NBER WORKING PAPER SERIES

DON'T TAKE THEIR WORD FOR IT: THE MISCLASSIFICATION OF BOND MUTUAL FUNDS

Huaizhi Chen
Lauren Cohen
Umit Gurun

Working Paper 26423
http://www.nber.org/papers/w26423

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2019

We thank Rob Battalio, Geoffrey Booth, Ilan Guttman, Tim Loughran, Bill McDonald, Rabih Moussawi, Bugra Ozel, Joshua Ronen, Stephen Ryan, Paul Zarowin, and seminar participants at New York University and the University Notre Dame for comments. We also thank James Ng for providing valuable research assistance. We are grateful for funding from the National Science Foundation, SciSIP 1535813. Please send correspondence to Huaizhi Chen (University of Notre Dame, 238 Mendoza College of Business, Notre Dame, IN 46556, phone: 574-631-3385, email: hchen11@nd.edu), Lauren Cohen (Harvard Business School, Baker Library 279, Soldiers Field, Boston, MA 02163, phone: 617-495- 3888, email: lcohen@hbs.edu), Umit G. Gurun (University of Texas at Dallas, 800 W Campbell Rd, SM 41, Richardson, TX 75080, phone: 972-8835917, email: umit.gurun@utdallas.edu).

Don’t Take Their Word For It: The Misclassification of Bond Mutual Funds
Huaizhi Chen, Lauren Cohen, and Umit Gurun
NBER Working Paper No. 26423
November 2019
JEL No. G11,G12,G23,G24,G4,K0

ABSTRACT

We provide evidence that mutual fund managers misclassify their holdings, and that these misclassifications have a real and significant impact on investor capital flows. In particular, we provide the first systematic study of bond funds’ reported asset profiles to Morningstar against their actual portfolios. Many funds report more investment grade assets than are actually held in their portfolios, making these funds appear significantly less risky. This results in pervasive misclassifications across the universe of US fixed income mutual funds by Morningstar, who relies on these reported holdings. The problem is widespread-resulting in about 30% of funds being misclassified with safer profiles, when compared against their actual, publicly reported holdings. “Misclassified funds” – i.e., those that hold risky bonds, but claim to hold safer bonds– outperform the actual low-risk funds in their peer groups. “Misclassified funds” therefore receive higher Morningstar Ratings (significantly more Morningstar Stars) and higher investor flows due to this perceived outperformance. However, when we correctly classify them based on their actual risk, these funds are mediocre performers. Misreporting is stronger following several quarters of large negative returns, and it is strong at the fund family level. We report those families that have the highest percentage of misreported funds in the sample.

Huaizhi Chen
University of Notre Dame
238 Mendoza College of Business
hchen11@nd.edu

Lauren Cohen
Harvard Business School
Baker Library 273
Soldiers Field
Boston, MA 02163
and NBER
lcohen@hbs.edu

Umit Gurun
University of Texas at Dallas
School of Management
800 W Campbell Rd. SM41
75080 Richardson, TX
umit.gurun@utdallas.edu

I. Introduction

Information acquisition is costly. However, the exact cost of collecting information on any given asset depends on timing, location, a person's private information set, etc. This is in addition to the idiosyncratic characteristics and complexities of the information signal itself. External agents – both public and private - have emerged to fill this role and reduce the cost of information acquisition. However, the value of these agents depends on how much additional information provision is needed. To this end, delegated portfolio management is the predominant way in which investors are being exposed to both equity and fixed income assets. With over 16 trillion dollars invested, the US mutual fund market, for instance, is made up of over 5,000 delegated funds and growing. While the SEC has mandated disclosure of many aspects of mutual fund pricing and attributes, different asset classes are better (and worse) served by this current disclosure level. Investors have thus turned to private information intermediaries to help fill these gaps.

In this paper, we show that for one of the largest markets in the world, US fixed income debt, this had led to large information chasms that have been filled by strategic-response information provision by funds. In particular, we show that this reliance on the information intermediary has resulted in systematic misreporting by funds. This misreporting has been persistent, widespread, and appears strategic – casting misreporting funds in a significantly more positive position than is in actuality. Moreover, the misreporting has real impact on investor behavior and mutual fund success.

Specifically, we focus on the fixed income mutual fund market. The entirety of the fixed income market is similarly sized as equites (e.g., 40 trillion dollars compared with 30 trillion dollars in equity assets worldwide). However, bonds are fundamental different as an asset, which are reflected in their delegated portfolios, as well. While equity funds hold predominantly the same security type (e.g., the common stock of IBM, Apple, Tesla, etc.), each of a fixed income funds' issues differ in yield, duration, covenants, etc. – even across issues of the same underlying firm - making them more bespoke and unique. Moreover, the average active equity fund holds roughly 100 positions, while the average active fixed income fund holds over 600 issues. While the SEC mandates disclosure of the portfolio constituents, this data is more complex in aggregating to measures of the fund itself.

This has led fund private information intermediaries to provide a level of aggregation and summary on the general riskiness, duration, etc. of fixed income funds that investors rely upon. We focus on the largest of intermediaries that provide data on categorization and riskiness at the fund level – Morningstar, Inc. In particular, we provide the first systematic study that compares fund reported asset profiles provided by Morningstar against their *actual* portfolio holdings. We find significant misclassification across the universe of all bond funds. This is on the order of roughly 30% of all funds (and rising) in recent years, and is pervasive across the funds being reported as overly safe by Morningstar.

How do these misclassifications occur? Morningstar "rates" each fixed income mutual fund into style boxes based on measures of risk. This – along with expenses, and other proprietary measures– are then used to classify and rank funds, and an aggregate rating is given in the form of "Morningstar Stars."[1] These Morningstar Star Ratings have been shown throughout the literature to have a strong and significant impact on investor flow from both retail and institutional investors

[1] The ratings methodology and proprietary adjustments and assumptions (e.g., tax burden) are described here: https://www.morningstar.com/content/dam/marketing/shared/research/methodology/771945_Morningstar_Rating_for_Funds_Methodology.pdf, but to a first-order approximation, the rating is determined by their risk and net return categorization (with high expenses detracting from net returns).

(Del Guercio and Tkac (2008), Evans and Sun (2018), Reuter and Zitzewitz (2015)).[2] In addition, they are used ubiquitously throughout the industry. For example, Figure 1 shows a fund fact page from Fidelity that lists prominently Morningstar's metrics about its fund offerings.

The central problem that we show empirically, however, is that Morningstar itself has become overly reliant on summary metrics, leading to significant misclassification across the fund universe. In particular, Morningstar requires data provision from each fund it rates (and categorizes) on the breakdown of the bonds the fund holds by risk rating classification. Specifically, what percentage of the fund's current holdings are in AAA bonds, AA bonds, BBB bonds, etc. One might think that Morningstar uses these Summary Reports data to augment the detailed holdings it acquires from the SEC filings on the firm. However, Morningstar makes its risk classifications, fund style categorizations, and even fund ratings, solely based on this self-reported data.

Now this would be no issue if funds were truthfully passing on a realistic view of the fund's actual holdings to Morningstar. Unfortunately, we show that this is not the case. We provide robust and systematic evidence that funds on average report significantly safer portfolios than they actually (verifiably) hold. In particular, funds report holding significantly higher percentages of AAA bonds, AA bonds, and all investment grade issues than they actually do. For some funds, this discrepancy is egregious – with their reported holdings of safe bonds being 100% while their r holdings are only 0.05% of their portfolios, as was the case with *CMG Tactical Fund* in the first quarters of 2018. Due to this misreporting, funds are then misclassified by Morningstar into safer categories than they otherwise should be.

[2] Investors also respond to other attention grabbing and easy to process external ranking signals, such as Wall Street Journal (Kaniel and Parham, 2017) and sustainability rankings (Hartzmark and Sussman, 2018).

