

Discretionary Information in ESG Investing: A Text Analysis of Mutual Fund Prospectuses

Angie Andrikogiannopoulou*

King's College London

Philipp Krueger†

University of Geneva & SFI

Shema Mitali‡

EPFL

Filippos Papakonstantinou§

King's College London

Abstract

We construct novel measures of funds' environmental, social, and governance (ESG) commitment by applying text analysis to the discretionary investment-strategy descriptions in their prospectuses. We find that investors respond strongly to text-based ESG measures. Using discrepancies between text- and fundamentals-based ESG measures, we identify greenwashing funds. We find greenwashing is more prevalent in the last five years and among funds with lower past flows and weaker oversight. Furthermore, greenwashing funds attract similar flows as funds that truthfully reveal their ESG commitment, suggesting that investors cannot distinguish between them. On the other hand, greenwashers have inferior performance than genuinely green funds.

Keywords: *ESG, Prospectus, Greenwashing, Text Analysis*

Mutual Funds, Fund Flows, Fund Performance

JEL Classification: *G11, G23*

*King's Business School, Department of Banking & Finance, Bush House, 30 Aldwych, London WC2B 4BG, UK, e-mail: aandriko@kcl.ac.uk.

†University of Geneva, Geneva School of Economics and Management, 40 Bd du Pont-d'Arve, CH 1211 Geneva 4, Switzerland, e-mail: philipp.krueger@unige.ch.

‡École Polytechnique Fédérale de Lausanne, EXTRA 214 (Extranef UNIL) Quartier, UNIL-Dorigny, CH 1015 Lausanne, Switzerland, e-mail: shema.mitali@epfl.ch.

§King's Business School, Department of Banking & Finance, Bush House, 30 Aldwych, London WC2B 4BG, UK, e-mail: fpapakon@kcl.ac.uk.

1 Introduction

In recent years, interest in ESG investing has grown exponentially. Google searches for the term “ESG” have grown at an average rate of forty percent per year, more than quintupling in the last five years. The rise in ESG interest has also translated into a rapid growth in capital flows to sustainable mutual funds, which have similarly more than quintupled in the last five years. This growing significance of ESG considerations in fund management heightens the importance of the information that is available about funds’ ESG investments, as this information should in principle facilitate the matching between like-minded investors and fund managers. But, while there is an extant literature that studies the *objective* information available to investors—e.g., information gleaned from portfolio holdings and returns—much less is known about the *discretionary* information that funds themselves release. In this paper, we study the discretionary information released through fund prospectuses, which enables us to make several contributions. Specifically, we conduct a text analysis of fund prospectuses to learn how and how much funds talk about ESG investments. We also compare the discretionary with the objective ESG-related information to understand if and when the two deviate. Subsequently, we analyze whether investors pay attention to this discretionary information and, in case it is manipulated, whether they are deceived by it or not. Succinctly, we find that there is often a discrepancy between funds’ words and actions and that, while investors pay attention to both, they can be deceived.

Our study is motivated by a growing literature showing that mutual fund investors have limited attention and information processing capacity so they attend more to fund information that is more salient, i.e., easier to access and process (Barber, Odean and Zheng, 2005; Hartzmark and Sussman, 2019; Kostovetsky and Warner, 2020). Fund prospectuses are easy-to-access through fund websites and easy-to-read as they need to comply with the SEC’s plain language requirement, hence they constitute a more salient source of information about funds’ ESG activities than information on fundamentals, such as holdings and returns.¹ As such, when choosing between funds, investors

¹Mutual fund prospectuses are mandated by SEC regulations as a means of promoting effective communication between funds and their prospective investors. These regulations specifically require that prospectuses contain key information, updated at least annually, about a fund’s investment objectives and strategies, risks, expenses, and performance. Furthermore, this information must be communicated in a clear and concise manner that can be understood by an average or typical investor who may not be sophisticated in legal or financial matters.

may react strongly to the information disclosed by prospectuses. We thus start our analysis by constructing a set of novel measures of each mutual fund’s engagement with ESG activities that is based on the number and choice of ESG-related words used to describe the fund’s Principal Investment Strategy (PIS) in its prospectus. Specifically, we calculate the number and frequency of ESG keywords, the length and location of the ESG-related text, as well as the readability, uniqueness, and tonality of the ESG-related text. Throughout the 2011–2020 period that we study, the number of funds that discuss ESG in their PIS has grown exponentially, mirroring the growth in the public’s interest in ESG investments and rising to about 450 funds (about 10% of the total) in 2020. Among funds that discuss ESG, the ESG-related portion of the PIS is, on average, about a fifth of the entire PIS and it appears about halfway through the PIS though in 13% of cases it starts at the very first sentence. In addition to our text-based measures of ESG, we also construct a second set of ESG measures that are based on objective or fundamental information. In line with prior research (e.g., Starks, Venkat and Zhu, 2017), we use fund holdings data to calculate a weighted average of the ESG scores of the stocks in each fund’s disclosed equity portfolio. Furthermore, we perform a style analysis on fund returns and calculate the weighted average of each fund’s style weights on stock portfolios sorted by their ESG scores. We refer to the first set of measures based on the text analysis of the prospectus as “text-based”, and to the holdings- and returns-based measures as “fundamentals-based” measures of funds’ ESG intensity.

The first step in our analysis is to study whether fund flows respond to our text- and fundamentals-based ESG intensity measures. Our first finding is that fund flows respond strongly to text-based ESG measures. This result delivers a valuable insight to fund managers, as it demonstrates the effectiveness of communicating their ESG commitment to potential investors through the prospectus. Furthermore, we find that fund flows respond more strongly to text-based than to fundamentals-based ESG measures. This confirms the finding that salience matters in mutual fund investing, and extends it to the space of ESG investing. We verify that our findings are robust to using alternative fundamentals-based ESG measures. For instance, our findings hold up to calculating our holdings-based ESG measures using data from a variety of providers that employ different methodologies to assess corporations’ ESG practices, as well as to using primary data on specific

ESG dimensions that investors may care about such as greenhouse gas emissions (see Bolton and Kacperczyk, 2021). Likewise, they hold up if we replace our measures with commercial fundamentals-based ESG fund ratings such as those issued by Morningstar. We also verify that the documented relations are robust to including fund fixed effects, suggesting that within-fund changes in the ESG narrative of the prospectus trigger an increase in investor flows.

