret.*Institutional		0.022	0.021	0.031**	0.006
		(0.015)	(0.017)	(0.016)	(0.022)
Fund unadj. return	0.069***		0.018*	0.042***	-0.002
	(0.011)		(0.009)	(0.013)	(0.012)
Fund unadj. return*Institutional	-0.016*		-0.017*	-0.021**	-0.003
	(0.009)		(0.009)	(0.009)	(0.010)
Institutional	-0.001	-0.003***	-0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Adjusted R-squared	0.088	0.089	0.043	0.089	0.097

Panel B: Effect of Passive Share on Flow Determinants

	(1)	(2)	(3)	(4)	(5)
Benchmark-adj. ret.		0.053***	0.059***	0.030*	0.072***
		(0.010)	(0.010)	(0.010)	(0.013)
Benchmark-adj. ret.*Passive %		0.032	0.118***	0.132**	0.089**
		(0.038)	(0.043)	(0.052)	(0.039)
Fund unadj. return	0.066***		0.014*	0.045***	0.004
	(0.012)		(0.008)	(0.013)	(0.011)
Fund unadj. return*Passive %	-0.058		-0.044**	-0.147***	-0.035*
	(0.039)		(0.018)	(0.048)	(0.018)
Passive %	-0.006	-0.008	0.007***	-0.003	0.016***
	(0.020)	(0.020)	(0.002)	(0.021)	(0.004)
Adjusted R-squared	0.088	0.088	0.044	0.089	0.098
Observations (both panels)	108,852	108,852	116,615	108,852	116,615
Date FE (both panels)	NO	NO	YES	NO	YES
Category-by-Date FE (both panels)	YES	YES	NO	YES	NO
Rating-by-Date FE (both panels)	YES	YES	NO	YES	NO
Fund FE (both panels)	NO	NO	NO	NO	YES
Controls (both panels)	YES	YES	YES	YES	YES

VI. Appendix

Table A1. Benchmark indexes included in alternative universe

The following 30 funds are used as a primary prospectus benchmark by at least five funds in our sample. We include only primary prospectus benchmarks, but in the cases where the fund has a secondary prospectus benchmark with a higher correlation to the fund's returns than the primary, we replace the primary prospectus benchmark with the secondary prospectus benchmark. These indexes are possible benchmarks in our first universe of alternatives.

Index Name	**Number of Funds**
Bloomberg US Agg Bond TR USD	396
Bloomberg US Govt/Credit 1-3 Yr TR USD	51
Bloomberg Global Aggregate TR USD	48
Bloomberg US HY 2% Issuer Cap TR USD	37
Bloomberg US Treasury US TIPS TR USD	36
Bloomberg US Universal TR USD	33
Bloomberg US Govt/Credit Interm TR USD	27
Bloomberg US Credit TR USD	20
Bloomberg US Corp Bond TR USD	16
ICE BofA US 3M Trsy Bill TR USD	12
Bloomberg US Government TR USD	11
Bloomberg U.S. Government/Credit TR	11
Bloomberg US Govt 1-3 Yr TR USD	11
Bloomberg US Govt Interm TR USD	9
Bloomberg US Govt/Credit 1-5 Yr TR USD	8
Bloomberg US Trsy Infl Note 1-10Y TR USD	8
Bloomberg Government 1-5 Yr TR USD	7
Bloomberg Municipal TR USD	7
Bloomberg US Aggregate 1-3 Yr TR USD	7
ICE BofA 1-5Y US CorpTR USD	6
Bloomberg US Agg Interm TR USD	6
Bloomberg US Interm Credit TR USD	6
Bloomberg U.S. Treasury TIPS 1-5Y TR USD	6
Bloomberg Credit 1-3 Yr TR USD	5
Bloomberg Credit 1-5 Yr TR USD	5
Bloomberg Global Credit TR Hdg USD	5
ICE BofA 1-3Y BB US Cash Pay HY TR USD	5
Bloomberg Multiverse TR USD	5
Bloomberg US Govt/Credit Long TR USD	5
Bloomberg US MBS TR USD	5

Table A2. Evidence of misbenchmarking using the pseudo benchmarks

In this table we present the performance of the funds that have greater correlation with a pseudo benchmark other than the Primary Prospectus Benchmark (PPB). Pseudo benchmarks are composed using 67 potential weighted combinations of the Bloomberg US Aggregate Index (AGG), the Bloomberg US Corporate Index (IG), and the ICE BOFA US High Yield Index (HY). ωAGG (IG, HY) is the weight of the AGG (IG, HY) components of the Pseudo Index with the highest correlation to the fund returns, or for the Pseudo Index with the highest correlation to the fund's primary prospectus benchmark (PPB).