We define "Misclassified Funds" in a straightforward way: namely as those funds that are classified into a different category than they should be if their *actual* holdings were used as opposed to the Summary Report percentages that are used to classify them. We show that misclassification is widespread, and continues through present-day, rising to 33% of high and medium credit quality funds in 2018. Moreover, as mentioned above misclassifications are overwhelmingly one-sided: 1% of all misstatements push funds down a category – 99% of misstatements push up to a safer category.

So, what are the characteristics of these "Misclassified Funds?" First, Misclassified Funds have higher average risk - and accompanying yields on their holdings - than its category peers. This is not completely surprising, as again Misclassified Funds are holding riskier bonds than the correctly classified peers in their risk category. Importantly, this translates into significantly higher returns earned by these Misclassified Funds relative to peer funds. They earn 10.3 basis points (t=2.64) per quarter more, implying a 14% higher return than peers.

In order to estimate what portion of this seeming return outperformance of Misclassified Funds comes from skill versus what comes from the unfair comparison to safer funds, we turn to the funds' *actual* holdings reported in their quarterly filings to the SEC. We use these actual holdings to calculate the correct risk category that the fund should be classified into were it to have truthfully reported the percentage of holdings in each risk category. When we re-run the same performance regression specification, but using correct peer-comparisons, we find that Misclassified Funds no longer exhibit any outperformance. In point estimate they even *underperform* by -11 basis points per quarter (t=1.64). Thus, it appears that 100% of the apparent outperformance of Misclassified Funds is coming from being misclassified to a less risky comparison group of funds.

However, the Misclassified Funds still reap significant real benefits from this incorrectly ascribed outperformance. Perhaps most impressively, Morningstar actually seems to fool itself, as they reward these Misclassified Funds with significantly more Morningstar Stars. In particular, these Misclassified Funds receive an additional 0.34 stars (t=5.75), or a 10.4% in the number of stars. Armed with higher returns relative to (incorrect) peers and higher Morningstar Ratings, Misclassified Funds then are able to charge significantly higher expenses. In particular, they charge expense ratios that are 9.98% higher than peers (t=5.77).

Lastly, we estimate to what extent – even with these higher fees – Misclassified Funds might be able to attract more investor flows due to the favorable comparison benefits of being misclassified. In order to do this, we run a two stage least squares procedure. In the first stage, we estimate – controlling for other fund, category, and time effects – the impact of being a Misclassified Fund on the number of Morningstar Stars that a fund receives. We then take this estimate of *just* the extra portion of Morningstar Stars a Misclassified Fund gets from being misclassified against a lower risk peer group, and take just this piece of their Stars – Misclassified Stars - to see if it has an impact on investor flows. We find that it has a significantly positive impact. In particular, a one Misclassified Star increase raises the probability of positive flows by almost 14% (t=2.45).

Stepping back, what makes this even somewhat more surprising is that funds actually *do* report holdings directly to Morningstar, and these holdings line up almost perfectly with the SEC-downloaded holdings. Thus, it is literally that Morningstar uses the Summary Reports itself (and not the other data also delivered directly to it by funds) instead of taking the extra step of calculating riskiness itself that contributes to classification.

The remainder of the paper proceeds as follows. Section II provides background for our study, while Section III describes the data, and methodology that Morningstar uses to classify funds into categories. Section IV then presents our main results on the misreporting of funds, and misclassification of these funds by Morningstar based on these faulty reports. Section IV also documents the return implications, along with the real benefits for funds in terms of expenses, Morningstar Stars, and investor flows. Section V concludes.

II. Background

Our results primarily contribute to three lines of literature. First, our evidence is related to studies on the implications of accuracy and completeness of data sources. Along these lines, Ljungqvist, Malloy, and Marston (2009) show that I/B/E/S analyst stock recommendations have various changes across vintages and these changes (alterations of recommendations, additions and deletions of records, and removal of analyst names) are non-random and likely to affect profitability of trading signals, e.g. profitability of consensus recommendation, among others. Other examples include Rosenberg and Houglet (1974), Bennin (1980), Shumway (1997), Canina et al. (1998), Shumway and Warther (1999), and Elton, Gruber, and Blake (2001). The asset management literature also documents biases in reporting. In the hedge fund setting, Bollen and Poole (2009, 2012) exploit a discontinuity at 0% for reported returns by fund managers (i.e., investors view 0% as a natural benchmark for evaluating hedge fund performance) and document a discontinuous jump in capital flows to hedge funds around this zero-return cut-off. There is also recent work that shows the mutual funds also exhibit considerable variation in their month-end valuations of identical corporate bonds (Cici, Gibson and Merrick, 2011). Similar biases have been shown for valuation of private companies by mutual funds (Agarwal, et al. 2019). Likewise, Choi, Kronland and Oh (2018) show that zero returns are prevalent in fixed income funds and that zero-return reporting is essentially driven high illiquidity of fund holdings.

Second, our finding on the association between misclassification and performance is related to studies on deviations from stated investment policies by equity funds. For example, Wermers (2012), Budiono and Martens (2009) and Swinkels and Tjong-A-Tjoe (2007) show that equity mutual funds that drift from the stated investment objective do better than counterparts. Brown, Harlow and Zhang (2009) and Chan, Chen, and Lakonishok (2002) show that funds that exhibit discipline in following a consistent investment mandate outperform less consistent funds. More recently, Bams, Otten, and Ramezanifer (2017) study performance and characteristics of funds that deviate from stated objectives in the prospectuses.

Lastly, our study contributes to the literatures on style investment. Barberis and Shleifer (2003) argue that investors tend to group assets into a small number of categories, causing correlated capital flows and correlated asset price movements. Vijh (1994) and Barberis, Shleifer, and Wurgler (2005) provide examples using S&P 500 Index membership changes. Other examples in the empirical literature include Froot and Dabora (1999), Cooper, Gulen, and Rau (2005), Boyer (2011), and Kruger, Landier, and Thesmar (2012), who find that mutual fund styles, industries, and countries all appear to be categories that have a substantial impact on investor behavior (and asset price movements). Our work complements these studies by showing that investors categorize bond funds along the credit risk dimension as provided by the mutual fund industry's primary data source, Morningstar.

III. Data

In this section, we describe in detail the three aspects of data used in this paper. Specifically, we combine (1) the Morningstar Direct database of mutual funds and their characteristics, (2) the

Morningstar database of Open-Ended Mutual Fund Holdings, and (3) our assembled collection of credit rating histories to document the substantial gap between the reported and the true portfolio compositions in fixed income funds.

III.1 The Morningstar Direct Database

Morningstar Direct contains our collection of fixed income mutual funds. These are US domicile, dollar denominated, mutual funds that belongs to the "US Fixed Income" global category. We filter out US government, agency, and municipal bond funds using lagged Morningstar sub-categories. The final collection is 1,294 unique fixed income mutual funds that span Q1 2003 to Q2 2019. The fund characteristics that underlying these funds also come from Morningstar Direct. This data service contains detailed characteristics that originates both from the regulatory open-ended mutual fund filings and from direct fund surveys. A key element of our study is the reported asset compositions from mutual fund companies. Figure 2 displays the survey used by Morningstar to collect this information from managers. Since the first quarter of 2017, Morningstar began calculating percent asset compositions directly from holdings, but as recently as August 2019, still use the surveyed compositions to place fixed income funds in style categories. Notably, we also obtain historical returns, share-level investor flow, and fixed income fund styles from this dataset. For a full list of variables used in this study, refer to Appendix A.

III.2. Open-Ended Mutual Fund Holdings

Our open-ended mutual fund holdings come directly from Morningstar. This service provides us with linkages of portfolio holdings to the Morningstar Direct funds. The fixed income portfolio positions are identified by fundid, security name, CUSIP, and portfolio date. Along with the identity of these positions, we use portfolio weight, long/short profile, and asset type from this data.

