The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. , NO. 0 • APRIL 2021

Don't Take Their Word for It: The Misclassification of Bond Mutual Funds

HUAIZHI CHEN, LAUREN COHEN, and UMIT G. GURUN*

ABSTRACT

We provide evidence that bond fund managers misclassify their holdings, and that these misclassifications have a real and significant impact on investor capital flows. The problem is widespread, resulting in up to 31.4% of funds being misclassified with safer profiles, compared to their true, publicly reported holdings. "Misclassified funds"—those that hold risky bonds but claim to hold safer bonds—appear to on-average outperform lower risk funds in their peer groups. Within category groups, misclassified funds receive more Morningstar stars and higher investor flows. However, when we correctly classify them based on actual risk, these funds are mediocre performers.

INFORMATION ACQUISITION IS COSTLY. HOWEVER, the exact cost of collecting any piece of information depends on factors, such as timing, location, a person's private information set, as well as idiosyncratic characteristics and complexities of the information signal and the underlying asset itself. External agents—both public and private—have emerged to fill this role and reduce the cost of

*Huaizhi Chen is with University of Notre Dame. Lauren Cohen is with Harvard Business School. Umit G. Gurun is with the University of Texas at Dallas. We would like to thank Stefan Nagel (the Editor) and two anonymous referees; Robert Battalio; Nick Bollen; Robert Burn; Geoffrey Booth; John Campbell; Bruce Carlin; Tom Chang; Christine Cuny; Alex Dontoh; Mark Egan; Michele Gambera; Defne Gurun; Ilan Guttman; Yael Hochberg; Samuel Hartzmark; Alan Isenberg; Robert Jackson; Christian Julliard; Mark Kamstra; Oguzhan Karakas; Craig Lewis; Dong Lou; Tim Loughran; Chris Malloy; Bill McDonald; Rabih Moussawi; Bugra Ozel; Jeff Pontiff; Melissa Prado; Joshua Ronen; Nick Roussanov; Stephen Ryan; David Solomon; Tarik Umar; Russell Wermers; Ingrid Werner; Bob Whaley; Paul Wildermuth; Paul Zarowin; and seminar participants at Drexel University, Michigan State University, London School of Economics, New York University, University of Notre Dame, Rice University, the United States Securities and Exchange Commission, Vanderbilt University, 2020 Conference on the Experimental and Behavioral Aspects of Financial Markets, 2020 Consortium for Asset Management Conference, 2020 Review of Asset Pricing Studies Winter Conference, the 2020 Future of Financial Information conference at Stockholm Business School, 2020 European Finance Association Meeting, and 2020 Northern Finance Association Meeting for helpful comments and suggestions. We also thank James Ng for providing valuable research assistance. We are grateful for funding from the National Science Foundation, SciSIP 1535813. Additionally, Umit Gurun gratefully acknowledges support from the Koc University Distinguished Research Fellowship. We have read *The Journal of Finance* disclosure policy and have no conflicts of interest to disclose.

Department of Finance, 238 Mendoza College of Business, University of Notre Dame, Notre Dame, IN 46556 (574) 631-3385; e-mail: hchen11@nd.edu.

DOI: 10.1111/jofi.13023

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 exposed to both equity and fixed-income assets. With over 16 trillion dollars invested, the U.S. mutual fund market, for instance, counts over 5,000 delegated funds and growing. Although the Securities Exchange Commission (SEC) has mandated disclosure of many aspects of mutual fund pricing and attributes, different asset classes are better (and worse) served by the current disclosure requirement level. Investors have thus turned to private information intermediaries to help fill existing gaps.

In this paper, we show that for one of the largest markets in the world, namely, U.S. fixed-income debt securities, this has led to large information gaps that have been filled by strategic-response information provision by funds. In particular, we show that the reliance on (and by) 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 actually the case. Moreover, the misreporting has a real impact on investor behavior and mutual fund success.

Specifically, we focus on the fixed-income mutual fund market. The fixed-income market is similarly sized to equites (e.g., 40 trillion dollars compared with 30 trillion dollars in equity assets worldwide). However, bonds are fundamentally different as an asset cash-flow claim and have different attributes in delegated portfolios. In particular, while equity funds hold predominantly the same security type (e.g., the common stock of IBM, Apple, Tesla, etc.), fixed-income fund issues differ in characteristics, such as yield, duration, and covenants, 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. To illustrate, in Figure 1 we include an excerpt from the AZL Enhanced Bond Index Fund's N-Q Schedule of Investments from September 30, 2018.[1] As can be seen, the fund held over 700 issues, including seven different bonds of McDonald's Corp each with differing yields, durations, and callable features. Thus, while the SEC mandates equivalent disclosure of portfolio constituents for equity and bond mutual funds, the data are more complex in terms of both processing and aggregating to fund-level measures for fixed income. This has led information intermediaries to bridge this gap, providing a level of aggregation and summary on the general riskiness, duration, and so on, of fixed-income funds.

In this paper, we focus on the largest of such intermediaries that provides data on categorization and riskiness at the fund level—Morningstar, Inc. In particular, we compare fund profiles provided to investors by the intermediary (Morningstar) against the funds' *actual* portfolio holdings. We find significant misclassification of fund riskiness across the universe of all bond funds, with

[1] The full filing, including all 11 pages of holdings, is available on the SEC website: https://www.sec.gov/Archives/edgar/data/1091439/000119312518338086/d615188dnq.htm .

Principal Amount		Fair Value
Corporate Bonds, continued		
Health Care Providers & Services, continued		
$ 1,320,000	UnitedHealth Group, Inc., 4.75%, 7/15/45	$ 1,412,659
220,000	WellPoint, Inc., 3.50%, 8/15/24, Callable 5/15/24 @ 100	215,543
		13,229,679
Hotels, Restaurants & Leisure (0.1%):		
10,000	McDonald's Corp., 2.75%, 12/9/20, Callable 11/9/20 @ 100	9,927
50,000	McDonald's Corp., 4.70%, 12/9/35, Callable 6/9/35 @ 100	51,974
310,000	McDonald's Corp., 3.70%, 2/15/42, MTN	271,124
120,000	McDonald's Corp., 4.60%, 5/26/45, Callable 11/26/44 @ 100	120,542
780,000	McDonald's Corp., 4.88%, 12/9/45, Callable 6/9/45 @ 100, MTN	815,713
800,000	McDonald's Corp., 4.45%, 3/1/47, Callable 9/1/46 @ 100, MTN^	785,569
160,000	McDonald's Corp., 4.45%, 9/1/48, Callable 3/1/48 @ 100, MTN^	156,882
315,000	Starbucks Corp., 3.80%, 8/15/25, Callable 6/15/25 @ 100	312,019
215,000	Starbucks Corp., 3.75%, 12/1/47, Callable 6/1/47 @ 100	187,160
120,000	Starbucks Corp., 4.50%, 11/15/48, Callable 5/15/48 @ 100	118,018
		2,829,828
Household Products (0.0%):		
740,000	Clorox Co. (The), 3.90%, 5/15/28, Callable 2/15/28 @ 100^	738,344
Industrial Conglomerates (0.3%):		
2,565,000	General Electric Capital Corp., Series A, 5.55%, 5/4/20, MTN	2,657,323
2,120,000	General Electric Co., Series D, 5.00%(US0003M+333bps), 12/31/49, Callable 1/21/21 @ 100^	2,065,410
164,000	Georgia-Pacific LLC, 5.40%, 11/1/20(a)	170,635
505,000	Georgia-Pacific LLC, 3.73%, 7/15/23, Callable 4/15/23 @ 100(a)	507,482
309,000	Georgia-Pacific LLC, 3.60%, 3/1/25, Callable 12/1/24 @ 100(a)	305,973
		5,706,823

Figure 1. Sample bond fund holding data. This figure contains an excerpt from the AZL Enhanced Bond Index Fund's September 30, 2018 N-Q schedule of investments held (Source: https:www.sec.gov/Archives/edgar/data/1091439/000119312518338086/d615188dnq.htm). (Color figure can be viewed at wileyonlinelibrary.com)

up to 31.4% of all funds misclassified in recent years. Moreover, this misclassification is pervasive across funds reported as overly safe by Morningstar.

How do these misclassifications occur? Morningstar "rates" each fixed-income mutual fund into a style box based their assessment of credit quality and interest rate sensitivity. For instance, a bond portfolio may be designated as a high-credit-quality fund with low interest rate sensitivity. Our analysis concerns the risk levels reflected in the style box. In addition to the style box, Morningstar places each fund into categories such as "Multisector Bond" or "Intermediate Core Bond." Within each of these fund categories, Morningstar uses a fund's realized returns and volatility to rank funds and assigns them an aggregate rating in the form of "Morningstar Stars."[2] Star ratings have been shown to have a strong and significant effect on both retail and institutional investor flows (Nanda, Wang, and Zheng (2004), Del Guercio and Tkac (2008), Evans and Sun (2018), Reuter and Zitzewitz (2015), Ben-David et al. (2019)).[3]

[2] The ratings methodology and proprietary adjustments and assumptions (e.g., tax burden) that Morningstar employ are described here:https://www.morningstar.com/content/dam/marketing/shared/research/methodology/771945_Morningstar_Rating_for_Funds_Methodology.pdf. To a first-order approximation, the rating is determined by a fund's risk and net return categorization (with high expenses detracting from net returns), within official Morningstar categories (see Section IV of the Internet Appendix, which is available in the online version of this article on *The Journal of Finance* website).