Repeating our analysis separately for funds targeted to institutional investors and funds targeted to retail investors, we establish that both types of investors respond to the ESG information in fund prospectuses, but institutional investors are more sensitive to it. On the other hand, only institutional investors respond significantly to holdings-based ESG measures. These findings are consistent with the idea that institutional investors are more likely to be sophisticated and/or pay attention to (and process) information than retail investors. They are also consistent with existing evidence that institutional investors chase more sophisticated performance measures such as multi-factor alphas while retail investors react to simple return measures like past raw returns (Del Guercio and Tkac, 2002; Evans and Fahlenbrach, 2012). Finally, we examine the effect of certain writing style characteristics of the ESG portion of the PIS. We find that a more readable and less ambiguous writing style strengthens the effect of text-based ESG intensity on fund flows, reinforcing the idea that investors have limited information processing capacity and are thus more likely to react to financial disclosure information that is easier to process.

In the second part of the paper, we study whether funds engage in greenwashing, i.e., if they represent their investment strategy in the prospectus in a way that makes it appear more heavily tilted toward ESG than it is in reality. This analysis produces important insights for asset allocators and regulators who seek to identify funds that do not follow through on their ESG commitments. To conduct our analysis, we leverage our text-based against the fundamentals-based ESG measures and construct a time-varying measure of greenwashing at the fund-level. The rationale of our greenwashing measure is that our text-based measure is based on discretionary information that a fund chooses to reveal in its prospectus, while fundamentals-based measures are based on objective information on a fund’s holdings and returns. Thus, it identifies funds that—relative to their actual fundamentals—are likely overstating their ESG commitments in their prospectuses.

Specifically, our baseline greenwashing measure is an indicator variable that identifies funds that include ESG-related words in their PIS—placing themselves in the top 5% of funds in terms of text-based ESG—but exhibit a holdings-based ESG measure below the median across all funds in the same investment category in the same month. In setting a high minimum discrepancy necessary to be classified as a greenwasher—top 5% versus bottom 50%—we essentially focus on relatively egregious greenwashers. In robustness checks, we set the discrepancy threshold even higher and we require that funds also have a low returns-based ESG measure to be classified as greenwashers.

Using our greenwashing measures, we explore (1) what types of funds are more likely to engage in greenwashing, (2) whether investors can distinguish between funds that greenwash versus those that talk about ESG in their prospectus and follow through in terms of investments, and (3) whether greenwashing is associated with higher or lower performance. First, we find that greenwashing behavior is more likely to be observed in the second half of our sample period (i.e., after 2016), and is associated with funds that have lower past flows and higher expense ratios. These findings are consistent with the hypothesis that funds are more likely to greenwash if they have (i) a potential benefit from doing so (as is the case later in the sample period, when ESG interest is particularly high), (ii) a strong incentive to do so (as is the case when past flows are low), and/or (iii) the opportunity to go undetected (as is the case for funds with weak governance, which often have high expense ratios; see Gil-Bazo and Ruiz-Verdú, 2009). Second, we find that, contrary to what rationality would imply, investors direct their flows to funds that talk about ESG in their prospectus regardless of whether these funds back their words with actions or not. Third, we find that greenwashing funds do not have better investment performance; this result serves as a useful sanity check of our greenwashing measure. In contrast, our analysis reveals that non-greenwashing funds—i.e., those that truthfully reveal and actually engage in ESG investing—exhibit better performance.

This paper is related to several strands of the literature. First, it is related to the growing literature that applies text analysis to information shared with investors (see Loughran and McDonald, 2016, for an early review). Most studies focus on corporate regulatory filings such as 10-K reports,² while a few recent studies apply text analysis to information shared by mutual funds in their

²For example, Hoberg and Phillips (2016) construct a firm’s set of competitors by calculating product similarities from 10-K report product descriptions, and Cohen, Malloy and Nguyen (2020) show that substantial active changes in firms’ 10-K reports predict important changes in their operations and returns.

prospectuses. Krakow and Schäfer (2020) and Sheng, Xu and Zheng (2022) study the information content of fund risk disclosures. Kostovetsky and Warner (2020) analyze investment strategy descriptions and find that product differentiation—measured by the uniqueness of a fund’s strategy description—has a short-term positive effect on fund flows. Furthermore, Abis and Lines (2022) analyze fund strategy descriptions to allocate funds to peer groups and find that funds’ portfolios do not deviate much from the average portfolio in their peer group, as such deviations are penalized by the market through capital outflows. Our paper contributes to this literature by studying the ESG content of fund prospectuses and the effect of greenwashing on fund flows. Similar to Kostovetsky and Warner (2020), we find that investors respond more to text-based than to fundamentals-based measures; however, focusing on the ESG-related portion of the prospectus, we do not find that product differentiation affects flows, possibly because an ESG focus is already a significant differentiating feature of a fund’s investment strategy. Contrary to Abis and Lines (2022), in the context of funds’ ESG investing, we do not find evidence of market discipline, rather we find that there *are* substantial deviations between what funds talk about and what they invest in, and that investors tend to pay more attention to the former and can be deceived. This is possibly because ESG is a hot investment theme, hence divergence between words and actions is more likely to occur as managers are tempted to attract investors’ attention, and it is still not well-understood by investors, hence evaluating such divergence may be more complicated than for traditional investment styles.