III.3. Credit Rating Histories

Our analysis centers on the misrepresentation of credit risk in reports heavily used by investors, therefore we collect credit rating histories from a large variety of data sources in order to achieve comprehensive coverage. Due to Dodd-Frank, credit rating agencies are required to post their rating histories as XBRL releases. These releases help us achieve coverage by Standard & Poor, Moody's, and Fitch of all CUSIP-linked securities after June 2012. Prior to this date, we obtain credit ratings from the Capital IQ and the Mergent FISD databases. Capital IQ contains credit rating histories from Standard & Poor for all of our sample history. In addition, Mergent FISD provides coverage of credit ratings from Moody's, Standard and Poor, and Fitch on corporates, supranational, agency, and treasury bonds. Table 1 Panel A lists these data sources, the rating agencies reported in these sources, and the time span of their respective coverage. Panel B and Panel C tabulates the actual (as calculated using our credit rating histories) and the reported percentage holding compositions of fixed income mutual funds in the various credit rating categories from Q1 2003 to the end of each respective samples.

Table 2 tabulates the time series of Fund Quarter observations in each Morningstar Credit Quality Category. The last column is the number of misclassified observations. Morningstar changed the way it calculated average credit risk in August 2010. Prior to this date, the average credit risk is a simple weighted average of the underlying linear bond scores, in which a AAA bond has a score of 2, AA has a score of 3, and so on. After August 2010, the credit risk variable attempts to describe a fund in terms of the returns and risks of a portfolio of rated bonds, and nonlinear scores are assigned to each category. We describe the weighing scheme in accordance to Morningstar's methodology in Appendix B. The result of this change in methodology is a much more composition dependent categorization of fixed income funds, and an increased number of mis-classified fund as according to the Morningstar Fund Style classification.

IV. Main Results

IV.1. Diagnostics Analysis

We start our analysis by examining histograms of fund reported percentage of holdings minus the calculated percentage holdings in various bond credit rating categories (Figure 3) between Q1 2017 and Q2 2019. The start of this diagnostic sample is dictated by the time Morningstar began calculating the percent holdings of assets in each credit risk category per each fixed income fund. Ideally, if Morningstar and the bond funds in its database kept the same reporting standards in credit ratings, the fund reported percent should be almost same as the calculated percent holdings. Therefore, these histograms should report a sharp spike around zero, and exhibit no significant variation. This simple diagnostic shows that, on the contrary, there is wide dispersion of discrepancy between the records of asset compositions. Most notably, for the assets above the investment grade (above BBB), the percentage of assets reported by funds is markedly higher than the percentage of assets calculated by Morningstar. When we check the same gap for below investment grade and especially in unrated assets, we see an opposite pattern, i.e. the percentage of assets reported by funds is significantly lower than the percentage of assets calculated by Morningstar.

In the remainder of our analysis, we will drill down the reasons as to why and how these patterns emerge. First, we will expand the sample to a larger period (e.g. to 2003) by calculating percent holdings of assets in each credit risk category directly from each fund portfolio. We will then investigate how these systematic patterns of over/under reporting affect the classification and marketing of risk profiles by mutual funds, through Morningstar, toward investors. Finally, we will investigate the implications of this reporting gap with respect to performance, expenses, and finally fund flows.

IV.2. Misclassification

In this subsection, we look at the major implication of the difference between reported and actual holding implied composition of fund portfolios. When a fund reports high levels of investment grade assets, it will get classified as an investment grade fund regardless of its actual holdings. We find that in our sample, 15% to 25% of the bond funds that were classified as high or medium in their credit risk quality are under-classified by Morningstar.

Specifically, we combine the credit rating history on each fixed income asset in every fund portfolio to calculate the actual percentage of assets held in each credit risk category. Then we follow the Morningstar methodology to calculate holding composition implied average credit risk for the entire fund.

In Figure 4, we plot the credit risk distribution of fund-quarter observations between first quarter of 2017 and the end of the second quarter of 2019. The dashed lines represent breaks in the fixed income fund style-box. AAA and AA credit quality funds are high credit quality; A and BBB credit quality funds are medium credit quality; and BB and B are low credit quality as deemed by Morningstar.

The first (blue) bar depicts the distribution of the actual credit risk category that Morningstar assigns to US Fixed Income funds. In other words, blue line is what mutual fund investors observe if they use Morningstar as a data provider. Per Morningstar's methodology, this official credit risk category is calculated as a function of the fund survey reported percentage holdings of assets in the various credit risk categories. That is, the average credit risk category assigned to each fund is scored on its percentage of holdings in AAA assets, percentage in AA, etc. Indeed, when we calculate the average credit risk of a fund using their self-reported percentage holdings (the second bar in red), we see almost an exact overlap.

In the third bar (gray), we calculate the counter-factual credit risk category that would result if we had used Morningstar calculated percentage holdings. In other words, if Morningstar relied on the holdings compositions it had already calculated using the "Morningstar database of Open-Ended Mutual Fund Holdings", it would have computed a vastly different categorization of fund-level credit risk. Many fixed income mutual funds would have fallen into a higher credit risk bucket in this counterfactual credit risk category.

In the final bar (yellow), we report the average credit rating that we would have assigned to funds by directly using their quarterly holdings. In calculating this bar, we combine our collection of the credit rating histories, the portfolio information for each bond mutual fund, and the Morningstar methodology for aggregating portfolio credit risk. The sample is limited from Q1 2017 to Q4 2018 because our portfolio data ends in Q4 2018.

Comparison of these four distributions clearly indicates that using fund self-reported credit risk composition has widely skewed the fund-level credit categorization in favor of lower credit risk. For example, almost half of funds that are marked as A should not be in this category if the fund-level credit rating were assigned based on actual holdings rather than self-reported compositions. Likewise, half of the AAA rated funds should have received a riskier categorization according to the counter factual holding implied aggregation.

IV.3. Fund Performance and Misclassification

In Table 3, we analyze whether the misclassified funds tend to have above average yields for their credit risk categories. Higher yields typically imply higher average risk. Specifically, we regress the various yield metrics on the misclassified dummy. We define a *Misclassified* dummy variable which take a value of one if the official credit quality (High or Medium) is higher than the counter factual credit quality, and zero otherwise. We use three different types of yield metric. In the first column, we use yields reported to Morningstar by the funds themselves. These yields are voluntarily reported. In the second column, we use the yields calculated by Morningstar. The sample size in this second column is limited because Morningstar began calculating the holding yields in 2017. In the last column, we use twelve-month yield which combines total interest, coupon, and dividend payments. In this last analysis, our sample period starts in the first quarter of 2003 and ends at the last quarter of 2018. We include duration of bonds (as reported by the funds) as a control variable to capture the interest rate risk of the bond portfolio. Most importantly, we include a Time x Official Fund Style fixed effect to our specification which absorbs the mean yield of each funds corresponding fund style at a given year. Doing so allows us to address the concern that a group of funds in a particular year systematically misclassify their riskiness and that misclassified dummy essentially captures this fund style related reporting choice. We cluster the standard errors by quarter and fund to address the time series cross-sectional and individual variation in risk. In all these three tests, there is a strong relation between misclassification and yields: Misclassified funds have higher yields. The annualized reported yield to maturity is 43.1 basis points higher (t = 9.16), whereas the calculated yield from the holdings (second column) and the payout yield are 27.1 and 27.2 basis points higher respectively for misclassified funds over their official peers.

To understand the implication of higher risk in the underlying holdings, we regress actual fund returns on the *Misclassified* dummy, reported duration, as well as two sets of fixed effects. In the first regression, we include Time x Official Fund Style fixed effect as we do in the previous table. In the second regression, we replace this fixed effect with Time x Counterfactual Fund Style fixed effect. The idea is to see to what extend the *Misclassified* dummy captures the return gap between the reported fund style and the counterfactual fund style. Put differently, because the misclassification occurs with a specific direction, i.e. more risky funds are reported to have lower risk when in fact they belong to a riskier category, the additional return performance of *Misclassified* funds in their reported category is due to the risk gap between counterfactual and reported fund styles. The results reported in Table 4 confirm this intuition. The statistical significance of the *Misclassified* dummy disappears when *Time x Official Fund Style* is replaced by *Time x Counterfactual Fund Style* fixed effect, i.e. any excess return earned by *Misclassified* funds can actually be attributed to the performance of its true (e.g. counterfactual) category.