[3] 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)).

In addition, Morningstar provides data releases that are used ubiquitously throughout the industry.

As we document empirically, the central problem is that Morningstar itself has become overly reliant on summary metrics, leading to significant misclassification of risk levels across the fund universe. In particular, Morningstar requires data provision from each fund it rates on the percentage holdings of the bonds by risk rating, that is, on the percentage of the fund's current holdings that are in AAA bonds, AA bonds, BBB bonds, and so on. One might think that Morningstar uses funds' self-reported "Summary Report," data to augment the detailed holdings information it obtained from SEC filings on fund holdings, but Morningstar bases its credit risk summaries solely on the self-reported data. Although this would be no issue if funds provided an accurate view of their holdings to Morningstar, we show that this is not the case—we provide robust evidence that funds on average report significantly safer portfolios than they actually (i.e., 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, as they have large holdings of noninvestment-grade bonds despite being rated AAA portfolios. Due to this misreporting, funds are then misclassified by Morningstar into safer style boxes than they otherwise should be.

For the purposes of our analysis, we define "misclassified funds" as funds that are classified into a different style box than they would be if their *actual* holdings were used to classify them as opposed to the self-reported Summary Report percentages. We show that misclassification is widespread throughout our sample period, with as many as 31.4% of high- and medium-credit-quality funds being misclassified in 2018. Moreover, as mentioned above, misclassifications are overwhelmingly one-sided—very few misstatements push funds toward a higher risk category, while the vast majority of misstatements push funds to a "safer" risk category.

We next examine the characteristics of these "misclassified funds." We first find that misclassified funds have higher average risk, and higher yields on their holdings, than their category peers. This result is not entirely surprising, as again misclassified funds hold riskier bonds than their correctly classified peers in their risk category. Importantly, this translates into misclassified funds earning 3.05 bps ($t = 3.47$), or 16% higher returns per month on average relative to peer funds.

To estimate the portion of misclassified funds' return outperformance that is due to skill versus the portion due to unfair comparison to safer funds, we turn to the *actual* holdings reported in funds' quarterly SEC filings. We use these actual holdings to calculate the correct risk category that the fund should be classified into if it truthfully reported the percentage of holdings in each risk category. When we rerun the same performance regression specification but using proper peer comparisons, we find that misclassified funds no longer exhibit outperformance, and indeed even *underperform* by 0.558 bps per month ($t = 0.65$). It therefore appears that all of the apparent outperformance of misclassified funds comes from being misclassified into a less risky comparison group of funds than they should be.

We also find that misclassified funds receive significantly more Morningstar stars than other funds. This is true even after controlling for Morningstar category and risk classification. In particular, misclassified funds receive an additional 0.38 ($t = 5.97$), or 12.3%, stars. This is surprising given that stars are awarded based on relative risk-adjusted performance within a Morningstar category. When we dig deeper into the relation between star rating and misclassification, we find that funds achieve a higher star rating by holding riskier securities, as the Morningstar methodology does not incorporate an appropriate risk adjustment. Consequently, the additional risk results in higher stars, with the misclassification allowing these funds to take on such risk while appearing to be a "safe" fund.

Armed with higher returns relative to (incorrect) peers and higher Morningstar Ratings, misclassified funds are able to charge significantly higher expenses. In particular, they charge 11.4 bps higher expense ratios than peers ($t = 6.36$).

We next explore the factors that drive this misclassification. Morningstar has posited that it can be explained almost entirely by their classification formula's treatment of nonrated bonds.[4] In the Internet Appendix,[5] however, we show even when we omit all funds that have any nonrated bonds, all of our results above remain economically and statistically significant (in fact, magnitudes are even larger in some cases). Looking more closely at the characteristics and behaviors of the nonrated bonds themselves, and of the misclassified funds that hold them, we find that (i) the yields of nonrated bonds look incredibly similar to those of junk bonds (and very little like those of the higher rated bonds that they are advertised to be by fund managers and that Morningstar takes their word, and (ii) the misclassified funds that hold these nonrated bonds underperform precisely when the junk bond market crashes but experience their greatest fund outperformance when the junk bond market surges (even though they supposedly hold predominantly highly rated, safe securities).

In our next set of analyses, we estimate the extent to which misclassification impacts investor behavior. In particular, we examine whether misclassified funds—even those with higher fees—attract more investor flows, presumably due to the favorable comparison benefits of being misclassified. We find this to be strongly true in the data—misclassified funds have a 12% higher probability of positive flows ($t = 4.95$). The reason is twofold. First, misclassified funds observe a boost in realized returns (on average) given the more aggressive positions in their portfolios. Second, importantly, they get the additional risk

[4] In Section III, we detail our ongoing conversations regarding these large misclassifications. We have been in contact with Morningstar since we first began this project. Included are their proposed explanations of the discrepancies, along with our responses and evidence on their proposed explanations.

[5] The Internet Appendix is available in the online version of this article on *The Journal of Finance* website.

for "free" in the sense that investors believe them to be of low risk, given Morningstar's incorrect risk classification (we show that investors invest significantly less in funds that they perceive to be riskier, conditional on the *same* Morningstar star rating).

In our final set of tests, we explore the characteristics of misclassifying funds. We find that younger managers who are earlier in their careers tend to misclassify more often. Moreover, the more distinct share classes a fund services, along with funds that are the only taxable income fund in their family, are more likely to be misclassifiers. In terms of when fund managers begin to misclassify funds, we find that misclassifications begin when younger fund managers of funds with numerous share classes realize a string of especially negative recent returns. Finally, we find evidence of significant and widespread flow responses among both individual and institutional investors. Although in point terms estimate retail investors (and in particular retirement investors) appear to be more swayed by misclassification, institutional investors also invest significantly more in funds misclassified as overly safe given their actual holdings.

The results we document fit within a number of literature streams. First, our findings on the association between misclassification and performance are 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 their stated investment objective do better than their counterparts, whereas 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 the performance and characteristics of funds that deviate from the stated objectives in the prospectuses. In the equity space, Sensoy (2009) shows that a fraction of size and value/growth benchmark indices disclosed in the prospectuses of U.S. equity mutual funds do not match the fund's actual style.

Second, our paper is related to the growing literature on reaching for yield of investors. Stein (2013) and Rajan (2013) note that an extended period of low interest rates can create incentives for investors to take a greater duration risk, which in turn could induce "fixed-income investors with minimum nominal return needs to migrate to riskier instruments." Along these lines, Becker and Ivashina (2015) study the holdings of insurance companies and show that these firms prefer to hold higher rated bonds because of higher capital requirement constraints, but, conditional on credit rating, their portfolios are systematically biased toward higher yield bonds. Similarly, Choi and Kronlund (2017) show that U.S. corporate bond mutual funds that tilt portfolios toward bonds with yields higher and are able to attract fund flows, especially during periods of low interest rates.[6]

[6] Another group of papers in this literature investigates whether financial intermediaries' institutional frictions affect their response to interest rates. See Drechsler, Savov, and Schnabl (2018) and Acharya and Naqvi (2019), who present models to study the conditions under which banks reach for yield by taking deposits from risk-averse investors. Prior studies consider similar

Our evidence is also related to studies on the implications of data accuracy and completeness. Along these lines, Ljungqvist, Malloy, and Marston (2009) show that I/B/E/S analyst stock recommendations are associated with various changes across vintages (alterations of recommendations, additions and deletions of records, and removal of analyst names) that are nonrandom and that likely affect the profitability of trading signals, for example, profitability of consensus recommendation. 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 cutoff. Recent work further shows that the mutual funds exhibit considerable variation in their month-end valuations of identical corporate bonds (Cici, Gibson, and Merrick (2011)). Similar biases have been documented for the 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 by high illiquidity of fund holdings.

Finally, our study contributes to the literature on style investment. Barberis and Shleifer (2003) argue that investors tend to group assets into a small number of categories, leading to 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 effect 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.

The remainder of the paper proceeds as follows. Section I describes the data, and methodology that Morningstar uses to classify funds into categories. Section II presents our main results on the misreporting of funds and misclassification of these funds by Morningstar based on the faulty reports. Section II also documents return implications as well as benefits for funds in terms of expenses, Morningstar stars, and investor flows, and explores characteristics of misclassified funds. Section III explores the role of nonrated securities and provides evidence on Morningstar's proposed explanations for misclassified funds. Finally, Section IV concludes.

mechanisms for life insurance companies (Ozdagli and Wang (2019)), pension funds holdings (Andanov, Bauer, and Cremers (2017)), and households (Lian, Ma, and Wang (2019)).

I. Data

In this section, we describe the three major databases used in this paper. Specifically, we combine (i) the Morningstar Direct database of mutual funds and their characteristics, (ii) the Morningstar Open-Ended Mutual Fund Holdings database, and (iii) our assembled collection of credit rating histories to document a substantial gap between the reported and true portfolio compositions in fixed-income funds.

A. The Morningstar Direct Database

Our initial sample of fixed-income mutual funds comes from Morningstar Direct. Specifically, we begin with U.S.-domiciled, dollar-denominated, mutual funds that belong to the "U.S. Fixed Income" global category where we filter out the U.S. government, agency, and municipal bond funds using lagged Morningstar subcategories. The full sample comprises 2,029 unique fixed-income mutual funds from Q1 2003 to Q2 2018. We next apply the following filters (i) more than 85% of each portfolio's total holdings are observable, (ii) the long side of each portfolio is no greater than 115% of its total value, (iii) the total net assets of each fund is over $10 million dollars in value, and (iv) each fund has no more than 35% in holdings on which we have no ratings information. These filters yield a sample of 675 unique funds. Morningstar Direct, contains detailed information on fund characteristics that comes from both regulatory open-ended mutual fund filings and direct fund surveys.