	≥ 0.05		≥ 0.10		≥ 0.25		≥ 0.50	
Misbenchmarked funds by correlation improvement	N	Mean	N	Mean	N	Mean	N	Mean
Fund correlation with PPB (1)	467	0.587	347	0.504	187	0.359	54	0.170
Fund correlation with best matched benchmark (2)	467	0.834	347	0.814	187	0.797	54	0.812
Best matched benchmark correlation improvement (2) - (1)	467	0.247	347	0.309	187	0.438	54	0.642
Best matched benchmark return minus PPB return 1m (bps)	467	15.700	347	18.319	187	22.583	54	27.859
Fund PPB adjusted gross return 1m (bps)	467	14.080	347	15.603	187	17.673	54	22.272
Fund best matched benchmark adjusted gross return 1m (bps)	467	-5.430	347	-6.733	187	-8.857	54	-7.633
Fund PPB adjusted net return 1m (bps)	467	8.336	347	9.590	187	11.507	54	16.215
Fund best matched benchmark adjusted net return 1m (bps)	467	-11.174	347	-12.746	187	-15.022	54	-13.690
ωAGG of fund's best matched pseudo benchmark	467	0.294	347	0.217	187	0.111	54	0.046
ωIG of fund's best matched pseudo benchmark	467	0.264	347	0.260	187	0.220	54	0.135
ωHY of fund's best matched pseudo benchmark	467	0.442	347	0.523	187	0.668	54	0.819
ωAGG of PPB's best matched pseudo benchmark	467	0.943	347	0.944	187	0.952	54	0.939
ωIG of PPB's best matched pseudo benchmark	467	0.029	347	0.025	187	0.025	54	0.035
ωHY of PPB's best matched pseudo benchmark	467	0.028	347	0.031	187	0.022	54	0.026

Table A3. Tests of differences in control and misbenchmarked funds: pseudo benchmarks

This table presents the difference in portfolio characteristics between misbenchmarked funds and a control sample. The table uses pseudo-indices composed of the Bloomberg U.S. Aggregate, Bloomberg U.S. Corporate Bond, and ICE BofA U.S. High Yield indices. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Pseudo Benchmarks

	Control	Mismatched	Difference
Flow (%)	-0.197	0.424	-0.621***
Fund age (years)	19.323	14.98	4.342***
Fund assets ($m)	1830.596	2136.802	-306.206***
Expense (%)	0.689	0.706	-0.017***
Turnover (%)	31.532	35.485	-3.952***
Net return 36m (%)	10.021	7.87	2.151***
Gross return 36m (%)	12.053	9.96	2.093***
PPB adjusted gross return 36m (%)	1.434	5.172	-3.738***
Match minus PPB return 36m (%)	0.06	6.16	-6.100***
Match adjusted gross return 36m (%)	1.372	-0.987	2.359***
Rating	6.424	5.459	0.966***
Yield to maturity	4.001	4.253	-0.252***
Duration	4.087	2.957	1.130***
Coupon	4.103	3.74	0.364***
Time to maturity	8.465	8.754	-0.289***
Option adjused spread	171.855	128.933	42.923***

Table A4. Misbenchmarked funds and gross adjusted performance: pseudo benchmarks

This table reports the results of monthly regression of gross primary prospectus benchmark adjusted returns on mismatched benchmark and proxies for reach for yield. The dependent variable is the difference between gross monthly fund return, computed as net returns plus one-twelfth of the expense ratio, net of the monthly return on the primary prospectus benchmark. *Misbenchmarked* is an indicator variable equal to one if a Pseudo benchmark, composed of the Bloomberg U.S. Aggregate, the Bloomberg U.S. Corporate, and the ICE BofA U.S. High Yield indices, improves gross return correlation by at least 0.10, and zero otherwise. YTM is the value-weighted yield to maturity of the fund holdings. All regressions include lagged monthly controls for expense ratio, fund turnover, log of fund assets, log of fund age, and flow, but the coefficients are not reported. Standard errors clustered at the fund and date level are presented in parenthesis below the coefficient.