Second, our paper is related to the literature documenting that funds take cosmetic actions to exploit investor sentiment. For example, Cooper, Gulen and Rau (2005) find that funds change their names to reflect “hot” investment styles without making corresponding changes to their holdings. Using different settings and samples, recent papers have highlighted that some institutional investors sign the United Nations Principles for Responsible Investment (PRI) but do not follow through on their ESG commitments (Gibson Brandon et al., 2021; Kim and Yoon, 2021; Liang, Sun and Teo, 2020). We contribute to this literature by using regulatory filings to construct a novel measure of greenwashing that is based on text analysis of fund prospectuses, which (i) relative to fund names, change more frequently, and (ii) relative to PRI sign-ups, tend to be a more nuanced way of signaling to investors the fund’s commitment to ESG investing.

Finally, our paper is related to research concerned with mutual fund ESG strategies, and particularly their effect on fund flows. For instance, Hong and Kostovetsky (2012) show that the political orientation of a fund manager affects the fund’s ESG strategy. Other papers have focused on the effect of commercial fundamentals-based ratings like Morningstar’s sustainability fund ratings on investor flows (e.g., Hartzmark and Sussman, 2019; Ammann et al., 2019). Our results confirm that commercial ESG ratings affect flows but suggest that the effect of text-based ESG measures is stronger. Furthermore, while the previously documented effect of Morningstar’s ratings appears to be short-term (Gantchev, Giannetti and Li, 2021), the effect of text-based ESG measures we find is long-term. More broadly, our finding on the effect of text-based ESG on fund flows contributes to a large literature that studies the determinants of mutual fund flows including past performance (e.g., Chevalier and Ellison, 1997), fees (Barber, Odean and Zheng, 2005), marketing (e.g., Cooper, Gulen and Rau, 2005, Jain and Wu, 2000, Kostovetsky and Warner, 2020), and fund manager characteristics (e.g., Kumar, Niessen-Ruenzi and Spalt, 2015), among others.

The rest of the paper is organized as follows. In Section 2, we discuss our data sources. In Section 3, we describe how we construct our text-based and fundamentals-based ESG measures and our greenwashing measure that captures the discrepancy between words and actions. In Section 4 we present the results from our analysis, and in Section 5 we conclude.

2 Data

To construct the data set we use in this paper, we draw from a variety of data sources that we describe below.

2.1 Fund prospectuses

We obtain data on funds’ investment strategies from the “Mutual Fund Prospectus Risk/Return Summary Data Sets” that the SEC makes available on its website.³ The data contain numeric and text information extracted from the risk/return summary section of mutual fund prospectuses

³See <https://www.sec.gov/dera/data/mutual-fund-prospectus-risk-return-summary-data-sets>.

filed with the SEC in a structured XBRL format.⁴ For each prospectus (485BPOS form) filed by a fund, we observe fund identifying information (e.g., fund name, fund family’s Central Index Key, and Series ID), the document’s filing date, and all narrative text blocks included in the summary section (e.g., a description of the fund’s investment strategies, exposed risks, and past performance). In our analysis, we focus on the *Principal Investment Strategy* (PIS) text block in which funds are required to provide information about their principal investment strategies, and specifically to identify the types of securities in which the fund principally invests; to describe any policy, practice, or technique used by the fund to achieve its investment objectives; and to explain in general terms how the fund’s adviser decides which securities to buy and sell.⁵ Our data start in January 2011 (when funds started submitting XBRL data) and stop in June 2020. Funds are required by the SEC to update their summary information at least once per year. Importantly, we note that, for the period up to 2018, all text blocks in the flattened XBRL data provided by the SEC are truncated at 2,048 characters, which affects about half of all PIS text blocks and would be problematic for our analysis as it relies on analyzing this text. To address this problem, we write our own code that downloads the original prospectuses from the SEC’s EDGAR website and extracts the full text under the PIS section. In the Appendix, we present some examples of fund PIS descriptions.

2.2 Fund information

We obtain fund-level information for the period January 2010 to June 2020 from the Center for Research in Security Prices (CRSP) Survivorship-Bias-Free U.S. Mutual Fund Database. We focus on actively managed funds operating in the United States that invest in domestic and foreign equities, so we exclude fixed income, money market, balanced, index, exchange-traded, and variable annuity funds. To exclude these, we use the CRSP objective codes and flags as well as keyword searches in

⁴The risk/return summary section appears at the beginning of the statutory prospectus (sometimes also filed as a separate summary prospectus) and contains key information about the fund in the following standardized order: (i) investment objectives; (ii) fees and expenses; (iii) principal investment strategies, risks, and performance; (iv) investment advisers and portfolio managers; (v) purchase and sale of fund shares; (vi) tax information; and (vii) financial intermediary compensation. Starting in 2011, all funds are required to report this information in the eXtensible Business Reporting Language (XBRL) format which organizes the filed data under different “tags” making it easier to analyze and compare over time and across funds. Funds are also required to make this information available on their own websites.

⁵See a reference copy of the form funds need to file, with instructions, at <https://www.sec.gov/files/formn-1a.pdf>.

fund names.⁶ The CRSP data include share class-level characteristics such as net returns at a daily frequency, total net assets (TNA) at a monthly frequency, and expense, portfolio turnover, 12b-1 and load fee ratios at an annual frequency. They also contain comprehensive information on all fund holdings at a quarterly frequency, including long and short positions in U.S. and international equities.⁷

Our analysis is conducted at the fund-month level. We use the CRSP_CIK_MAP file provided by WRDS to link the CRSP share class identifiers (crsp_fundno) with the fund identifiers in the SEC filings (CIK-Series ID pairs).⁸ We then aggregate all share class data at the fund level using the fund's Series ID. While all fund share classes have the same prospectus and holdings, they have different TNA, fees, and net returns. We compute a fund's TNA as the sum of its classes' TNAs, and a fund's net return (fees) as the weighted average of its classes' returns (fees), with weights equal to the beginning-of-month TNA value of each class. To improve the accuracy of our return data, we omit any monthly return that directly follows a missing one, as it may compound multiple months' returns.

We calculate monthly fund flows for fund i during month t as

$$Flows_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1} (1 + r_{i,t})}{TNA_{i,t-1} (1 + r_{i,t})}, \quad (1)$$

where $TNA_{i,t}$ is total net assets and $r_{i,t}$ is net return of fund i over month t . To reduce the effect of outliers, we winsorize fund flows at the 1% and 99% level.