A key element of our study is the self-reported asset compositions from mutual fund companies. Figure 2 displays the survey used by Morningstar to collect this information from managers. The date of the survey ("Survey As Of Date") is clearly communicated to the funds to be a month-end. We check this date against the month-ends corresponding to the exact quarter-ends of holding period reporting dates associated with SEC filings. In the first quarter of 2017, Morningstar began calculating percentage asset compositions directly from holdings, but as of March 2020, they continue to use the self-reported asset compositions to place fixed-income funds in risk classification styles. Notably, we also obtain historical returns, share-level investor flow, and fixed-income fund styles from these data. We provide a full list of the variables used in this study in Section I of the Internet Appendix.

B. 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 obtain portfolio weight, long/short profile, and asset type from these data. We focus on positions listed as "Bond" broad-types, and we exclude assets listed as swaps, futures, or options.

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. Morningstar survey. This figure contains a portion of the fixed-income template sent by Morningstar to survey mutual funds in August 2019. (Color figure can be viewed at wileyonlinelibrary.com)

C. Credit Rating Histories

As our analysis centers on the presentation 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 the Dodd-Frank Act, credit rating agencies are required to post their rating histories within a year of each ratings announcement as XBRL releases. These releases enable us to achieve coverage by Standard & Poor's, Moody's, and Fitch for all CUSIP-linked securities after June 2012. In addition to these three main nationally recognized statistical rating organizations (NRSROs), we also have coverage of Ambest, DBRS, Egan-Jones, Kroll, and Morningstar credit rating services, covering all of the designated U.S. domicile NRSROs during our sample period. We obtain credit ratings for pre-June 2012 from Capital IQ and the Mergent Fixed Income Securities Database (FISD). Capital IQ contains credit rating histories from Standard & Poor's for all of our sample history. Mergent FISD provides coverage of credit ratings from Moody's, Standard & Poor's, and Fitch on corporates, supranational, agency, and Treasury bonds. Table I, Panel A, lists these data sources, the rating agencies reported in these sources, and the time span of their respective coverage. Panels B and C tabulate the actual (as calculated using our credit rating histories) and reported percentage holding compositions of fixed-income mutual funds in the various credit rating categories from Q1 2003 to the end of the respective sample.

II. Main Results

A. Diagnostics Analysis

We start our analysis by examining histograms, presented in Figure 3, of fund-reported percentage holdings minus the calculated percentage holdings in various bond credit rating categories between Q1 2017 and Q2 2018. The start of this diagnostic sample is dictated by the time Morningstar began calculating the percentage holdings of assets in each credit risk category per each fixed-income fund. Ideally, if Morningstar and the bond funds in its database maintained the same reporting standards in credit ratings, the fund reported percentage holding would be almost same as the calculated percentage holdings. In this case, the histograms would show a sharp spike around zero (e.g., no discrepancies) and exhibit no significant variation. However, this simple diagnostic shows that there are wide discrepancies in the recorded asset compositions. Most notably, for assets above investment-grade (i.e., 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 assets below investment grade, especially unrated assets, the opposite pattern obtains; that is, the percentage of assets reported by funds is significantly lower than the percentage of assets calculated by Morningstar.

Table I

Description of Data

We obtain credit ratings from three sources. The Dodd-Frank Act that requires all credit rating agencies release their rating data history through XBRL filings with a one-year delay. Capital IQ provides S&P's rating history. Mergent FISD contains corporates, supranational, and agency/Treasury Bonds. Portfolio histories come directly from Morningstar's filings and surveys for each fund. The surveyed holdings percentage on individual fixed-income funds comes from the Morningstar Direct database from Q1 2003 to Q2 2018.

Panel A: Sources of Credit Ratings

Coverage Period	Source	Coverage Description
June 2012 to June 2018	XBRL Filing	All NRSROs Rated Bonds
January 2003 to June 2018	Capital IQ	S&P Rating History
January 2003 to June 2018	Mergent FISD	S&P, Moody's, Fitch Ratings for Corporations and Treasuries

Panel B: Actual Holdings of U.S. Fixed-Income Funds from Q1 2003 to Q2 2018

	10^{th} P	Median	90^{th} P	Mean	*SD*	*N*
AAA	0.00%	40.8%	81.4%	39.0%	31.2%	18,508
AA	0.00%	2.48%	9.15%	3.73%	4.92%	18,508
A	0.00%	7.97%	22.7%	9.58%	9.94%	18,508
BBB	0.326%	12.6%	35.8%	15.9%	15.9%	18,508
BB	0.00%	3.88%	28.2%	9.10%	11.6%	18,508
B	0.00%	1.52%	44.8%	11.4%	18.3%	18,508
Below B	0.00%	0.537%	18.1%	4.71%	8.08%	18,508
Unrated	0.0743%	4.12%	15.7%	6.50%	7.42%	18,508

Panel C: Surveyed Holdings of U.S. Fixed-Income Funds from Q1 2003 to Q2 2018

	10^{th} P	Median	90^{th} P	Mean	*SD*	*N*
AAA	0.00%	41.1%	83.9%	40.1%	31.5%	18,508
AA	0.00%	3.56%	12.8%	5.51%	7.97%	18,508
A	0.00%	9.34%	25.6%	10.9%	10.7%	18,508
BBB	0.50%	12.5%	34.6%	15.7%	15.1%	18,508
BB	0.00%	4.20%	32.0%	10.3%	13.3%	18,508
B	0.00%	1.70%	46.0%	11.8%	18.6%	18,508
Below B	0.00%	0.39%	14.6%	3.99%	7.16%	18,508
Unrated	0.00%	0.32%	5.26%	1.67%	3.61%	18,508

B. Implications of Composition Disagreement: Misclassification

We next examine the major implication of the difference between reported and actual compositions of fund portfolios, namely, the misclassification of these funds. Specifically, in Figure 4, we plot the credit risk distribution of fund-quarter observations between Q1 2017 and the end of Q2 2018. The dashed lines represent breaks in the fixed-income fund style box. AAA and AA funds are classified as high-credit-quality funds, A and BBB funds are classified as medium-credit-quality funds, BB and B funds are classified as low-credit-quality funds by Morningstar.

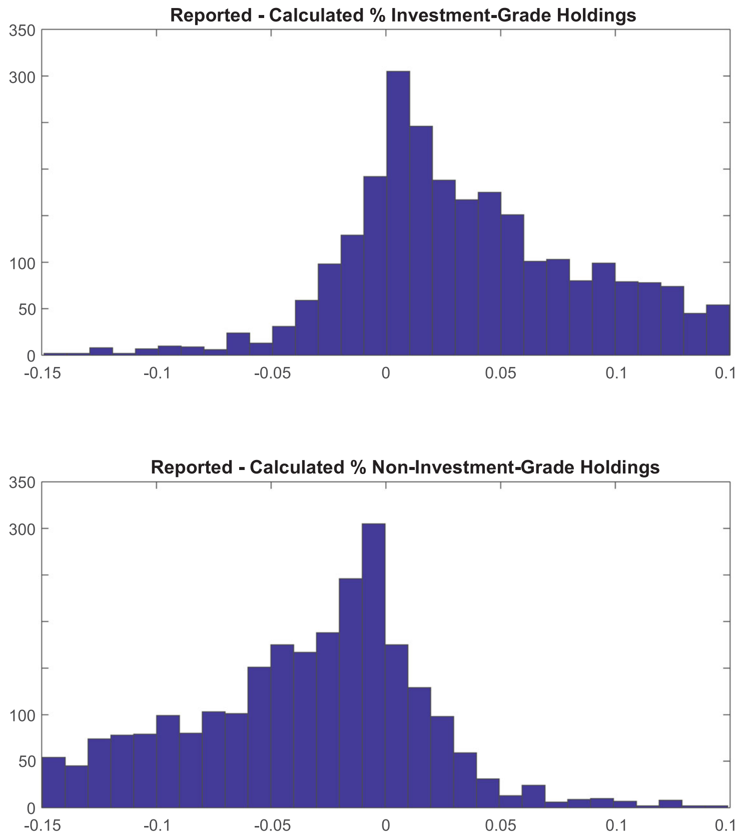

Figure 3. Distribution of difference between reported and calculated holdings. This figure provides histograms of fund reported percentage holdings minus calculated percentage holdings in various bond credit rating categories. The sample period begins in Q1 2017, when Morningstar began calculating percentage holdings of assets in each credit risk category for each fixed-income fund, and ends in Q2 2018. Observations for which fund reported percentage is exactly the same as the calculated percentage holdings are removed to aid readability. (Color figure can be viewed at wileyonlinelibrary.com)

The first (blue) bar depicts the distribution of the Morningstar Assigned Credit Risk Category of the fixed-income funds. In other words, the blue bar is what mutual fund investors observe if they use Morningstar as a data provider. The second (orange) bar depicts the same category distribution, calculated using the fund's self-reported percentage of holdings in the various credit risk categories (from Figure 2). Specifically, using Morningstar's published methodology, this credit risk categorization is calculated as a function of a nonlinear score assigned to each category by Morningstar (see Section II of the Internet Appendix) multiplied by the fund's self-reported percentage of holdings in AAA assets, AA assets, and so on. Finally, the third (gray) bar is calculated using the