IV.4. Incentives to Misclassify

In our next analysis, we test whether misclassified funds obtain various benefits from being classified in the less-riskier groups of funds. One such benefit is that they get better Morningstar star ratings that can be used as a powerful marketing tool. We test this hypothesis formally by regressing various Morningstar rating metrics on the *Misclassified dummy*, reported duration, average expense ratio, and Time x Official Fund Style fixed effect. This fixed effect absorbs the mean Morningstar rating given that that particular fund style in a given year. Because the ratings and expenses are reported at the share class level, the fund level Morningstar Ratings and the Average Expense ratio are calculated as the value weighted average of their respective share-class level values. In this analysis, we use the full sample periods, i.e. 2003 to 2018 and cluster t-statistics at quarterly level just like we did in the return analysis.

The results reported in Table 5 shows that there are economically large benefits for a risker fund to be pooled in a less-risky category. *Misclassified* funds receive 0.18 to 0.34 higher Morningstar stars compared to their peer funds. This amount of higher rating corresponds to 20% to 37% of the one standard deviation in Morningstar ratings.

In Table 6, we investigate whether misclassified funds have higher expense ratios than their peers. The idea here is simply to test if these funds charge higher expenses to their investors because their "reported" (but not actual) performance is better and that they have higher Morningstar ratings.

Prior research has paid considerable attention to the question of whether equity mutual funds are able to consistently earn positive risk-adjusted returns.[3] Given that mutual fund fees pay for the services provided to investors by the fund and because the main service provided by a mutual fund is portfolio management, fees should reflect funds' risk-adjusted performance. This line of arguments suggests there should be a positive relation between before-fee risk-adjusted expected returns and fees. On the other hand, Gil-Bazo and Ruiz-Verdu (2009) argue funds of ten engage in strategic fee-setting in the presence of investors with different degrees of sensitivity to performance and this could lead to a negative relation between fund performance and fee. In the second and third column, we repeat our analysis with duration of the fund and fund return in the specification. By explicitly controlling for risk and fund return, we are able to capture the component of misclassification that is attributed to these two factors. As reported in Table 6, we

[3] See, for example, Brown and Goetzmann (1995); Carhart (1997); Daniel et al. (1997); Wermers (2000); Cohen, Coval, and Pastor (2005); Kacperczyk, Sialm, and Zheng (2005); Kosowski et al. (2006).

find that, on average, the misclassified funds have 9.6 basis point higher ($t =$ 5.71) average annual expenses than funds within the same style-category and that our results remain when we add proxy to control for the riskiness and returns of the fund.[4]

In Table 7, we investigate whether fund flows to misclassified funds differ from the rest of the funds in an economically significant way. There are several reasons why misclassification matters for the bond fund flows. First, Barberis and Shleifer (2003) argue that investors tend to group assets into a small number of categories, causing correlated capital flows and correlated asset price movements. If an asset ends up being in the wrong bucket then it may receive a disproportionately higher (or lower) investments that its original bucket. Several papers in the literature show the power of style investment in explaining asset flows. Froot and Dabora (1999), Cooper, Gulen, and Rau (2005), Boyer (2011), and Kruger, Landier, and Thesmar (2012), find that mutual fund styles, industries, and countries all appear to be categories that have a substantial impact on investor behavior (and asset price movements).

It is hard to conclude a causal link between misclassification and its flow implications because of endogeneity, i.e. while the investors are responding to excess performance of excess funds, it is also plausibly possible that fund flows lead to strategic reporting of funds that leads to misclassification. One way to get around this problem is to use the link between misclassification and Morningstar ratings that is closely followed as the key metric to allocate funds. Specifically,

[4] Past research in the equity space has investigated whether funds alter their investment style and whether funds with characteristics are more likely to deviate from stated objectives in their mandate due to various reasons including fund manager incentives. In particular, DiBartolomeo and Witkowskip (1997) show that younger mutual funds are particularly prone to misclassification and Frijns et al. (2013) show that funds which switch across fund objectives aggressively tend to have higher expense ratios. Along these lines, Huang, Sialm and Zhang (2011) argue that funds with higher expense ratios experience more severe performance consequences when they alter risk.

we use misclassification as an instrument to predict Morningstar ratings and use this predicted Morningstar ratings to see if variation in predicted Morningstar ratings (Morningstar ratings solely attributable to misclassification) generates significant flows. More formally, the estimates reported in Table 5 help us calculate how much extra "Stars" a Misclassified Fund gets from being misclassified against too easy of a peer group. We use these Misclassified Stars to see if it has an impact on investor flows, i.e. we regress the direction of investor flows into funds on predicted component obtained in the first stage. In this analysis, we include reported duration, expense ratio, Time x Official Fund style, and Fund Fixed effects as control variables. The results in Table 7 suggests that there is an economically large relationship between Morningstar ratings and the sign of flows. One notch higher rating that is solely attributed to misclassification increases the changes of higher flows by 13.8% (t = 2.45) which suggests – even with these higher fees – Misclassified Funds are be able to attract more investor flows due to the favorable comparison benefits of being misclassified.

In our final test, we investigate the correlation between the funds decision to misclassify their holdings and their past fund performance after controlling for fund size and fund style fixed effects. We define two variables to capture the change in fund reporting behavior over time. The first variable, *Starts*, takes a value of one if a previously correctly classified fund starts misclassifying its holdings. In addition to this variable, we define another indicator variable, *Stops*, which takes a value of one if a previously misclassified fund starts correctly classifying its holdings. The negative coefficient for 3-year return, -0.140 (t = 2.27), indicates that funds with worse performance, in comparison to better performing funds, are more likely to start misclassifying their holdings (Column 1, Table 8). Likewise, the corresponding coefficient in the second column, 0.309

(t = 4.31), suggests funds with better past returns are more likely to stop misclassification. These results collectively suggest that misclassification strategy does vary with fund performance.

IV.5. Fund Family level evidence of Misclassified Funds

Lastly, we aggregate the extent of misclassified funds by the fund family level to examine whether there is any family level correlation in the extent to which funds appear to misstate holdings to a large enough extent to cause misclassification. We find that this is indeed the case. In particular, we find that some fund families essentially never misstate holdings, while some families engage in misstatement regularly – and for nearly all of the funds in the their family. The list of the most frequent misstating and misclassified funds is in Table 9. As can be seen, 100% of certain families' fixed income funds are misclassified into safer categories as compared to what their actual holdings imply.

V. Conclusion

Investors rely on external information intermediaries to lower their cost of information acquisition. While *prima facie* this brings up no issues, if the information that the intermediary is passing on is biased, these biases propagate throughout markets and can cause real distortions in investor behavior and market outcomes. We document precisely this in the market for fixed income mutual funds. In particular, we show that investors' reliance on Morningstar has resulted in significant investment based on verifiably incorrect reports by fund managers that Morningstar simply passes on as truth.

We provide the first systematic study that compares fund reported asset profiles provided by Morningstar against their *actual* portfolio holdings, and show evidence of significant misclassification across the universe of all bond funds. A large portion of bond funds are not truthfully passing on a realistic view of the fund's actual holdings to Morningstar and Morningstar creates its important risk classifications, fund categorizations, and even fund ratings, based on this self-reported data. Roughly 30% of all funds (and rising) in recent years, are reported as overly safe by Morningstar. This misreporting has been not only persistent and widespread, but also appears to be strategic. We show that misclassified funds have higher average risk - and accompanying yields on their holdings - than its category peers. We also show evidence suggesting the misreporting has real impacts on investor behavior and mutual fund success. Misclassified funds reap significant real benefits from this incorrectly ascribed outperformance in terms of being able to charge higher fees, receiving "extra" undeserved Morningstar Stars, and ultimately receiving higher flows from investors.