We proxy for funds' marketing expenses by combining a fund's 12b-1 fee and front load, similar to Roussanov, Ruan and Wei (2021).

Finally, we calculate monthly fund alphas based on the 4-factor model of Carhart (1997) using

⁶CRSP objective codes are constructed using objective codes from three different sources (Wiesenberger, Strategic Insight and Lipper) in order to provide consistency in style classifications. We exclude ETFs and variable annuity funds based on the *et_flag* and *vau_fund* variables in CRSP. We remove index funds using the *index_fund_flag* variable, by searching for the words “INDEX”, “S&P”, “DOW 30”, and “NASDAQ” in fund names, and by removing all funds in specific fund families (Dimensional Fund Advisors, Direxion, Potomac, ProFunds, and Rydex) which follow index-like strategies.

⁷We note that, starting in 2008 (hence for our entire sample period), the CRSP mutual fund holdings database is more complete and reliable than the Thomson Reuters s12 database.

⁸We match each prospectus to the CRSP fund-month observation in the month after the filing, and retain this prospectus for subsequent CRSP fund-months until a new prospectus is filed.

daily fund and factor returns. We use the CRSP NYSE/Amex/NASDAQ value-weighted index as the market factor, the one-month Treasury bill rate as the risk-free rate, and the factor-mimicking portfolios for size, book-to-market, and momentum, all downloaded from Kenneth French’s website.

Our final sample contains 4,863 funds, which have filed 37,399 prospectuses between January 2011 and June 2020. In Panel A of Table 1, we present summary statistics for the characteristics of all funds in our sample. The average fund in our sample is about 13 years old, has total net asset value of about 1.6 billion U.S. dollars, fund inflows of about 0.39% per month, and net returns (4-factor alpha) of about 0.78% (-0.10%) per month, and is equally likely to be primarily targeted to institutional or to retail investors.

2.3 ESG data

We obtain ESG data from several providers, namely MSCI, Sustainalytics, Refinitiv (formerly Thomson Reuters ASSET4), and Morningstar.

From MSCI, Sustainalytics, and Refinitiv, we obtain annual firm-level ESG scores, which we use to calculate fund-level scores (see Section 3.2 below). MSCI ESG scores capture a company’s resilience to ESG risks, measured at the GICS (Global Industry Classification Standard) sub-industry level, on a scale from 0 to 10. Sustainalytics computes a score, from 0 to 100, based on a company’s ESG risk relative to its industry peers and how well it manages that risk. Refinitiv measures, on a scale from 0 to 100, a company’s ESG performance relative to its TRBC (The Refinitiv Business Classification) industry group peers, using several themes such as carbon emissions, human rights, or green innovation. Most often, data providers construct these ESG scores by combining firm-reported data (e.g., from sustainability reports) on how firms manage issues such as environmental impacts, workplace safety, or human rights concerns, and information reported by third parties (e.g., non-governmental organizations). While methodologies can vary across data providers, all these measures generally capture the quality of a firm’s ESG policies and practices. Since the MSCI database has the best cross-sectional coverage and has been found to contain the least amount of noise for U.S. firms (Berg et al., 2021), we use it as our primary source. However, as a robustness check, we also construct a combined score that averages the standardized z -scores from all three data providers. For

more details on the construction of ESG scores from different providers and their correlations see, for instance, Gibson Brandon, Krueger and Schmidt (2021) and Berg, Koelbel and Rigobon (2022).

From Morningstar, we obtain monthly fund-level globe ratings, which were first made available to investors in March 2016. The globe rating of a fund is based on the percentile rank of its portfolio sustainability score relative to other funds in the same Morningstar category. This portfolio sustainability score is a value-weighted average based on firm-level ESG scores from Sustainalytics and an adjustment that accounts for a company's involvement in ESG-related controversies.

Finally, we complement the ESG score data with firm-level carbon emissions (CO_2) data from Refinitiv. Specifically, for each firm, we obtain its annual total emissions, defined as the sum of direct and indirect emissions. Direct emissions are the equivalent of Scope 1 emissions from the Greenhouse Gas (GHG) Protocol, that is emissions from sources that are owned or controlled by the firm, from the generation of intermediate energy, from transportation/business travel where the transportation methods are owned by the firm, emissions from the operation of buildings, and emissions associated with the sale of own-generated electricity to another firm. Indirect emissions are the equivalent of Scope 2 GHG Protocol emissions, that is they include emissions arising from the generation of imported (purchased) electricity, heat, or steam consumed by the firm; they exclude Scope 3 GHG Protocol emissions, as these emissions are not controlled or owned by the firm and it has no influence over them.

3 Constructing fund-level ESG measures

We construct two sets of ESG measures for each fund. The first utilizes fund prospectuses to capture how and how much funds talk about ESG investments. The second utilizes fund fundamentals (holdings and returns) to capture funds' actual ESG investments. Subsequently, we use our text-based and fundamentals-based ESG measures to devise a measure of greenwashing for each fund, that captures the discrepancy between words and actions.

3.1 Text-based ESG measures

To construct text-based ESG scores for funds, we perform text analysis on the PIS text block of fund prospectuses. First, we clean each fund’s PIS text by removing common non-letter characters and punctuation marks that do not end sentences, and by capitalizing all letters. Each PIS is then split into sentences. Then, we form a list of ESG keywords/phrases, and finally we search for these in funds’ PIS text. In Table 2, we show our list of ESG keywords and the frequency with which we encounter each of these words across all PISs in our sample. Using this list of ESG keywords, we then calculate several text-based measures of ESG intensity, positioning, readability, tonality, and uniqueness.

ESG intensity. Our main measure of ESG intensity is the relative frequency of ESG mentions in each fund’s PIS, i.e., the number of all occurrences of ESG keywords divided by the total number of words in the PIS. As our list of ESG keywords is not exhaustive, we also construct an alternative measure which captures the relative length of the ESG-related portion of the PIS, i.e., the number of words in sentences containing at least one ESG keyword divided by the total number of words in the PIS.