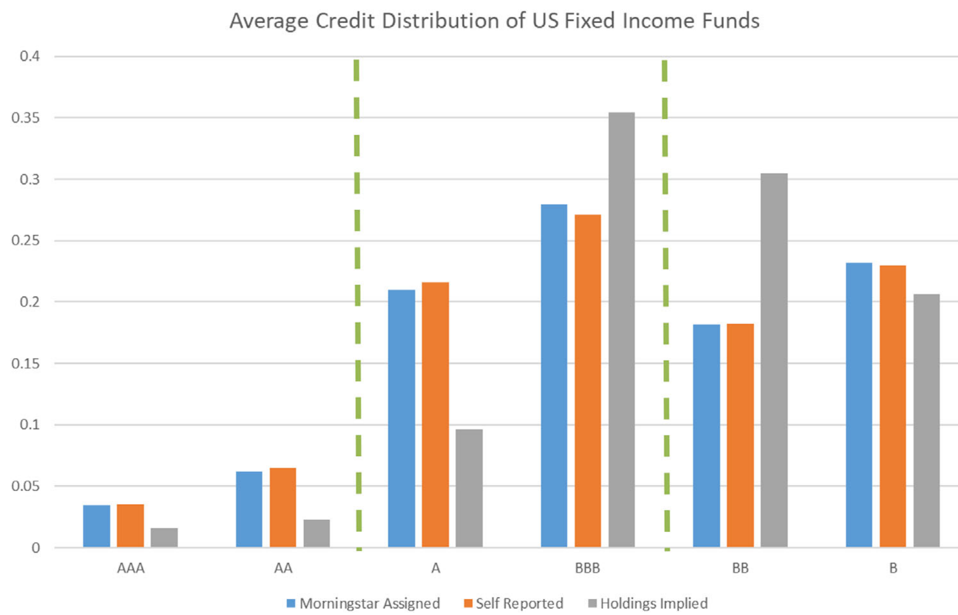

Figure 4. Credit risk distribution of us fixed-income funds. This figure plots the credit risk distribution of fund-quarter observations between Q1 2017 and Q2 2018. Blue bars provide the distribution of the official average credit quality category that Morningstar assigns to U.S. fixed-income funds. According to Morningstar's methodology, this official credit quality category is calculated using fund survey reported percentage holdings of assets in the various credit risk categories. Red bars depict the official credit quality category using fund survey-reported percentage holdings. Gray bars show the counterfactual credit risk category that would result if we had used Morningstar calculated percentage holdings. The dashed lines represent breaks in the fixed-income fund style box. AAA and AA funds are high credit quality, A and BBB funds are medium credit quality, and BB and B funds are low credit quality as rated by Morningstar. (Color figure can be viewed at wileyonlinelibrary.com)

fund's *actual* holdings and their rating (multiplied by the same scores assigned to each rating type as in the orange bar).

If Morningstar relies on the actual holding compositions of the funds, the blue bar should track the gray bar. If, instead, it simply "takes the funds' word for it," simply multiplying the appropriate risk score times the funds' self-reported percentages, the blue bar should track more closely the orange bar. As can be seen in Figure 4, the blue bar tracks the orange bar almost exactly. As a result, many fixed-income mutual funds that would have fallen into a higher credit risk bucket, are classified into safer categories.

Looking at these three distributions more closely indicate that using funds' self-reported credit risk compositions has skewed the fund-level credit categorization in favor of lower perceived credit risk. For example, almost half of funds A-rated funds would not be in this category if the fund-level credit rating were assigned based on the actual holdings-implied, rather than self-reported, compositions. Likewise, half of the AAA-rated funds should have received a riskier categorization according to the actual calculated holdings. Taken together, the evidence in this subsection suggests that when a fund reports high levels of investment grade assets, it will be classified as an investment-grade fund regardless of its actual holdings.

C. Misclassification in Detail

In this section, we explain how systematic patterns of over/underreporting vary with assumptions regarding (i) how we select our sample and (ii) how we match credit ratings to securities. We discuss the baseline analysis in detail below. In Section III of the Internet Appendix, we outline the degree of the misclassification in each scenario.

We combine the credit rating history of each fixed-income asset in every bond fund portfolio in order to calculate the actual percentage of assets held in each credit risk category. In other words, we match the bond positions of mutual fund portfolios to their respective ratings to calculate their average credit risk classification. These are positions that are listed as "Bond" broad-types in the Morningstar Holdings database. In our baseline analysis, we exclude assets that are listed as swaps, futures, or options, that is we do not classify these assets as a specific rated type or as unrated. When multiple credit rating agencies rate a single asset, we aggregate using the Bloomberg/Barclays method as prescribed by Morningstar's own methodology document. According to this method, if a security is rated by only one agency, then that rating used as the composite. If a security is rated by two agencies, then the more conservative rating is used. If all three rating agencies are present, then the median rating is assigned. Additionally, government-backed securities, such as Agency Pass-thru's, Agency collateralized mortgage obligations, and Agency adjustable-rate, are automatically classified as AAA-rated assets. We also search for Treasuries and potentially overlooked government-backed securities by searching keywords such as "FNMA," "U.S. Treasuries," "REFCORP," assigning them a AAA-rating. We then use these holdings-calculated compositions to determine the implied average credit risk. According to this method, roughly 24.1% of bond funds receive counterfactual credit risk categorizations that are riskier than their official credit risk categorizations in the post-2016 sample. In Section III of the Internet Appendix, we list the potential assumptions one can make and its corresponding misclassified bond ratios.

In Table II, we tabulate the time series of fund-quarter observations in each Morningstar Credit Quality Category using the longest time series we can obtain (2003 to 2018). Morningstar's fund level credit ratings are calculated by weighing the funds' self-reported percentage of assets under management (AUM) in the different credit rating categories using static scores and then assigning credit risk ratings using score cutoffs. Morningstar changed its scoring weights and cutoffs for classifying funds in Q3 2010. Prior to the change, assets were weighed by assigning categorical scores that corresponded linearly to their credit ratings—AAA bonds weighed 2 points, AA 3, A 4, and so on. The final portfolio designations were then determined over specific ranges

Table II

Time Series of Misclassification

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 medium credit quality categories. Morningstar changed the way it calculates average credit quality in August 2010. Prior to August 2010, average credit quality was a simple weighted average of the underlying linear bond scores, where a AAA bond had a score of 2, AA had a score of 3, and so on. After August 2010, the credit risk variable attempts to describe a fund in terms of the return and risk of a portfolio of rated bonds, and nonlinear scores are assigned to each category. The sample period is Q1 2003 to Q2 2018. We record the weighting scheme used after August 2010 in Section III of the Internet Appendix.

Year	High Credit Quality	Med Credit Quality	Low Credit Quality	# Misclassified
2003	251	412	321	7
2004	262	396	337	4
2005	255	364	282	4
2006	315	414	332	5
2007	322	516	422	7
2008	359	610	468	8
2009	246	698	548	9
2010	209	705	583	147
2011	189	765	658	307
2012	194	857	708	283
2013	191	887	824	297
2014	178	920	891	348
2015	181	1,056	1,022	321
2016	209	1,195	1,024	360
2017	225	1,215	993	370
2018	123	581	484	191

of scores—portfolios scoring less than 2.5 were marked AAA, those between 2.5 and 3.5 marked AA, and so on.

Starting in Q3 2010 (through to the end of the sample period), nonlinear scores that correspond to default probabilities were assigned to each rating category. At the low-risk end, AAA bonds received a weight of zero, with AA bonds a weight of 0.56, while at the higher risk end, BB bonds received a weight of 17.78, B and unrated bonds a weight of 49.44, and B-bonds a weight of 100. The classification cutoffs were then changed to correspond to the new scores of the respective bond classes. This effectively means that any reporting of low-credit quality bond assets would likely move a portfolio toward a higher risk category. In effect, the methodology change made it very difficult for portfolios to have high yield bonds while still maintaining a low credit risk classification.

In Table II, the final column # *Misclassified*, is then the number of observations per year that have riskier counterfactual ratings than their official ratings. These numbers suggest that number of misclassified funds increased dramatically over time but most notably after August 2010, the year Morningstar changed the way it calculated average credit risk. We reproduce the weighting scheme in accordance with Morningstar's published methodology in

Section II of the Internet Appendix. The result of this change in methodology (as seen in the Internet Appendix and described above) was a much higher relative penalty placed on lower rated bonds versus higher rated bonds. This resulted in a much more composition-dependent categorization of fixed-income funds (given the drastic ratings penalty spreads). In our main regression analysis, we focus on the sample of funds that are misclassified from Q3 2010, when Morningstar began its new bond credit risk classification system, to Q2 2018.

D. Fund Performance and Misclassification

A natural follow-up question is whether these misclassified funds are, in fact, different than their risk category peers, given that they hold a larger percentage of lower credit-quality assets than their risk category peers (and lower credit quality assets than their classifications suggest they should hold). We explore both the risk and the return characteristics of these misclassified funds versus their correctly classified peers in this section.