Stepping back, as the costs of producing, disseminating, and delivering information continue to fall, we have seen firms continue to ramp up the production of such information. Given this, the need for information aggregation will only accelerate. We exploit a novel setting in which investors reliance on external information intermediaries can lead to predictable patterns in fund ratings and capital flows, *and* in which we can ex-post verify the veracity of the information conveyed. We believe that our study is a first step to think about a market design in which information intermediaries have more aligned incentives to better package the information they gather from market constituents. Future research should explore alternate monitoring and verification mechanisms for increasingly complex information aggregation in financial markets, and ways that investors can engage as important partners in information collection and price-setting in modern capital markets.

References

Agarwal, V., Barber, B.M., Cheng, S., Hameed, A. and Yasuda, A., 2019. Private company valuations by mutual funds. Available at SSRN 3066449.

Barberis, N., Shleifer, A., 2003. Style investing, Journal of Financial Economics 68, 161–199.

Brown, Stephen J., and William N. Goetzmann, 1997. Mutual fund styles, Journal of Financial Economics 43, no. 3, 373-399.

Brown, Keith C., Harlow, W. Van and Zhang, Hanjiang, 2009. Staying the course: The role of investment style consistency in the performance of mutual funds. Available at SSRN 1364737.

Bams, Dennis, Otten, Roger, and Ramezanifar, Ehsan, 2017. Investment style misclassification and mutual fund performance. In 28th Australasian Finance and Banking Conference.

Bennin, Robert, 1980. Error rates in CRSP and COMPUSTAT: A second look, Journal of Finance 35, 1267–1271.

Boyer, Brian H., 2011. Style-related comovement: Fundamentals or labels?, Journal of Finance 66, 307-332.

Bollen, Nicholas and Pool, Veronica, 2009. Do Hedge Fund Managers Misreport Returns? Evidence from the Pooled Distribution, Journal of Finance, 2257-2288.

Bollen Nicholas and Pool, Veronica, 2012. Suspicious Patterns in Hedge Fund Returns and the Risk of Fraud, Review of Financial Studies 25, 2673-2702.

Budiono, Diana and Martens, Martens, 2009. Mutual fund style timing skills and alpha. Available at SSRN 1341740.

Canina, Linda, Roni Michaely, Richard Thaler, and Kent Womack, 1998. Caveat compounder: A warning about using the daily CRSP equal-weighted index to compute long-run excess returns, Journal of Finance 53, 403–416.

Carhart, Mark M, 1997. On persistence in mutual fund performance. The Journal of Finance 52, no. 1, 57-82.

Chan, Louis K., Chen, Hsiu-Lang, and Lakonishok, Joseph, 2002. On mutual fund investment styles. The Review of Financial Studies 15, pp.1407-1437.

Choi, Jaweon, Kronlund, Mathias and Oh, Ji Y.J., 2018. Sitting Bucks: Zero Returns in Fixed Income Funds, Working Paper.

Cici, Gjergji, Gibson, Scott and Merrick Jr, John J., 2011. Missing the marks? Dispersion in corporate bond valuations across mutual funds. Journal of Financial Economics, 101(1), pp.206-226

Cohen, Randolph B., Coval, Joshua D. and Pástor, Lubos, 2005. Judging fund managers by the company they keep. The Journal of Finance, 60(3), pp.1057-1096. Vancouver

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

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

Del Guercio, Diane, and Paula A. Tkac, 2008. Star power: The effect of Morningstar ratings on mutual fund flow, Journal of Financial and Quantitative Analysis 43, 907-936.

Di Bartolomeo, Dan, and Erik Witkowski, 1997. Mutual fund misclassification: Evidence based on style analysis, Financial Analysts Journal 53, no. 5, 32-43.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 2001. A first look at the accuracy of the CRSP Mutual Fund Database and a comparison of the CRSP and Morningstar Mutual Fund Databases, Journal of Finance 56, 2415–2430.

Evans, Richard B., and Yang Sun, 2018. Models or stars: The role of asset pricing models and heuristics in investor risk adjustment, Working paper, University of Virginia.

Frijns, Bart, Aaron B. Gilbert, and Remco CJ Zwinkels, 2013. On the Style-based Feedback Trading of Mutual Fund Managers. Available at SSRN 2114094.

Froot, Kenneth, Dabora, Emil, 1999. How are stock prices affected by the location of trade? Journal of Financial Economics 53, 189–216.

Gil-Bazo, Javier and Pablo Ruiz-Verdu, 2009. The relation between price and performance in the mutual fund industry, The Journal of Finance 64, no. 5, 2153-2183.

Hartzmark, Samuel M., and Abigail Sussman, 2018. Do investors value sustainability? A natural experiment examining ranking and fund flows, Working paper, University of Chicago.

Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang, 2011. Risk shifting and mutual fund performance. Review of Financial Studies 24, no. 8, 2575-2616.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2008. Unobserved actions of mutual funds, Review of Financial Studies 21, no. 6, 2379-2416.

Kaniel, Ron, and Robert Parham, 2017. WSJ category kings: The impact of media attention on consumer and mutual fund investment decisions, Journal of Financial Economics 123, 337-356.

Kruger, P., Landier, A., and Thesmar, D., 2012. Categorization bias in the stock market. Unpublished working paper. University of Geneva, Toulouse School of Economics, and HEC.

Kosowski, Robert, Timmermann, Allan, Wermers, Russ and White, Hal, 2006. Can mutual fund “stars” really pick stocks? New evidence from a bootstrap analysis. The Journal of Finance, 61(6), pp.2551-2595.

Ljungqvist, Alexander, Malloy, Christopher and Marston, F., 2009. Rewriting history. The Journal of Finance, 64(4), pp.1935-1960.

Reuter, Jonathan, and Eric Zitzewitz, 2015. How much does size erode mutual fund performance? A regression discontinuity approach, Working paper, Boston College.

Rosenberg, Barr, and Michel Houglet, 1974. Error rates in CRSP and Compustat data bases and their implications, Journal of Finance 29, 1303–1310.

Shumway, Tyler, 1997. The delisting bias in CRSP data, Journal of Finance 52, 327–340.

Shumway, Tyler, and Vincent A. Warther, 1999. The delisting bias in CRSP’s NASDAQ data and its implications for interpretation of the size effect, Journal of Finance 54, 2361–2379.

Swinkels, Laurens, and Liam Tjong-A-Tjoe, 2007. Can mutual funds time investment styles?, Journal of Asset Management 8, no. 2, 123-132.

Vijh, Anand, 1994. S&P 500 trading strategies and stock betas, Review of Financial Studies 7, 215–251.

Wermers, Russ, 2012. Matter of style: The causes and consequences of style drift in institutional portfolios. Available at SSRN 2024259.