Panel A: YTM and returns

	PPB adjusted return				Gross fund return	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Misbenchmarked	7.764*	8.091*	6.507*	6.386*	-0.010	-0.008
	(4.049)	(4.322)	(3.335)	(3.440)	(0.010)	(0.010)
Misbenchmarked * YTM(m-1)			0.272	0.323		
			(0.609)	(0.648)		
YTM(m-1)			0.008	-0.084		
			(0.315)	(0.374)		
Observations	146,270	135,615	105,305	96,280	146,270	135,615
Adjusted R-squared	0.402	0.457	0.364	0.432	0.874	0.894
Category-by-Date FE	YES	NO	YES	NO	YES	NO
Match bmrk-by-Date FE	YES	NO	NO	NO	NO	NO
Category-by-Match bmrk-by-Date FE	NO	YES	NO	YES	NO	YES
Controls	YES	YES	YES	YES	YES	YES

Panel B: YTM by benchmark category

	Yield to maturity	
VARIABLES	(1)	(2)
Misbenchmarked	0.769***	0.070
	(0.197)	(0.189)
Observations	97,261	105,638
Adjusted R-squared	0.443	0.458
PPB-by-Date FE	YES	NO
Best-matched-benchmark-by-Date FE	NO	YES
Controls	YES	YES

Table A5. Persistence in Portfolio Characteristics

This table reports the results of monthly regression of fund characteristics on 12-month lagged characteristics. The dependent variable in columns 1 and 2 is fund duration in month t. In Columns 3 and 4 (5 and 6) the dependent variable is % Corporate (% Treasury), which is the percentage of fund assets invested in corporate (Treasury) bonds. In even columns, we include only meaningfully misbenchmarked funds, representing those whose return correlation would be improved by at least 0.10 by switching to the proper benchmark; in odd columns, we include all other funds only. All regressions include lagged monthly controls for expense ratio, fund turnover, log of fund assets, log of fund age, and flow, but the coefficients are not reported. Each column uses Lipper category by benchmark by date fixed effects. Standard errors clustered at the fund and date level are presented in parenthesis below the coefficient.

	Portfolio Duration t		% Corporate t		% Treasury t	
	(1)	(2)	(3)	(4)	(5)	(6)
Portfolio Duration t-12	0.774***	0.838***				
	(0.024)	(0.026)				
% Corporate t-12			0.813***	0.909***		
			(0.015)	(0.011)		
% Treasury t-12					0.772***	0.712***
					(0.030)	(0.038)
Observations	46,799	18,879	32,341	25,163	32,341	25,163
Adjusted R-squared	0.696	0.938	0.839	0.862	0.754	0.760
Category-by-PPB-by-Date FE	YES	YES	YES	YES	YES	YES
Misbenchmarked	NO	YES	NO	YES	NO	YES

Table A6. Persistence in Pseudo Benchmark Weights

This table reports the results of deviation between the weights for the best matched pseudo benchmark compared to the weights in the primary prospectus benchmark. A negative (positive) number indicates that a fund's chosen benchmark has a higher (lower) weight than the best matched pseudo benchmark. Weights are allocated, in increments of 0.1, to a combination of the Bloomberg U.S. Aggregate Index (AGG), the Bloomberg U.S. Corporate Index (IG), and the ICE BofA U.S. High Yield Index (HY). Panel A calculates the deviation based on the best matched pseudo benchmark for the entire sample, while Panel B calculates the average deviation using rolling 36-month periods.

Panel A: Best matched pseudo bmrk based on full sample

Deviation	Mean	SD	Min	p10	p25	p50	p75	p90	Max
AGG	-0.303	0.363	-1	-0.9	-0.5	-0.2	0	0	0.7
IG	0.117	0.221	-0.9	0	0	0	0.2	0.4	1
HY	0.185	0.267	-0.5	0	0	0.1	0.3	0.6	1

Panel B: Best matched pseudo bmrk based rolling 36-months

Deviation	Mean	SD	Min	p10	p25	p50	p75	p90	Max
AGG	-0.297	0.382	-1	-1	-0.6	-0.2	0	0	0.9
IG	0.118	0.250	-1	0	0	0	0.2	0.5	1
HY	0.179	0.284	-0.9	0	0	0.1	0.3	0.6	1