In Table III, we first regress the yield metrics of a fund on our misclassification measure. Specifically, we define *Misclassified* as a dummy variable that takes a value of 1 if the Morningstar credit quality (high or medium) is higher than the counterfactual credit quality calculated using the actual underlying holdings, and 0 otherwise. We use three types of yield metrics. 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 calculated yields only became available in 2017. In the third column, we use the 12-month yield that combines total interest, coupon, and dividend payments. We also include a credit score variable (the reported compositions score that is used to classify fund credit risks), with increasing values signifying greater credit risk, and the bond duration (as reported by the funds) to control for the interest rate risk of the bond portfolio. In addition, we include *Time* $\times$ *Morningstar Category* fixed effects to control for common variation in returns and risk due to category-time specific variation (Section IV of the Internet Appendix lists the official Morningstar categories). In columns (1) to (3), we also include *Time* $\times$ *Morningstar Reported Risk Style* fixed effects in our specification to absorb the mean yield of each fund's corresponding Morningstar fund-calculated risk classification in the 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 the misclassified dummy essentially captures this fund-style-related reporting choice. We cluster standard errors by time and fund to address the time series cross-sectional and individual variation in risk.

From Table III, all three yield columns point to the same empirical regularity, namely, that misclassified funds have significantly higher yields. The annualized reported yield to maturity is 27.7 bps higher (t = 5.49), whereas the calculated yield from the holdings (second column) and the payout yield are 23.7 and 19.0 bps higher, respectively, for misclassified funds over their official peers.

Table III

Yields and Misclassification

In this table, we regress various yield metrics on *Misclassified*, a dummy equal to 1 if the official credit quality (high or medium) is higher than the counterfactual credit quality, and 0 otherwise, and control variables. Funds voluntarily report their portfolio yields (1) and (4) to Morningstar. Morningstar began calculating the holding yields (2) and (5) in 2017. The 12-month total interest, coupon, and dividend payments constitute the 12-month yield (3) and (6). The sample period is Q3 2010 to Q2 2018. t-Statistics are double-clustered by time and fund. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) Reported $Yield_t$	(2) Calculated $Yield_t$	(3) 12-Month $Yield_{t+11}$	(4) Reported $Yield_t$	(5) Calculated $Yield_t$	(6) 12-Month $Yield_{t+11}$
$Misclassified_{t\text{-}1}$	0.277***	0.237***	0.190***	0.0106	0.0130	–0.0735
	(5.494)	(5.372)	(3.344)	(0.157)	(0.273)	(−1.106)
Reported Credit $Score_{t\text{-}1}$	0.112***	0.0569***	0.0551***	0.0727***	0.0486***	0.0552***
	(8.394)	(6.188)	(4.744)	(7.861)	(9.229)	(6.755)
Reported $Duration_{t\text{-}1}$	0.127***	0.0229**	0.107***	0.138***	0.0359**	0.110***
	(4.263)	(3.083)	(3.272)	(4.820)	(3.116)	(3.637)
Time × Morningstar Reported Risk Style FE	Yes	Yes	Yes	No	No	No
Time × Correct Fund Risk Style FE	No	No	No	Yes	Yes	Yes
Time × Morningstar Category FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,402	1,303	7,127	7,957	1,542	8,800
Adjusted R^2	0.673	0.816	0.587	0.736	0.873	0.607

Table IV

Counterfactuals and Misclassification

In this table, we regress monthly fund returns on *Misclassified*, a dummy equal to 1 if the official credit quality (high or medium) is higher than the counterfactual credit quality, and 0 otherwise, and control variables. The sample period is Q3 2010 to Q2 2018. *t*-Statistics are clustered quarterly. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) *Fund Return$_t$*	(2) *Fund Return$_t$*	(3) *Fund Return$_t$*	(4) *Fund Return$_t$*
Misclassified$_{t-1}$	3.579***	3.038***	–2.341**	–0.558
	(2.951)	(3.472)	(−2.003)	(−0.646)
Reported Credit Score$_{t-1}$		0.411**		0.611**
		(2.419)		(2.259)
Reported Duration$_{t-1}$		1.522		1.468
		(1.065)		(1.012)
Average Expense$_{t-1}$		–3.551***		–3.392***
		(−3.393)		(−3.774)
Time × Morningstar Reported Risk Style FE	Yes	Yes	No	No
Time × Correct Fund Risk Style FE	No	No	Yes	Yes
Time × Morningstar Category FE	Yes	Yes	Yes	Yes
Observations	25,318	22,671	31,196	27,941
Adjusted R^2	0.874	0.874	0.841	0.844

In columns (4) to (6), we then explore how these misclassified funds would compare were we to compare them against their *correctly* classified risk peers. In particular, for each fund, we use its underlying holdings to calculate its *Correct Fund Risk Style*—note that for funds that are already correctly classified, this will be the same as in columns (1) to (3), and hence will only change, and correctly reflect the risk of the underlying holdings, for misclassified funds.

Columns (4) to (6) of Table III then conduct the same tests as columns (1) to (3), but replace the *Time* × *Morningstar Reported Risk Style* fixed effects with *Time* × *Correct Fund Risk Style* fixed effects. From columns (4) to (6), the *Misclassified* dummy variable drops in magnitude to near zero and is statistically insignificant. This result implies that when one properly accounts for the *true* risk of the funds' underlying holdings (based on their actual holdings, as opposed to what they self-report to Morningstar and that Morningstar bases their risk classifications), they have identical yields as their true peer funds.

Next, we examine the performance of these misclassified funds versus their correctly risk-classified peer funds. In Table IV, we regress actual fund returns on the *Misclassified* dummy, along with the same controls and fixed effects as in Table III. In the columns (1) and (2), we include *Time* × *Morningstar Reported Risk Style* fixed effects as in the previous table. We find that misclassified funds significantly outperform their risk style and Morningstar fund category peers, controlling for other determinants of returns. In particular, column 2 implies that these funds outperform by 3.04 bps per month (t = 3.42), which represents a 16% higher return than peer funds.

In columns (3) and (4), we then replace this *Morningstar Reported Risk Style* fixed effect with *Time* × *Correct Fund Risk Style* fixed effects. The idea is to estimate the percentage of the documented return outperformance of misclassified funds that comes from skill versus percentage that comes from the unfair comparison to safer funds. From columns (3) and (4), when we compare misclassified funds against their correctly classified peers, we find that they exhibit no outperformance. In fact, in column (4), when compared against their correct risk peers, misclassified funds actually slightly underperform by 0.558 bps per month (t = 0.65), though insignificantly so. In sum, the results in Table IV suggest that misclassified funds appear to outperform, but all of that outperformance comes from being compared against an incorrect (overly safe) set of category peers.

E. Morningstar Stars, Fees Charged, and Flows to Misclassified Funds

In our next set of analyses, we explore a number of other characteristics of misclassified funds. The first characteristic is the number of Morningstar stars received by the funds from Morningstar itself. As referenced above, Morningstar uses their star rating system to reward funds for "true outperformance" in their designated Morningstar category (see Section IV of the Internet Appendix). These Morningstar stars have been shown by a vast literature to have a strong relationship to investor fund flows (e.g., Del Guercio and Tkac (2008), Evans and Sun (2018), Reuter and Zitzewitz (2015), Ben-David et al. (2019)), and by revealed preference are used by many fund companies as an explicit part of their marketing strategy.

We explore this relationship by regressing various Morningstar rating metrics on the *Misclassified* dummy, the reported credit rating score, reported duration, average expense ratio, *Time* × *Morningstar Reported Risk Style* fixed effects, and importantly the *Time* × *Morningstar Category* fixed effects (as this is the peer group against which Morningstar asserts to make its risk and net return comparison). 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. The results are reported in Table V. We find that misclassified funds are associated with significantly higher levels of Morningstar Stars. On average, misclassified funds receive 0.17 (t = 3.77) to 0.38 (t = 5.97) more Morningstar Stars compared to their peer funds. This rating change corresponds to 18% to 41% of a standard deviation in Morningstar stars ratings, or up to a 12.3% increase in the number of stars.

A natural question to arise is why funds that misclassify their risk-style boxes would receive more Morningstar stars, given that stars are assigned based on risk-adjusted return rankings within Morningstar categories. In Section IX of the Internet Appendix, we show that this arises from Morningstar not adequately "penalizing" fixed-income funds for risk taken. Therefore, as

Table V

Morningstar Star Ratings and Misclassification

In this table, we regress Morningstar ratings on the *Misclassified* dummy and controls. Since 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 Q3 2010 to Q2 2018. t-Statistics are double-clustered by time and fund. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) *Morningstar Rating3 Yr$_t$*	(2) *Morningstar Rating3 Yr$_t$*	(3) *Morningstar RatingOverall$_t$*	(4) *Morningstar RatingOverall$_t$*
Misclassified$_{t\text{-}1}$	0.383***	0.170***	0.341***	0.182***
	(5.971)	(3.774)	(4.660)	(3.218)
Reported Credit Score$_{t\text{-}1}$	0.0698***	0.0299**	0.0588***	0.0289*
	(4.355)	(2.553)	(3.090)	(1.774)
Reported Duration$_{t\text{-}1}$	0.107***	−0.0277	0.113***	0.0122
	(3.679)	(−1.138)	(2.752)	(0.386)
Average Expenses$_{t\text{-}1}$	−1.024***	−0.755***	−0.822***	−0.622***
	(−6.915)	(−6.966)	(−5.045)	(−4.566)
3 Year Returns$_{t\text{-}1}$		15.22***		11.36***
		(8.036)		(6.202)
Time × Morningstar Reported Risk Style FE	Yes	Yes	Yes	Yes
Time × Morningstar Category FE	Yes	Yes	Yes	Yes
Observations	7,391	7,391	7,391	7,391
Adjusted R^2	0.211	0.541	0.170	0.373

misclassified funds hold significantly more high-risk bonds, the commensurate higher (on average) returns on these translate into more Morningstar stars. This is true even though, as we show, misclassified funds have zero outperformance when correctly benchmarked against same-risk peers and despite their returns crashing precisely at times when junk bonds crash (see in Table XI). Still, this leaves open the question of why funds would misclassify their risk level at all, as opposed to simply holding more risky assets while correctly reporting risk levels to Morningstar. Section IX of the Internet Appendix explores this question and shows that for a given level of Morningstar rating, investors direct significantly more flows to those funds that they "perceive" to have done this with their lower risk. Investors' perceptions are explained entirely by Morningstar's reported risk classification styles, not what the funds are actually holding. Thus, funds that misclassify by self-reporting overly safe risk levels take on higher risk (which translates into increased Morningstar stars) for "free" in terms of investors' perceptions, which translates into greater inflows.