Appendix A: Variable Definitions

Variable Name	Definition	Data Source
Reported AAA %	% of holdings in AAA assets as reported by a fund	Morningstar Direct
Calculated AAA %	% of holdings in AAA assets as calculated by Morningstar	Morningstar Direct
AAA %	% of holdings in AAA assets as calculated by us from the portfolio	Constructed
…	…	…
Official Credit Group	The average credit risk of a portfolio assigned by MS	Morningstar Direct
Surveyed Credit Risk Group	The average credit risk of a portfolio calculated using the formula provided by MS and the survey-reported %	Constructed
MS Calculated Credit Risk Group	The average credit risk of a portfolio calculated using the formula provided by MS and the MS calculated %	Constructed
Holdings Calculated Credit Risk Group	The average credit risk of a portfolio calculated using the formula provided by MS and holdings	Constructed
Official Fund Style	A fixed fund could be categorized as any of the following: “High Limited”, “Medium Limited”, “Low Limited”, “High Moderate”, “Medium Moderate”, “Low Moderate”, “High Extensive”, “Medium Extensive”, and “Low Extensive”	Morningstar Direct
Counterfactual Fund Style	We re-evaluate a fund as either high, medium, or low credit quality using their holdings. This counterfactual fund style is the fund style as indicated by the re-evaluated credit quality.	Constructed
Misclassified Dummy	Dummy variable that indicates whether a fund is misclassified in their fund credit quality dimension. It is 1 if the official credit style (High or Medium) is higher than the counter factual credit quality as indicated by holdings, and 0 otherwise.	Constructed
Reported Duration	The reported effective duration of a portfolio	Morningstar Direct
Reported Yield	The reported yield to maturity of a portfolio (in % points)	Morningstar Direct
Calculated Yield	Morningstar calculated average yield to maturity of a portfolio (in % points)	Morningstar Direct
12-Month Yield	The total coupon and dividend payment from the past 12 months (in % points)	Morningstar Direct

Fund Return	The fund return is the value weighted average of the share class returns. Share class returns come from Morningstar Direct	Constructed
Morningstar Rating 3-yr	The fund level Morningstar Rating is the value weighted average of share level Morningstar Ratings	Constructed
Morningstar Rating General	The fund level Morningstar Rating is the value weighted average of share level Morningstar Ratings	Constructed
Average Expense	Average expense at the fund level is calculated by taking the value weighted average of the share-class level expense ratios	Constructed
Monthly Flow	Monthly fund level investor flows	Morningstar Direct
Flow	Quarterly fund level investor flow is the quarterly sum of monthly flow	Constructed

Appendix B: Credit Risk of a Fund Portfolio

Morningstar defines a bond portfolio's average credit risk using a weighted average score using the credit rating of the underlying assets. Prior to August 2010, a bond asset's score is defined using the following table:

Bond Quality	AAA	AA	A	BBB	BB	B	Below B	Not Rated	Not Rated Muni
Score	2	3	4	5	6	7	8	7	6

The portfolio's average position size weighted score then defines its credit quality using the following breakpoints.

Portfolio Average Score	0 to 2.5	2.5 to 3.5	3.5 to 4.5	4.5 to 5.5	5.5 to 6.5	6.5 to 7.5	>7.5
Quality Rating	AAA	AA	A	BBB	BB	B	Below B
Fund Style Quality	High	High	Medium	Medium	Low	Low	Low

After August 2010, the scores are based on a relative default rate:

Bond Quality	AAA	AA	A	BBB	BB	B	Below B	Not Rated	Not Rated Muni
Score	0	0.56	2.22	5.00	17.78	49.44	100.00	49.44	17.78

The respective breakpoints for post August 2010 are then:

Portfolio Average Score	0 to 0.13889	0.13889 to 1.25000	1.25000 to 3.47223	3.47223 to 9.02778	9.02778 to 31.25000	31.25000 to 72.36112	≥ 72.36112
Quality Rating	AAA	AA	A	BBB	BB	B	Below B
Fund Style Quality	High	High	Medium	Medium	Low	Low	Low

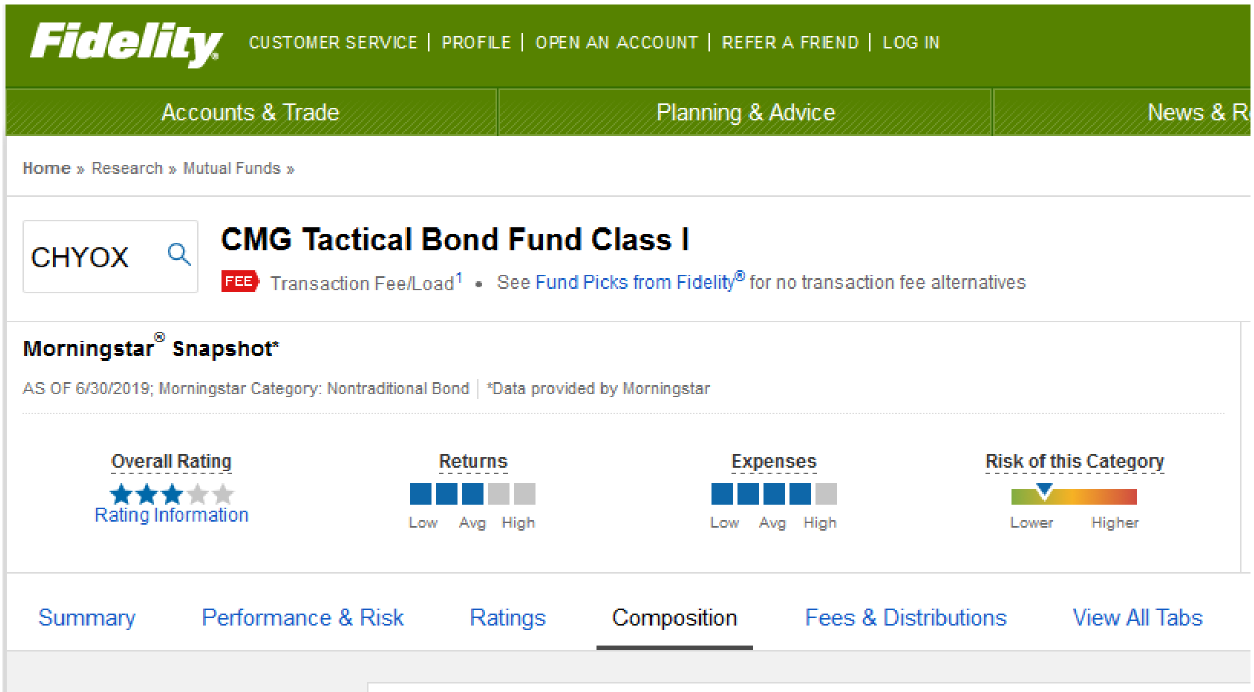

Figure 1.

This figure contains a screenshot of the Fidelity webpage, which uses data provided by Morningstar.

SURVEY AS OF DATE	UNIQUE IDENTIFIER	FUND NAME	CREDIT RATINGS RISK							
			AAA	AA	A	BBB	BB	B	BELOW B	NOT RATED
(mm/dd/yyyy) or (yyyy-mm-dd) Must be a month-end date. All other dates will be rounded (rounded forward where dd =>16). Do not include formulas.	**Char(75)** This MUST be the same identifier used for submitting portfolio holdings. Please contact us if you do not have this information.	**Char(75)** Must be kept consistent per fund, per fund file. Can vary from delivery to delivery due to name changes.	**Numeric(5.2)** % of assets for this rating assigned by a Nationally Recognized Statistical Rating Organizations.	**Numeric(5.2)** % of assets for this rating assigned by a Nationally Recognized Statistical Rating Organizations.	**Numeric(5.2)** % of assets for this rating assigned by a Nationally Recognized Statistical Rating Organizations.	**Numeric(5.2)** % of assets for this rating assigned by a Nationally Recognized Statistical Rating Organizations.	**Numeric(5.2)** % of assets for this rating assigned by a Nationally Recognized Statistical Rating Organizations.	**Numeric(5.2)** % of assets for this rating assigned by a Nationally Recognized Statistical Rating Organizations.	**Numeric(5.2)** % of assets for this rating assigned by a Nationally Recognized Statistical Rating Organizations.	**Numeric(5.2)** % of assets that do not have a rating.

Figure 2.

This figure contains a portion of the fixed income template sent by Morningstar to survey mutual funds.