ESG positioning. This measure is based on the location of ESG keywords in the PIS. Our main measure is defined as the proportion of the text (or, in robustness checks, the number of words) from the beginning of the corpus to the beginning of the first sentence containing at least one ESG keyword. As an alternative, we also consider a dummy indicating if at least one ESG keyword appears in the first sentence of the PIS.

ESG readability. We construct two measures of readability for the ESG portion of each fund’s PIS (i.e., all sentences containing at least one ESG keyword): the Flesch Reading Ease and the Gunning Fog Index.⁹

ESG tonality. We measure the vagueness of the ESG strategy description based on the frequency

⁹Flesch Reading Ease is defined as $206.835 - 1.015 \frac{\#words}{\#sentences} - 84.6 \frac{\#syllables}{\#words}$ (Flesch, 1948). The Gunning Fog Index (Gunning, 1952) estimates the years of formal education a person needs to understand a text on the first reading, with a value of 6 corresponding to sixth grade and a value of 17 corresponding to a College graduate. It is defined as $0.4 \cdot \left(\frac{\#words}{\#sentences} + 100 \frac{\#complex\ words}{\#words} \right)$, where complex words are those with three or more syllables, excluding proper nouns or compound words and common suffixes (e.g., -es, -ed, -ing). In our analysis, we flip this measure’s sign so that larger numbers correspond to more readable text (as is the case for the Flesch Reading Ease measure).

of words indicating uncertainty in the ESG portion of the PIS, e.g., some funds write “we *may* consider ESG factors in our investment strategy.”¹⁰

ESG uniqueness. We construct a measure that captures the text uniqueness of a fund’s ESG description. Similar to Kostovetsky and Warner (2020), after we remove from all PISs generic stop words that have no information content (e.g., “and” or “the”),¹¹ (i) we calculate the pairwise unique-word overlap between the ESG portion of a fund’s PIS and that of all other funds in the same year, (ii) we divide this overlap by the pairwise minimum of the unique-word counts to turn it into a ratio in [0, 1], and (iii) we calculate the mean across all pairs and flip its sign to turn it from a *similarity* to a *uniqueness* measure. We repeat this procedure for all funds with ESG keywords in their PIS to obtain a raw uniqueness measure for each, which we then standardize to obtain a measure with mean zero and variance one. Finally, we regress each text’s standardized uniqueness measure on the number of words it contains and use the residuals of this regression as a pure uniqueness measure, which captures uniqueness above and beyond what would be expected for a given text’s length.

To control in our analysis for a fund’s overall writing style, we also calculate the above measures of readability, tonality, and uniqueness for the entire PIS as well as for the non-ESG portion of the PIS (i.e., all sentences that do not contain ESG keywords) separately. All definitions are similar to the above, except that the uniqueness of each fund’s entire PIS is calculated relative to funds in the same CRSP objective code (rather than the universe of all funds) in the same year in order to account for systematic differences across investment styles in the content of the PIS.¹²

In Table 3 we present summary statistics for our measures of ESG positioning as well as our measures of readability, tonality, and uniqueness of the entire PIS and the ESG portion of the PIS. In Panel A, we see that the entire PIS has a mean (median) length of 404 (347) words, is difficult but readable by college graduates (a Flesch Reading Ease score of 19), and, on average, 2.3%

¹⁰We identify uncertain words using the Loughran-McDonald sentiment word list taken from Bill McDonald’s website at <https://sraf.nd.edu/textual-analysis/resources/#Master%20Dictionary>.

¹¹For a complete list of stop words we remove, see `StopWords_Generic.txt` and `StopWords_genericLong.txt` on Bill McDonald’s website at <https://sraf.nd.edu/textual-analysis/resources/#Master%20Dictionary>.

¹²We do not calculate ESG text uniqueness separately for funds with different investment objectives for two reasons. First, a fund’s “flavor” of ESG investment may depend on its investment objective and if this flavor is unique relative to that of the universe of funds then this is something we would like to capture. Second, the number of funds that include ESG keywords in their PIS is relatively small, so calculating ESG text uniqueness separately for each investment objective would result in more noisy measures.

of its words are vague, 1.1% are positive, and 0.5% are negative. Focusing on prospectuses that contain an ESG keyword, in Panel B, we see that the ESG-related portion of the PIS has a mean (median) length of 107 (65) words, so about a fifth of the entire PIS. The first sentence containing an ESG keyword appears on average about halfway through (more precisely, at the 42% mark), though in 13% of the cases an ESG keyword appears in the very first sentence. Furthermore, we see that the ESG portion of the PIS has substantially lower readability than the entire PIS (Flesch Reading Ease score of -16 vs. 19 for the entire PIS). Finally, 1.6% of words in the ESG portion of the PIS are uncertain, which is somewhat lower than the 2.3% of uncertain words encountered in the entire PIS. The differences in text readability and tonality between the ESG portion and the entire PIS are both statistically significant at the 1% level.

In Panel A of Table 4, we present summary statistics for our text-based ESG intensity measures. ESG keywords make up 0.04% of the words in the average PIS, and 0.62% of the words in a PIS that contains at least one ESG keyword. The PIS contains at least one ESG keyword in a little over 4% of all fund-months.

Finally, in Panel B of Table 1, we present summary statistics for the characteristics of the funds whose PIS contains at least one ESG keyword. Comparing the mean characteristics of the subsample of funds whose PIS contains at least one ESG keyword (in Panel B) with those of the entire sample of funds (in Panel A), we see that the former are younger (11.5 vs. 13.3 years), smaller (632 vs. 1,604 million U.S. dollars in total net asset value), have a lower turnover ratio (57% vs. 74%), higher fund inflows (0.67% vs. 0.39% per month), and higher net returns (1.11% vs. 0.78% per month), with all differences statistically significant at the 1% level. On the other hand, the expense ratio, marketing expenses, and the 4-factor alphas are quite similar, and the proportion of funds targeted to institutional investors is 50% for both samples.