In Table VI, we investigate whether misclassified funds are able to charge higher expense ratios than their peers. Perhaps intuitively, we explore whether misclassified funds charge their investors higher expenses because their "reported" (but not actual) performance is better and they are able to be rewarded higher Morningstar star ratings.

Table VI

Expense Ratios and Misclassification

In this table, we analyze whether misclassified funds are more expensive than usual. We regress average expense ratio on the *Misclassified* dummy and control variables. The average expense ratio is calculated at the fund level as the value-weighted average of the respective share-class level values. The sample period is Q3 2010 to Q2 2018. *t*-Statistics are double-clustered by time and fund. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) *Average* $Expense_t$	(2) *Average* $Expense_t$	(3) *Average* $Expense_t$
$Misclassified_{t\text{-}1}$	0.114***	0.0765***	0.0760***
	(6.356)	(4.186)	(4.172)
Reported Credit $Score_{t\text{-}1}$		0.0224***	0.0222***
		(3.611)	(3.592)
Reported $Duration_{t\text{-}1}$			−0.00790
			(−0.754)
Time × *Morningstar Reported Risk Style FE*	Yes	Yes	Yes
Time × *Morningstar Category FE*	Yes	Yes	Yes
Observations	8,373	7,586	7,586
Adjusted R^2	0.125	0.153	0.154

Prior research explored whether equity mutual funds are able to consistently earn positive risk-adjusted returns, and if so, whether funds are able to charge, in equilibrium, higher fees for this outperformance.[7] These studies often suggest that there should be a positive relation between before-fee risk-adjusted expected returns and fees. Gil-Bazo and Ruiz-Verdu (2009), in contrast, argue that funds often engage in strategic fee-setting in the presence of investors with different degrees of sensitivity to performance and that this could lead to an ambiguous or even negative relation between fund performance and fee.

Table VI presents our results on the fees of misclassified funds. As can be seen in column (3), we find that on average misclassified funds are associated with 7.6 bps higher ($t = 4.17$) average annual expenses than funds within the same style category, which implies that they are able to charge 10.8% higher fees than peers.[8]

[7] See, for example, Brown and Goetzmann (1997); Carhart (1997); Daniel et al. (1997); Wermers (2012); Cohen, Coval, and Pastor (2005); Kacperczyk, Sialm, and Zheng (2008); and Kosowski et al. (2006).

[8] Past research in the equity space investigates whether funds change 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. DiBartolomeo and Witkowskip (1997), for example, show that younger mutual funds are particularly prone to misclassification. Frijns et al. (2013) show that funds that switch fund objectives aggressively tend to have higher expense ratios, and Huang, Sialm, and Zhang (2011) argue that funds with higher expense ratios experience more severe performance consequences when they alter their risk. Related, Deli (2002) and Coles, Suay, and Woodbury (2000) argue that fee structures could vary across funds because of difficulty of managing a riskier portfolio. To test these ideas, in Sections X and XI of Internet Appendix, we explore both fund age, along with separating fees into advisor and distribution fees charged by managers (where available and reported).

In Table VII, we investigate fund flows to misclassified funds. Misclassification might be related to bond fund flows for several reasons. First, Barberis and Shleifer (2003) argue that investors tend to group assets into a small number of categories, leading to correlated capital flows and correlated asset price movements. If an asset ends up being in the wrong classification category, it may receive disproportionately higher (or lower) investment than its correct bucket especially if it has a favorable ranking attribute within that category (e.g., reported returns). Several papers in the literature document 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).

We test for the relationship between misclassification and flows in two ways. First, we simply test whether misclassified funds receive higher inflows. We find that they do—significantly so. In particular, in column (1) of Table VII, the coefficient on *Misclassified* is 0.0637 ($t = 4.95$), which implies that probability of positive flows is over 12% higher for misclassified funds controlling for other determinants. However, given that misclassification is also related to other attributes that drive flows (e.g., Morningstar stars), it is difficult to interpret how much of the flows might be coming from the misclassification itself. We therefore also employ a two-stage least squares procedure. In the first stage, controlling for other fund, category, and time effects we estimate the association between being a misclassified fund and the number of Morningstar stars that a fund receives (see Table V). We then take this estimate of *just* the extra portion of Morningstar stars a misclassified fund receives, and take 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, column (2) of Table VII implies that a one unit increase in *Misclassified Star* increases the probability of positive flows by almost 17.1% ($t = 5.16$).

We also examine if there is a difference between investor types (e.g., institutional versus retail) with respect to their responses to misclassified funds. From Morningstar Direct, we can classify share classes into a number of specific categories: in particular, into institutions, retirement, and retail classes. Results for these three investor types are reported in columns (3) to (5) of Table VII, respectively. We find that the positive flows accruing to misclassified funds appear to be coming from all types of investors. In particular, the coefficient on *Misclassified* is large and highly significant for all three shareclass categories. That said, individual investors appear to be tilted toward misclassifying funds to a slightly greater extent than institutions—while misclassified institutional share classes are 11.4% more likely to receive positive investor flows than other funds of their same share class, misclassified retail and retirement share classes increased their probabilities more than over 20% from their unconditional means. Among individual investors, the fact that retirement investors appear to be most influenced by misclassified funds in terms of flows is consistent with findings in Fisch, Lusardi, and Hasler (2019)

Table VII

Fund Flows and Misclassification

In this table, we analyze whether fund flows are related to misclassifications. There are two specifications for fund-level regressions. The first column regresses flow indicator on the *Misclassified* dummy directly. The second column regresses the flow indicator on misclassified stars. We then separately regress the flow indicator at the share-class level for institutional (3), retail (4), and retirement (5) classes against the *Misclassified* dummy. The sample period is Q3 2010 to Q2 2018. t-Statistics are clustered quarterly. *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.

	(1)	(2)	(3)	(4)	(5)
	Fund Portfolio		Institutional Share Class	RetailShare Class	Retirement Share Class
	$Flow_t > 0$	$Flow_t > 0$	$Flow_t > 0$	$Flow_t > 0$	$Flow_t > 0$
$Misclassified_{t\text{-}1}$	0.0637***		0.0639***	0.0905***	0.129***
	(4.947)		(3.639)	(4.368)	(5.356)
$Misclassified\ Stars_t$		0.171***			
		(5.155)			
$Reported\ Credit\ Score_{t\text{-}1}$	0.00438	–0.00422	0.00736*	–0.00435	–0.0117***
	(1.198)	(−0.757)	(1.864)	(−0.945)	(−2.906)
$Reported\ Duration_{t\text{-}1}$	0.0191***	0.00201	0.0145***	0.00537	–0.0259**
	(3.998)	(0.261)	(2.855)	(0.388)	(−2.590)
$Average\ Expenses_{t\text{-}1}$	–0.238***	–0.0685	–0.160***	–0.204***	–0.104**
	(−7.431)	(−1.409)	(−4.776)	(−5.826)	(−2.159)
Time × Morningstar Reported Risk Style FE	Yes	Yes	Yes	Yes	Yes
Time × Morningstar Category FE	Yes	Yes	Yes	Yes	Yes
Observations	7,766	7,391	7,248	4,306	5,733
Adjusted R^2	0.068	0.086	0.048	0.079	0.019

Table VIII

Characteristics of Misclassified Funds

In this table, we regress whether a bond fund is misclassified against various contemporaneous fund characteristics. *New Fund* indicates whether a fund has less than three years of history. *Log Size* is the log of total fund level AUM. *Number of Managers* and their *Average Tenure Length* are calculated using Morningstar Direct. *Only Taxable Bond Fund* indicates whether a fund is the only taxable bond fund present within a fund family. This is calculated by matching a fund to its family history information in the CRSP mutual fund database. *Number of Share Classes* is calculated from data provided by Morningstar Direct. *Market Share* is a fund's AUM as a percentage of the total AUM placed in all funds of a respective Morningstar category. *Past 3 Year Returns* is a fund's past three-year value-weighted net return for respective share classes. The sample period is Q3 2010 to Q2 2018. t-Statistics are clustered quarterly. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) *Misclassified*	(2) *Misclassified*	(3) *Misclassified*
New Fund	0.0668***	0.0785***	0.161***
	(3.834)	(4.257)	(5.673)
Log Size	0.0363***	0.0132**	0.00921
	(7.484)	(2.294)	(1.628)
Average Tenure Length	−0.000263**	−0.000232	−0.000350***
	(−2.054)	(−1.647)	(−2.829)
Number of Managers	0.000937	0.00589**	0.00347
	(0.490)	(2.662)	(1.506)
Number of Share Classes	0.0185***	0.0150***	0.0152***
	(11.81)	(10.34)	(9.867)
Only Taxable Bond Fund	0.0331**	0.0263*	0.0361**
	(2.357)	(1.947)	(2.500)
Market Share	−0.906***	1.548***	1.766***
	(−3.143)	(2.842)	(3.273)
Past 3 Year Returns			1.650***
			(6.834)
Time FE	Yes	No	No
Time × Morningstar Reported Risk Style FE	No	Yes	Yes
Time × Morningstar Category FE	No	Yes	Yes
Observations	7,612	7,543	7,543
Adjusted R^2	0.030	0.155	0.178

that financial literacy is significantly lower for retirement investors than other types of retail investors.