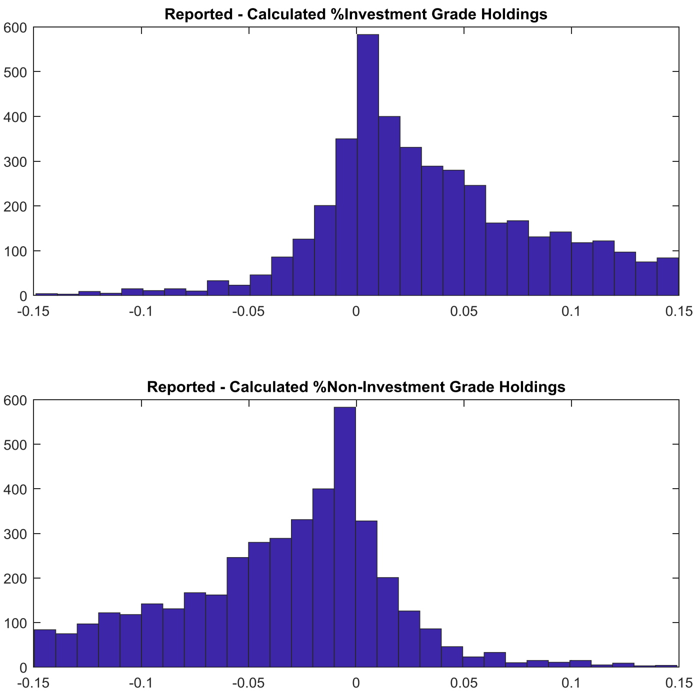

Figure 3.

This graph plots the histograms of fund reported % holdings minus the calculated % holdings in the various bond credit rating categories. The sample period begins in Q1 2017, when Morningstar began calculating % holdings of assets in each credit risk category per each fixed income fund, and ends in Q2 2019. Observations where fund reported % is exactly the same as the calculated % holdings are removed to aid readability.

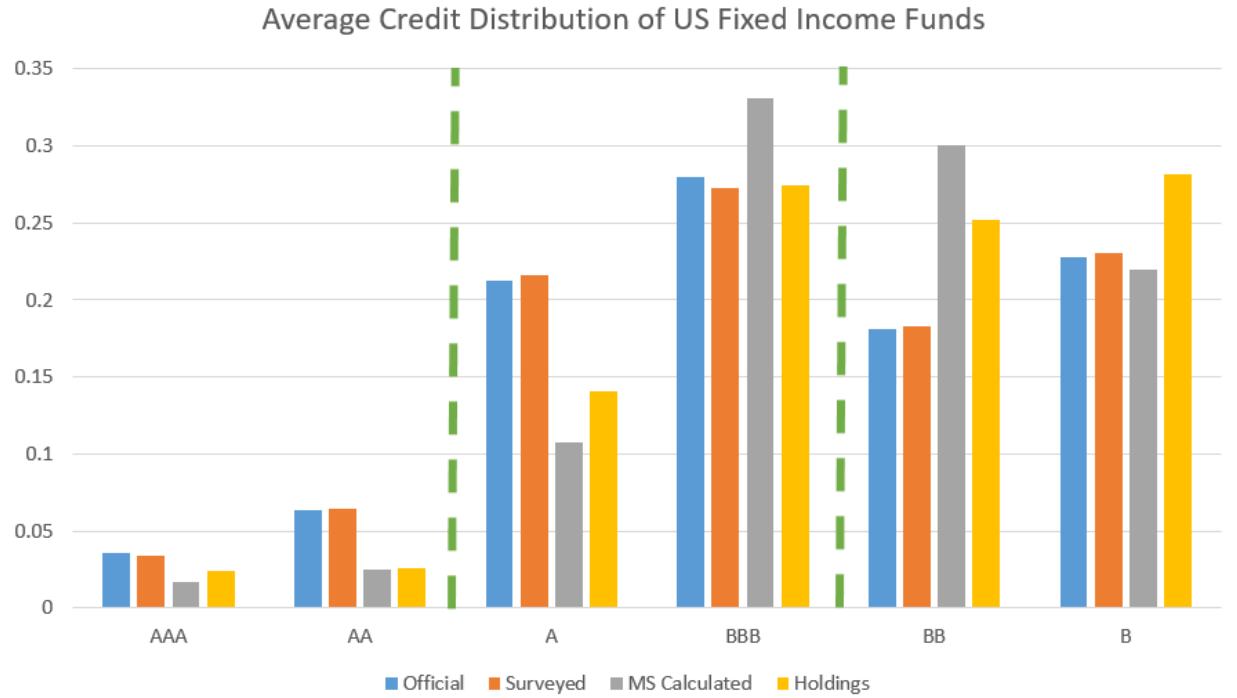

Figure 4.

This figure plots the credit risk distribution of fund-quarter observations between Q1 2017 and Q2 2019. The blue is the distribution of the official average Credit Risk category that Morningstar assigns to US Fixed Income funds. According to MS's methodology, this official credit risk category is calculated using fund survey reported % holdings of assets in the various credit risk categories. In red, we replicate the official credit risk category using the fund survey-reported % holdings. The grey is the counter-factual credit risk category that would result if we had used MS calculated % holdings. The yellow is the counter-factual credit risk category calculated directly from quarterly holdings. The dashed lines represent breaks in the fixed income fund style-box. AAA and AA credit quality funds are high credit quality; A and BBB credit quality funds are medium credit quality; and BB and B are low credit quality as deemed by Morningstar.

Table 1.
Description of Data

We obtain credit ratings from 3 sources. Dodd-Frank requires all credit rating agencies to release their rating data through XBRL filings. Capital IQ subscription contains the S&P rating history. Mergent FISD contains corporates, supranational, and agency/treasuries debts. Portfolio history is directly from Morningstar's collection of filings and surveys for each fund. The surveyed holdings % on individual fixed income funds comes from the Morningstar Direct database from Q1 2003 to Q2 2019.

Panel A. Sources of Credit Ratings:

Dates	Source	Coverage Description
Jun 2012 to Dec 2018	XBRL Filing	All NRSROs Rated Bonds
Jan 2003 to Dec 2018	Capital IQ	S&P Rating History
Jan 2003 to Dec 2018	Mergent FISD	S&P, Moody's, and Fitch Ratings for Corporates and Treasuries

Panel B. Actual Holdings of US Fixed Income Funds from Q1 2003 to Q4 2018

	10th P	Median	90th P	Mean	Std.	N
AAA	0.00%	32.62%	79.05%	35.06%	32.38%	35,926
AA	0.00%	1.40%	7.16%	2.70%	4.06%	35,926
A	0.00%	6.62%	24.52%	9.50%	11.17%	35,926
BBB	0.00%	8.57%	31.72%	13.01%	14.39%	35,926
BB	0.00%	2.93%	23.33%	7.41%	10.17%	35,926
B	0.00%	1.28%	39.63%	9.36%	15.97%	35,926
Below B	0.00%	0.44%	15.55%	4.14%	7.61%	35,926
Unrated	0.00%	4.50%	36.83%	13.20%	23.45%	35,926

Panel C. Surveyed Holdings of US Fixed Income Funds from Q1 2003 to Q2 2019

	10th P	Median	90th P	Mean	Std.	N
AAA	0.00%	42.94%	83.55%	41.00%	31.38%	34,420
AA	0.00%	3.61%	13.00%	5.68%	8.27%	34,420
A	0.00%	9.01%	24.86%	10.71%	10.66%	34,420
BBB	0.49%	11.30%	32.00%	14.43%	13.86%	34,420
BB	0.00%	4.00%	33.11%	10.46%	13.91%	34,420
B	0.00%	1.70%	45.56%	11.94%	18.47%	34,420
Below B	0.00%	0.30%	12.44%	3.54%	6.61%	34,420
Unrated	0.00%	0.40%	6.61%	2.25%	5.29%	34,420

Table 2.
Time Series

In this table, we report the time series of Fund Quarter observations in each Morningstar Credit Quality Category. The last column is the number of funds that are misclassified into the high or med credit quality category. MS changed the way it calculated average credit risk in August 2010. Prior to August 2010, the average credit risk is a simple weighted average of the underlying linear bond scores, in which a AAA bond has a score of 2, AA has a score of 3, and so on. After August 2010, the credit risk variable attempts to describe a fund in terms of the returns and risks of a portfolio of rated bonds, and nonlinear scores are assigned to each category. We record the weighing scheme used after August 2010 in Appendix B.