3.2 Fundamentals-based ESG measures

Holdings-based ESG score. Our primary measure of the ESG intensity of fund fundamentals is based on the ESG scores of the fund's most recent portfolio holdings. This holdings-based ESG measure is defined as the value-weighted average of the stock-level ESG scores of the stocks

included in a fund's portfolio. Formally, we define the measure for fund i in month t as

$$ESG_{i,t}^H := \frac{\sum_{s \in S_t} V_{i,s,t} ESG_{s,t}}{\sum_{s \in S_t} |V_{i,s,t}|}, \quad (2)$$

where $V_{i,s,t}$ is the value of fund i 's most-recently reported holdings in stock s , $ESG_{s,t}$ is the ESG score for stock s in month t , and S_t is the universe of stocks for which we have an ESG score in month t .¹³ As noted previously, in our baseline calculation we use stock-level ESG scores from MSCI but, for robustness, we also use a combined stock-level ESG score that averages the standardized z -scores from MSCI, Sustainalytics, and Refinitiv. Furthermore, we repeat the same calculation replacing stock-level ESG scores with stock-level carbon emissions levels. Finally, we use the fund-level globe ratings which are calculated according to Morningstar's methodology using Sustainalytics' stock-level ESG and controversy scores.

Returns-based ESG score. We also calculate a returns-based ESG measure using an approach inspired by the returns-based style analysis of Sharpe (1992). While the holdings-based measure discussed above has the advantage that it provides a more direct measure of a fund's level of ESG investment, the returns-based measure we discuss here has the advantage that it is immune to potential window dressing in fund holdings and that it is based on a fund's entire portfolio rather than just the part for which we have stock-level ESG scores.

To calculate the returns-based ESG measure, we start by estimating for each fund the asset-class factor model

$$r_{i,d} = \alpha_i + \mathbf{F}'_d \mathbf{w}_i + \varepsilon_{i,d}, \quad (3)$$

where $r_{i,d}$ is the net rate of return of fund i in day d , \mathbf{F}_d is a vector of returns of different asset classes on day d , \mathbf{w}_i is the vector of fund-specific style-weights that are positive and required to sum to 1, and $\varepsilon_{i,d}$ is the residual component that represents the return attributable to selection rather than style. The model we use contains several asset classes, similar to Sharpe (1992): U.S. stocks, emerging markets stocks, developed markets stocks, government bonds, corporate bonds,

¹³Note that $V_{i,s,t}$ is positive for long and negative for short positions. We divide by the sum of the absolute value of the holdings in each stock, because otherwise the measure is not well defined (e.g., if the total value of longs equals that of shorts). In any case, short positions are very rare in the funds we study, with about 0.5% (0.2% by value) of all positions being shorts.

and mortgage-backed securities. U.S. stocks in particular are partitioned into quintiles based on their MSCI ESG scores; specifically, the return of each ESG quintile is the market-capitalization value-weighted return of the stocks allocated to this quintile.

Subsequently, for each fund, we calculate the time-varying portfolio weight of each asset class on a monthly rolling basis using 12 months of daily returns. For each fund-month, we then use the style weights of the ESG quintiles to calculate a value-weighted measure of ESG intensity. That is, the ESG measure is $ESG_{i,t}^R = \sum_{q=1}^5 q \cdot w_{i,t}^q$, where $w_{i,t}^q$ is the style weight of the q^{th} ESG quintile for fund i in month t , with 1 (5) corresponding to stocks with the lowest (highest) ESG scores.

In Table 4, we present summary statistics for all our fundamentals-based ESG intensity measures (Panel B) and correlations between the text-based and fundamentals-based ESG measures (Panel C). We see that the average fund-level ESG measure based on MSCI is 4.72, which is near the midpoint of the [0, 10] scale that MSCI uses for firm-level ESG scores; this is also the case for our other two fundamentals-based ESG measures. Looking at the standard deviations and the percentiles across all three fundamentals-based ESG measures, we see that their spread is quite narrow, and certainly much narrower than that of the firm-level ESG scores. Furthermore, all correlations are positive and statistically significant at the 1% level. The correlation of our text-based ESG measure with the fundamentals-bases measures is somewhat low at 0.08, while the fundamentals-based measures are more highly correlated with each other, with a correlation of 0.45 for the two holdings-based measures and a correlation of 0.30 between the holdings-based and returns-based measures.

3.3 Measures of greenwashing

Finally, we use our text-based and fundamentals-based ESG measures to devise a measure of greenwashing for each fund in each month ($GW_{i,t}$) that captures whether funds overstate in their PIS their level of actual ESG investing. Specifically, our baseline measure of greenwashing is an indicator variable that indicates funds that talk about ESG in their PIS—placing themselves in the extreme top 5% of funds in terms of what they say about ESG—but the ESG measure based on their most-recently reported holdings is below the median across all funds in the same CRSP

investment category in the same month.¹⁴

In setting a high minimum discrepancy necessary to be classified as a greenwasher—top 5% versus bottom 50%—we essentially focus on the more egregious greenwashers. In robustness checks, we set the discrepancy threshold even higher but the trade-off is that doing so restricts the sample of greenwashers and so lowers the statistical power of our analysis. Specifically, we classify as greenwashers funds that talk about ESG in their prospectus (i.e., they are in the top 5% in terms of what they say) but (i) their holdings-based ESG score is in the bottom quartile (i.e., they are in the bottom 25% in terms of what their holdings indicate) or (ii) both their holdings-based ESG score and their returns-based score based on the returns of the subsequent 12 months is below the median (i.e., they are in the bottom 50% in terms of what their most-recently disclosed portfolio and their subsequent returns indicate). In all these greenwashing measures, we use our baseline holdings-based ESG measure calculated from stock-level ESG scores from MSCI. In a further robustness check, we also construct a greenwashing measure that uses the holdings-based measure calculated from the combined stock-level ESG scores from multiple data providers (MSCI, Sustainalytics, and Refinitiv). Results are similar in all cases.