F. Who Misclassifies?

We turn next to examine the characteristics that correlate with a fund being a misclassified as well as the determinants of fund misclassification over time. Specifically, we first run characteristics regressions in which the dependent variable is a dummy indicating whether the fund is *misclassified*, to examine which characteristics are more related to misclassification. The results are reported in Table VIII. From Table VIII, we note a number of characteristics of misclassifiers. In particular, from the full specification in column (3), younger and larger funds tend to misclassify, as do managers earlier in their careers, that is, with less tenure. Moreover, the more share classes a fund has, the more likely it is to be a misclassifier. In addition, if the fund is the only taxable fixed-income fund in the family, it has a higher likelihood of being a misclassifier. Finally, consistent with the advantages of misclassifying that we document (e.g., being able to hold higher yielding bonds than correctly classified peers, resulting in higher returns and flows), we find that misclassifying funds enjoy a significantly higher share of the fund's risk category (*Market Share*) and observe higher realized returns when holding the (misclassified) riskier positions.

To explore the time series decisions of funds that begin and end misclassifying, we construct two variables to capture fund reporting behavior over time. The first, *Start Being Misclassified*, takes a value of 1 if a fund that was previously correctly classified starts to misclassify its holdings. Similarly, *End Being Misclassified*, takes a value of 1 if a previously misclassified fund starts correctly classify its holdings. We then test the determinants of these two variables. As can be seen in Panel A of Table IX, misclassification tends to start (end) when younger managers of funds that offer more share classes, have experienced particularly poor (positive) recent performance.

In Panel B of Table IX, we explore the geographic location of misclassifying (vs. nonmisclassifying) funds. We find that, relative to the Northeast (which has the highest prevalence of mutual funds and is the omitted category), funds in the Midwest appear less likely to misclassify, on average, while funds in the South appear more likely to misclassify.

Finally, in Panel C of Table IX, we explore the impact of a "family specific" effects on the misclassification of funds. We find that the inclusion of family fixed effect explains a large part of the variation in misclassification. In column (1), we find that Year-Quarter fixed effects explains 0.3% of the variation. When we also include family fixed effects in column (2), the R^2 increases to 22.7%. Thus, family-specific factors appear to explain over a fifth of the variation in which funds misclassify over time (controlling for any time-specific variation that might impact all funds, such as the Fed lowering target interest rates or a pervasive change in ratings). In column (3) in which we further add a fund specific fixed effects, the R^2 rises to 49.4%. This result suggests that even with the importance of family effects in determining misclassification, a sizable amount of the variation in misclassification remains determined at the fund level (as also suggested in Table VIII).

III. Misclassified Funds Returns across Junk Bond Regimes, Nonrated Securities, and Morningstar's Response

We have been in contact with Morningstar since the beginning of this project. We were first referred to technical support teams with whom we verified details about the self-reported surveys that fund managers fill out as well as Morningstar's scoring process to ensure that each step of our analysis was correct. After we posted a draft of our work, Morningstar released an official organizational response shown in Section V of the Internet Appendix.

Table IX

Further Determinants of Misclassifying over Time, Geographic Location, and across Families

In this table, we explore further determinants of misclassification. The sample period is Q3 2010 to Q2 2018. In Panel A, we explore how the start and end of misclassification related to various characteristics. In column (1), the left-hand-side variable is a dummy that indicates when a fund that was previously correctly classified starts misclassifying. In column (2), the left-hand-side variable is a dummy indicating for when a previously misclassified fund starts correctly classifying. In column (3), we regress (1) minus (2). *New Fund* indicates whether a fund has less than three years of history. *Log Size* is the log of total fund-level AUM. N*umber of Managers* and their *Average Tenure Length* are calculated using Morningstar Direct. *Only Taxable Bond Fund* indicates whether a fund is the only taxable bond fund present within a fund family. This is calculated by matching a fund to its family history information in the CRSP mutual fund database. *Number of Share Classes* is calculated from data provided by Morningstar Direct. *Market Share* is a fund's AUM as a percentage of the total AUM placed in all funds of a respective Morningstar category. *Past 3 Year Returns* is a fund's past three-year value-weighted net return for respective share classes. In Panel B, we regress the misclassification indicator against controls, category fixed effects, and geographic indicators, *Northwest*, *West*, *South*, and *Midwest* correspond to U.S. Census Bureau statistical regions. In Panel C, we regress the misclassification indicator against time, fund family, and fund fixed effects. t-Statistics are clustered quarterly. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Characteristics of Misclassified Funds

	(1) *Start Being Misclassified*	(2) *End Being Misclassified*	(3) *(Start-End) Misclassified*
New Fund	–0.00665	0.00366	–0.0103
	(−0.684)	(0.255)	(−0.534)
Log Size	0.00174	–0.00228	0.00402
	(0.553)	(−0.685)	(0.912)
Average Tenure Length	–0.000128***	–0.000142***	1.38e-05
	(−2.927)	(−2.947)	(0.197)
Number of Managers	0.000885	0.00335**	–0.00247
	(0.747)	(2.349)	(−1.478)
Number of Share Classes	0.00294***	0.00341***	–0.000475
	(2.843)	(3.767)	(−0.326)
Only Taxable Bond Fund	0.00445	–0.00588	0.0103
	(0.589)	(−0.651)	(0.857)
Market Share	0.454	1.010	–0.557
	(0.820)	(1.513)	(−0.538)
Past 3-Year Returns	–0.0738*	0.186**	–0.260**
	(−1.820)	(2.193)	(−2.731)
Time × Morningstar Reported Risk Style FE	Yes	Yes	Yes
Time × Morningstar Category FE	Yes	Yes	Yes
Observations	7,941	7,941	7,941
Adjusted R^2	0.004	0.028	0.011

(Continued)

Table IX—*Continued*

Panel B: Geography of Misclassification

	(1) *Misclassified*
Northeast	–
West	–0.0115
	(–0.727)
South	0.0677***
	(5.025)
Midwest	–0.0177
	(–1.655)
Controls	Yes
Time × Morningstar Reported Risk Style FE	Yes
Time × Morningstar Category FE	Yes
Observations	6,774
Adjusted R^2	0.153

Panel C: Misclassification and Fund Family Fixed Effects

	(1)	(2)	(3)
	Misclassified	*Misclassified*	*Misclassified*
Time FE	Yes	Yes	Yes
Family FE	No	Yes	Yes
Fund FE	No	No	Yes
Observations	6,923	6,919	6,906
Adjusted R^2	0.003	0.227	0.494

In Section VI of the Internet Appendix, we provide our reply to Morningstar's initial comments. Morningstar then released a second response as shown in Section VII of the Internet Appendix. Our reply to these comments is available in Section VIII of the Internet Appendix.

In their first response, Morningstar made two main points. First, they argued that our star analysis was misspecified due to not comparing within Morningstar official fund category (see Section IV of the Internet Appendix).[9] In this paper, all specifications include official *Morningstar category* fixed effects. From these tests, comparing within categories, all of our results are strong and significant. Misclassified funds receive significantly more stars than peer-group funds within an official Morningstar category. Second, Morningstar posited that the discrepancies are due largely to how their classification formula treats nonrated bonds. But as we show in Section VI of the Internet

[9] In the internet appendix, we replicate the Morningstar Star Rating methodology itself. We show that Misclassified funds receive significantly more Stars from taking on more risk in their underlying portfolios, and get these Stars for "free" in the sense investors perceive these funds as being less risky and so allocate significantly more flows to them as a result (as we show that even conditional on the same number of Stars, investors allocate significantly more flows to funds that they believe attain these flows while taking on lower risk).

Table X

Characteristics of Unrated Bonds Held by Funds

This table summarizes the corporate bonds in the Mergent FISD database that were issued between 2010 and 2016. Each box describes the mean offering yield and the number of bonds in different ranges of offering maturities and credit qualities. A bond's credit rating at issuance is the Barclays/Bloomberg composite of Fitch, Moody's, and S&P's ratings that were available within 30 trading days of the offering date. N is the number of issue observations in each box.

	Issuing Maturity		
	0 to 3.5 Years	3.5 to 6 Years	6 to 10 Years
High Investment Grade (AA to AAA)	1.44%	2.21%	2.70%
	$N = 113$	$N = 146$	$N = 33$
Medium Investment Grade (BBB to A)	1.75%	2.80%	3.75%
	$N = 483$	$N = 1{,}110$	$N = 370$
Junk Grade Bonds (BB and Below)	5.14%	8.08%	7.69%
	$N = 43$	$N = 563$	$N = 1{,}655$
Unrated Bonds	7.81%	6.43%	7.09%
	$N = 81$	$N = 245$	$N = 356$

Appendix, when we omit all funds that have *any* nonrated bonds, our results remain large and significant (in fact larger in point-estimate in some cases).