Year	High Credit Quality	Med Credit Quality	Low Credit Quality	Misclassified
2003	251	412	321	6
2004	263	396	339	6
2005	255	364	286	6
2006	315	414	335	8
2007	325	521	427	16
2008	363	615	470	26
2009	250	704	550	17
2010	213	714	591	188
2011	192	771	673	375
2012	197	862	727	300
2013	193	898	851	359
2014	181	929	919	375
2015	186	1,065	1,046	345
2016	212	1,198	1,052	391
2017	230	1,218	1,025	432
2018	226	1,119	980	431

Table 3.
Yields and Misclassification

In this table, we regress various yield metrics on misclassified dummy and control variables. Misclassified dummy is 1 if the official credit quality (High or Medium) is higher than the counter factual credit quality, and 0 otherwise. Funds voluntarily report their portfolio yields (1) to Morningstar. Morningstar began calculating the holding yields (2) in 2017. The 12-month total interest, coupon, and dividend payments constitute the 12-month yield (3). The sample period is Q1 2003 to Q4 2018. t-statistics are double-clustered by time and fund.

	(1) Reported Yield_t	(2) Calculated Yield_t	(3) 12-Month Yield_{t+11}
Misclassified Dummy_{t-1}	0.431***	0.271***	0.272***
	(9.161)	(4.360)	(5.180)
Reported Duration_{t-1}	0.155***	0.0710***	0.177***
	(6.446)	(3.533)	(4.239)
Time x Official Fund Style FE	Yes	Yes	Yes
Observations	9,275	1,674	16,239
Adjusted R-squared	0.665	0.616	0.645

Table 4.
Counterfactuals and Misclassification

In this table, we regress fund returns on misclassified dummy and control variables. Misclassified dummy is 1 if the official credit quality (High or Medium) is higher than the counter factual credit quality, and 0 otherwise. The sample period is Q1 2003 to Q4 2018. t-statistics are clustered quarterly.

	(1) Fund $Return_t$	(2) Fund $Return_t$
Misclassified $Dummy_{t-1}$	0.103**	-0.111
	(2.644)	(-1.641)
Reported $Duration_{t-1}$	0.106	0.109
	(1.549)	(1.562)
Average Expense $_{t-1}$	-0.106	-0.115*
	(-1.602)	(-1.987)
Time x Official Fund Style FE	Yes	No
Time x Correct Fund Style FE	No	Yes
Observations	16,848	16,839
Adjusted R-squared	0.701	0.715

Table 5.
Morningstar Star Ratings and Misclassification

In this table, we regress Morningstar ratings on the misclassified dummy and controls. Since the ratings and expenses are reported at the share class level, the fund level Morningstar Ratings and the Average Expense ratio are calculated as the value weighted average of their respective share-class level values. The sample period is Q1 2003 to Q4 2018. t-statistics are double-clustered by time and fund.

	(1) Morningstar Rating 3 Yr_t	(2) Morningstar Rating 3 Yr_t	(3) Morningstar Rating $Overall_t$	(4) Morningstar Rating $Overall_t$
Misclassified $Dummy_{t-1}$	0.402***	0.217***	0.340***	0.184***
	(6.839)	(4.830)	(5.745)	(3.662)
Reported $Duration_{t-1}$	0.0392**	-0.103***	0.0429*	-0.0773***
	(2.101)	(-4.521)	(1.916)	(-3.900)
Average $Expenses_{t-1}$	-1.049***	-0.939***	-0.997***	-0.904***
	(-9.126)	(-10.15)	(-7.670)	(-8.354)
3 Year Returns $_{t-1}$		11.56***		9.745***
		(14.77)		(11.81)
Time x Official Fund Style FE	Yes	Yes	Yes	Yes
Observations	15,689	15,689	15,689	15,689
Adjusted R-squared	0.141	0.432	0.120	0.346

Table 6.
Expense Ratios and Misclassification

In this table, we analyze whether misclassified funds are more expensive than usual. We regress average expense ratio on misclassified dummy and control variables. The average expense ratio is calculated at the fund level as the value weighted average of their respective share-class level values. The sample period is Q1 2003 to Q4 2018. t-statistics are double-clustered by time and fund.

	(1) Average $Expense_t$	(2) Average $Expense_t$	(3) Average $Expense_t$
Misclassified $Dummy_{t-1}$	0.0961***	0.0988***	0.0998***
	(5.706)	(5.703)	(5.766)
Reported $Duration_{t-1}$		0.00718	0.00768
		(0.734)	(0.779)
Fund $Return_{t-1}$			-0.523
			(-1.411)
Time x Official Fund Style FE	Yes	Yes	Yes
Observations	17,194	16,647	16,615
Adjusted R-squared	0.097	0.095	0.096

Table 7.
Fund Flows and Misclassification

In this table, we regress the direction of investor flows into share classes on their concurrent Morningstar Rating, as instrumented on lagged fund misclassifications. There are two specifications. The first column controls for time interacted with the official Morningstar Fund-Style as fixed effects, while the second column, in addition, also controls for fund fixed effects (i.e. using the within fund variation). The sample period is Q1 2003 to Q4 2018 for the first column, and Q4 2010 to Q4 2018 for the second column. t-statistics are clustered quarterly.

	(1) $Flow_t$>0	(2) $Flow_t$>0
Misclassified Stars	0.0637***	0.138**
	(3.501)	(2.447)
Reported $Duration_{t-1}$	0.00714**	0.0106
	(2.631)	(1.495)
Average $Expenses_{t-1}$	-0.160***	-0.117
	(-8.702)	(-1.681)
Time x Official Fund Style FE	Yes	Yes
Fund Fixed Effect	No	Yes
Observations	52,503	34,510
Adjusted R-squared	0.122	0.176

Table 8.
Characteristics of Misclassified Funds

In this table, we regress indicators of when a share class start and stop misclassifying their holdings on past returns and other characteristics. The sample period covers Q4 2010 to Q4 2018, when Morningstar used its most recent classification scheme. In column 1, *Starts* is an indicator representing when a previously correctly classified fund starts misclassifying. In column 2, *Stops* is an indicator for when a previously misclassified fund starts correctly classifying. In column 3, we regress *Starts* minus *Stops*. t-statistics are clustered quarterly.

	(1) Starts	(2) Stops	(3) Starts-Stops
3 Year Return	-0.140**	0.309***	-0.449***
	(-2.266)	(4.306)	(-4.283)
Time x Official Fund Style FE	Yes	Yes	Yes
Observations	40,401	40,401	40,401
Adjusted R-squared	0.046	0.063	0.059

Table 9.
Fund Families with the Most Misclassified Funds

This table lists the top thirty fund families, ranked by the percentage of active US bond funds in their family structure that are misclassified for the most recent two years of our sample period (2017-2018).

Firm Name	High	Medium	% Misclassified
UBS Asset Management	0	4	100%
Saratoga	4	0	100%
Waddell & Reed	0	5	100%
Cornerstone	0	8	100%
Diamond Hill Funds	0	1	100%
DoubleLine	0	7	100%
Hartford Mutual Funds	0	9	100%
Yorktown Funds	0	7	100%
Sit	7	0	100%
KP Funds	0	5	100%
Macquarie Investment Management	0	5	100%
Angel Oak	0	1	100%
Allianz Funds	1	1	100%
Muzinich	0	1	100%
LEADER	6	5	91%
Loomis Sayles Funds	1	8	89%
New Covenant	0	7	86%
TCW	11	2	85%
American Century Investments	6	21	78%
Semper	0	4	75%
Putnam	0	6	67%
MassMutual	5	15	63%
Thompson IM Funds Inc	0	8	63%
Capital Group	8	0	63%
Lord Abbett	6	45	61%
John Hancock	9	28	56%
Pioneer Investments	0	16	50%
American Funds	32	27	49%
Voya	9	32	49%
AllianceBernstein	15	37	49%