In Panel D of Table 4, we present summary statistics for our greenwashing measures. Our baseline measure indicates that, across the years, on average 1.5% of all funds greenwash. More specifically, Figure 5b shows that the proportion of greenwashers starts at about 1.2% in 2011, drops to a low of about 0.8% in 2015, and then rises almost exponentially year-by-year until it reaches a high of 5.6% in 2020.

4 Results

4.1 Do flows respond to text-based or fundamentals-based ESG scores?

We start by studying how investors respond to a fund’s ESG commitment—stated or actual. Do investors pay attention to information contained in fund prospectuses? Or do they pay attention

¹⁴To avoid misclassifying as greenwashers funds that have just started making ESG investments, updated their PIS to include ESG mentions, but have not disclosed their updated portfolio holdings yet, our greenwashing dummy is not defined for all months after the PIS change until the fund discloses its new portfolio holdings.

to fund fundamentals such as information revealed by portfolio holdings?

Our baseline specification is

$$Flows_{i,t} = \alpha_{s,t} + \beta_T ESG_{i,t-1}^T + \beta_H ESG_{i,t-1}^H + \beta' \mathbf{x}_{i,t-1} + \varepsilon_{i,t}, \quad (4)$$

where $Flows_{i,t}$ are the flows of fund i in month t defined in Equation 1, $ESG_{i,t-1}^T$ is the fund's text-based ESG intensity score in month $t - 1$ based on its most recent prospectus (see Section 3.1), and $ESG_{i,t-1}^H$ is the fund's holdings-based ESG score in month $t - 1$ based on its most recently disclosed portfolio (see Equation 2). $\alpha_{s,t}$ are category-by-month fixed effects (with fund category defined using the CRSP objective code) that control for unobserved time-varying heterogeneity at the category level. $\mathbf{x}_{i,t-1}$ is a vector of fund characteristics that might affect fund flows, such as age, size, expense ratio, marketing expenses, and past performance. Standard errors are heteroskedasticity-consistent and clustered by fund and year-month.

4.1.1 Baseline

In Table 5, we report the results from estimating Equation 4. Across all specifications, our text-based ESG measures have a consistent strong positive effect on fund inflows, which is robust to controlling for a wide variety of fundamentals-based ESG measures, past performance, and other fund characteristics.

Specifically, in Columns 1–3, the text-based ESG measure we use is the frequency of ESG keywords in the PIS; in Column 4 we use the relative length of the ESG-related portion of the PIS; and in Columns 5–7 we replace the continuous ESG measure with a dummy variable indicating if the PIS contains at least one ESG keyword or a dummy variable indicating if the frequency of ESG keywords is high (above median across funds whose PIS contains at least one ESG keyword).¹⁵ We control for the ESG intensity of fund holdings using the raw ESG score (Columns 2 and 4), the percentile rank of the ESG score within the fund's investment category-month (Column 3), and dummies indicating if the fund's ESG score is high (above median) or exceptionally high (above

¹⁵We obtain similar results if we use variations of these measures, such as the frequency of unique (rather than all) ESG keywords or a simple count (rather than the frequency) of ESG keywords.

the 90th percentile) across funds in the same investment category-month (Columns 5–7).¹⁶

We find that investors respond strongly to text-based ESG measures. The coefficients on our text-based ESG measures (β_T) are economically large and statistically significant at the 1% level (t -statistics above 3.0) across all specifications. Specifically, since the mean (standard deviation) of ESG-keyword frequency among funds that talk about ESG investments in their PIS is about 1% (1%), a coefficient of about 0.3 on ESG-keyword frequency in Columns 2–3 suggests that (i) funds that talk about ESG attract additional inflows of about 0.3% of their net asset value per month (i.e., 3.6% per year)—translating to an additional 4.8 (0.7) million U.S. dollars per month for the mean (median) fund—and (ii) funds with ESG-keyword frequency a standard deviation above the mean get another 0.3% additional inflows per month. The estimated effect of the ESG text’s relative length in Column 4 is consistent with these results. Also consistent with these results are the findings in Columns 5–6, where we see that funds that have at least one ESG keyword in their PIS attract about 0.3%–0.4% higher monthly flows than funds that do not talk about ESG at all, essentially doubling their inflows as the mean monthly flows are about 0.39% (see Table 1A). In Column 7, we see that monthly flows are 0.6% higher for funds with above-median ESG-keyword frequency than for those with low or zero ESG-keyword frequency. Thus, our results indicate that investors direct their flows to funds that discuss, and even more so to those that extensively discuss, their ESG strategy in their PIS.

In Table 6, we repeat our analysis using alternative measures for fundamentals-based ESG intensity. In Column 1, we use the Morningstar globe ratings that have previously been shown to have a significant positive effect on fund inflows (see Hartzmark and Sussman, 2019); in Column 2 we use a holdings-based measure calculated from multiple data providers of stock-level ESG scores; in Column 3 we use a holdings-based measure calculated from stock-level carbon emissions; and in Column 4 we use a returns-based measure calculated from a “style” analysis of fund returns over the preceding 12 months (see Section 3.2). The estimate of the coefficient on ESG-keyword frequency remains highly significant in all these specifications.

Our analysis also shows that fund flows significantly react to the ESG intensity of fund holdings, but only for exceptionally high levels of holdings-based ESG measures (see Columns 6–7 of Table

¹⁶The ESG percentile ranks are defined for months in which there are at least 100 funds in the fund’s CRSP objective code.

5 and Column 1 of Table 6), and the effect is economically and statistically less significant than that of text-based ESG measures. Specifically, we estimate that extreme levels of holdings-based ESG are associated with additional inflows of about 0.2% per month, with coefficient t -statistics between 2.6 and 2.8. Our finding that extreme holdings-based ESG measures matter is consistent with that of Hartzmark and Sussman (2019), who show that funds with a five-globe Morningstar rating experience greater inflows in the eleven months after the introduction of the globe ratings (from March 2016 to January 2017). But our results are also consistent with those of Gantchev, Giannetti and Li (2021), who show that this effect is smaller when the sample is extended to December 2017. In Panel A of Table 7, we repeat our analysis for the period after the publication of the Morningstar globe ratings (March 2016 onward), which is when the media and investors started paying more attention to ESG investing. We see that, in more recent years, the text-based ESG measures have become even more important, while the holdings-based measures are weaker and lose their significance.