Notwithstanding, we examine the characteristics and behaviors of the nonrated bonds themselves, and the misclassified funds that hold them. In Table X, we look at the nonrated bonds. As can be seen, the yields of nonrated bonds look incredibly similar to those of junk bonds, and very little like those of higher rated bonds that they are purported to be by fund managers, who Morningstar takes at their word. Second, in Table XI, we examine the performance of misclassified funds around times of junk bond crashes and junk bond outperformance. If the funds classified into "safer" categories by Morningstar truly hold the high-credit-quality bond issues they claim as represented by Morningstar in their relatively safe risk classifications of the funds, the funds should not be sensitive to the movements of junk bonds. This is not what we see, however, in Table XI. Rather, Table XI shows that misclassified funds' over- and underperformance relative to their peer funds relates strongly to junk bond returns (captured by the return on a junk bond index—*JNK*). In particular, misclassified funds significantly underperform precisely when the junk bond market crashes, while they experience their greatest fund outperformance when the junk bond market surges (even though they are represented as holding primarily highly rated, safe securities).

Morningstar's second response (Section VII of the Internet Appendix) shifts focus to more technical points, stating that "To that end, we were able to largely reproduce the authors' multivariate analysis of the binary 'misclassified' dummy variable they defined and various ratings metrics." In Section VIII of the Internet Appendix, we explore the points in their second response in more detail by taking them to the data, and unfortunately do not find strong support for their claims.

Table XI

Misclassified Fund Performance around Junk Bond Crashes (and Outperformance)

In this table, we regress monthly fund returns on Misclassified, a dummy equal to 1 if the official credit quality (high or medium) is higher than the counterfactual credit quality, and 0 otherwise, and control variables. In the columns, we regress separately the sample months when JNK, the SPDR Bloomberg Barclays High Yield Bond ETF, had (1) major negative returns, (2) close to zero returns, and (3) substantial positive returns. The sample period is Q3 2010 to Q2 2018. t-Statistics are clustered quarterly. *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively.

	(1) *Fund* $Return_t JNK <$ -1%	(2) *Fund* $Return_t -1\% <$ $JNK < 1\%$	(3) *Fund* $Return_t JNK >$ 1%
Misclassified$_{t-1}$	−4.672***	2.849**	7.630***
	(−3.117)	(2.270)	(6.141)
Reported Credit Score$_{t-1}$	−1.421**	0.263	2.024***
	(−2.270)	(0.828)	(5.232)
Reported Duration$_{t-1}$	−4.784	0.378	5.227**
	(−1.246)	(0.187)	(2.327)
Average Expense$_{t-1}$	−9.683***	−2.234*	−1.146
	(−3.457)	(−2.011)	(−0.945)
Time × Morningstar Reported Risk Style FE	Yes	Yes	Yes
Time × Morningstar Category FE	Yes	Yes	Yes
Observations	4,522	9,972	8,177
Adjusted R^2	0.855	0.879	0.820

IV. Conclusion

Investors rely on external information intermediaries to lower their cost of information acquisition. Although prima facie this does not pose a problem, if the information that the intermediary is passing on is biased, such bias will propagate throughout markets and can cause real distortions in investor behavior and market outcomes. We document precisely this effect 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 biased reports by fund managers that Morningstar simply passes on as truth.

This paper is the first systematic study to compare fund-reported asset profiles provided by Morningstar against their *actual* portfolio holdings. We document that evidence of significant misclassification across the universe of bond funds. Specifically, a large portion of bond funds are not passing on a realistic view of the fund's actual holdings to Morningstar, which creates its risk classifications based on the self-reported data. As a result, up to 31.4% of funds 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 their category peers. We also find evidence suggesting that the misreporting has real effects on investor behavior and mutual fund success—misclassified funds charge significantly higher fees but receive higher flows from investors.

We exploit a novel setting in which investor 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 reported information. We view our study is a first step toward thinking about a market design in which information intermediaries have more aligned incentives to better process the information they gather from market constituents and share with investors. Future research should explore alternative monitoring and verification mechanisms for increasingly complex information aggregation in modern financial markets, along with ways that investors can engage as partners in information collection and price-setting.

Initial submission: April 6, 2020; Accepted: October 1, 2020
Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

REFERENCES

Acharya, Viral, and Hassan Naqvi, 2019, On reaching for yield and the coexistence of bubbles and negative bubbles, *Journal of Financial Intermediation* 38, 1–10.

Agarwal, Vikas, Brad M. Barber, Si Cheng, Allaudeen Hameed, and Ayako Yasuda, 2019, Private company valuations by mutual funds, Working paper, Harvard Business School.

Andonov, Aleksandar, Rob M. M. J. Bauer, and K. J. Martijn Cremers, 2017, Pension fund asset allocation and liability discount rates, *Review of Financial Studies* 30, 2555–2595.

Bams, Dennis, Roger Otten, and Ehsan Ramezanifar, 2017, Investment style misclassification and mutual fund performance. Paper presented at 28th Australasian Finance and Banking Conference. https://doi.org/10.2139/ssrn.2648375

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161–199.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2005, Comovement, *Journal of Financial Economics* 75, 283–317.

Becker, Bo, and Victoria Ivashina, 2015, Reaching for yield in the bond market, *Journal of Finance* 70, 1863–1901.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2019, What do mutual fund investors really care about? Fisher College of Business Working Paper No. 2019-03-005.

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

Bollen, Nicholas, and Veronica Pool, 2009, Do hedge fund managers misreport returns? Evidence from the pooled distribution, *Journal of Finance* 64, 2257–2288.

Bollen Nicholas, and Veronica Pool, 2012, Suspicious patterns in hedge fund returns and risk of fraud, *Review of Financial Studies* 25, 2673–2702.

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

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

Brown, Stephen J., and William N. Goetzmann, 1997, Mutual fund styles, *Journal of Financial Economics* 43, 373–399.

Budiono, Diana, and Martin Martens, 2009, Mutual fund style timing skills and alpha, Fisher College of Business Working paper.

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, *Journal of Finance* 52, 57–82.

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

Choi, Jaewon, and Mathias Kronlund, 2017, Reaching for yield in corporate bond mutual funds, *Review of Financial Studies* 31, 1930–1965.

Choi, Jaweon, Mathias Kronlund, and Ji Y.J. Oh, 2018, Sitting bucks: Zero returns in fixed income funds, SSRN Working paper.

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.

Cohen, Randolph B., Joshua D. Coval, and Lubos Pástor, 2005, Judging fund managers by the company they keep, *Journal of Finance* 60, 1057–1096.

Coles, Jeffrey L., Jose Suay, and Denise Woodbury, 2000, Fund advisor compensation in closed-end funds, *Journal of Finance* 55, 1385–1414.

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

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.

Deli, Daniel N., 2002, Mutual fund advisory contracts: An empirical investigation, *Journal of Finance* 57, 109–133.

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

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, 2018, A model of monetary policy and risk premia, *Journal of Finance* 73, 317–373.

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.

Fisch, Jill E., Annamaria Lusardi, and Andrea Hasler, 2019, Defined contribution plans and the challenge of financial illiteracy, *Cornell Law Review* 105, 19–22.

Frijns, Bart, Aaron B. Gilbert, and Remco C. J. Zwinkels, 2013, On the style-based feedback trading of mutual fund managers, *Journal of Financial and Quantitative Analysis* 51, 771–800.

Froot, Kenneth, and Emil Dabora, 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, *Journal of Finance* 64, 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, 2575–2616.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2008, Unobserved actions of mutual funds, *Review of Financial Studies* 21, 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.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White, 2006, Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance* 61, 2551–2595.

Kruger, Phillip, Augustin Landier, and David Thesmar, 2012, Categorization bias in the stock market, Working paper, University of Geneva, Toulouse School of Economics, and HEC.

Lian, Chen, Yueran Ma, and Carmen Wang, 2019, Low interest rates and risk-taking: Evidence from individual investment decisions, *Review of Financial Studies* 32, 2107–2148.

Ljungqvist, Alexander, Christopher Malloy, and Felicia Marston, 2009, Rewriting history, *Journal of Finance* 64, 1935–1960.

Nanda, Vikram, Z. Jay Wang, and Lu Zheng, 2004, Family values and star phenomenon: Strategies for mutual fund families, *Review of Financial Studies* 17, 667–698.

Ozdagli, Ali, and Zixuan Wang, 2019, Interest rates and insurance company investment behavior, Working paper, Federal Reserve Bank of Boston and Harvard Business School.

Rajan, Raghuram, 2013, A step in the dark: Unconventional monetary policy after the crisis, *Andrew Crockett Memorial Lecture, Bank for International Settlements, Basel*.

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 databases and their implications, *Journal of Finance* 29, 1303–1310.

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

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.

Stein, Jeremy, 2013, Overheating in credit markets: Origins, measurement, and policy responses, Research Symposium, Federal Reserve Bank of St. Louis.

Swinkels, Laurens, and Liam Tjong-A-Tjoe, 2007, Can mutual funds time investment styles? *Journal of Asset Management* 8, 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, SSRN Working paper.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Internet Appendix.
Replication Code.