Overall, our finding that text-based ESG measures matter and that they matter more than fundamentals-based ESG measures is consistent with the idea that investors have limited attention and pay attention to more salient information (see Barber, Odean and Zheng, 2005; Hartzmark and Sussman, 2019; Kostovetsky and Warner, 2020 for related evidence on mutual fund investing and the survey by Barber and Odean, 2013 for related evidence on other investment decisions).

4.1.2 Fund fixed-effects

Next, we repeat our analysis including fund fixed effects in addition to the category-by-month fixed effects. This analysis has the advantage that it accounts for unobserved fund heterogeneity, but the disadvantage that it is much less efficient as it identifies the effect of text-based ESG measures on fund flows through within-fund time-series variation only, which is rather limited. To be precise, about 6% of funds have switched from none to at least one ESG mention in their PIS during our 10-year sample period. Despite this shortcoming, it is very encouraging that this analysis of the effects of text-based ESG measures on fund flows yields results that are very similar to those from our baseline analysis that exploits cross-sectional variation within category-month.

In Panel B of Table 7, we present our results from this analysis. We see that the effect of ESG-keyword frequency is close to 0.3 both with and without fund fixed effects (see the first row in Tables 7B and 5, respectively), though, as expected, the statistical significance is lower with than without fund fixed effects (t -statistic of 1.9 vs. 4.3, respectively). Note also that the coefficient of interest in the specification that uses a dummy to identify funds with above-median ESG-keyword frequency remains highly significant (see Column 4, Table 7B). On the other hand, the effects of holdings-based ESG measures do not seem to carry over. For example, the dummy indicating exceptionally high levels of holdings-based ESG intensity is now smaller in magnitude (0.001 with vs. 0.002 without fund fixed effects) and not statistically significant (t -statistic of 0.7 vs. about 2.8).

4.1.3 Institutional versus retail investors

Next, we study whether the effect of the ESG content of the prospectus on fund flows is different for funds targeted to institutional versus retail investors. For example, this effect may be stronger for institutional funds because institutional investors might be better at parsing the information in fund prospectuses, or it may be weaker for these funds because institutional investors may focus more on fund fundamentals, potentially rendering (some) information in the prospectus superfluous.

We classify a fund as institutional or retail using CRSP's institutional class indicator for its largest (by net asset value) share class. Using this approach, about half of the funds each year are classified in each category. In Table 8, we replicate the analysis shown in Table 5, separately for institutional funds (in Panel A) and for retail funds (in Panel B). We see that our text-based ESG measures significantly affect the flows to both types of funds, but the estimated effect is larger for the institutional funds. A one standard deviation increase in the frequency of ESG mentions leads to higher monthly flows of about 0.35% for institutional funds and of 0.2% for retail funds (both effects are significant at the 1% level).¹⁷ As both types of investors significantly respond to the ESG information contained in fund prospectuses, we conclude that institutions are not the sole

¹⁷We note that, while the estimated effect for institutional funds is larger than for retail funds, the difference is not statistically significant. Using observations for both types of funds to estimate a model that includes a term that interacts ESG measures (both text- and holdings-based) with an institutional-fund dummy, we find that the t -statistic on the interaction terms is around 1 in all specifications. Results from this estimation are reported in the Internet Appendix.

driver of the effect of text-based ESG measures on fund flows. Furthermore, we find that, while flows to institutional funds significantly respond to exceptionally high levels of holdings-based ESG, this is not the case for retail funds.

Our finding that institutional investors respond more strongly than retail investors to both text- and holdings-based ESG measures could be explained by the idea that the former are more likely to be sophisticated and pay attention to (and process) information than the latter. Consistent with this idea, Del Guercio and Tkac (2002), Evans and Fahlenbrach (2012), and Salganik-Shoshan (2016) all show that institutional investors chase more sophisticated performance measures such as multi-factor alphas while retail investors react to simple return measures like past raw returns. An alternative explanation could be that retail investors are *more* sophisticated than institutional investors and avoid investing in ESG funds because they underperform; but this is not consistent with our findings from the performance analysis (see Section 4.2.3 below), where we find that ESG funds in fact overperform. Yet another possible explanation is that retail investors do not care (as much) about ESG investments. While Bauer, Ruof and Smeets (2021) find that the majority of retail investors in their study do care about ESG investments even if they may be detrimental to fund performance, institutional investors may indeed care even more, as they sometimes face constraints and/or pressure to make such investments.

4.1.4 Text style

In this section, we examine whether fund flows respond to the writing style of the ESG portion of the PIS. Does the readability, tonality, uniqueness, and positioning of the ESG text in the PIS affect flows? In Table 9, we present the results of this analysis. We find that (i) the readability of the ESG text has a positive and significant effect on fund flows, (ii) the uncertainty in the tone of the ESG text has a marginally significant negative effect on flows, and (iii) the uniqueness and positioning of the ESG text do not significantly affect fund flows. Throughout this analysis, we control for the writing style (readability, tonality, and uniqueness) of the entire PIS or the non-ESG portion of the PIS.

In Columns 1–2, we see that the readability of the passage containing ESG keywords has a strong positive effect on fund flows. For example, an increase of 1 in the ESG passage's (sign-

flipped) Gunning Fog index—corresponding to a text that requires one less year of education to be easily understood—is associated with additional inflows of 0.1% of a fund’s total net asset value per month. Using the Flesch Reading Ease measure of text readability yields very similar and statistically very significant results (both t -statistics around 7). This result reinforces our earlier findings that are consistent with the idea that investors have limited information processing capacity and are thus more likely to react to financial disclosure information that is easier to process. Furthermore, in Column 3, we see that a one standard deviation increase in the frequency of uncertain words in the ESG passage leads to flows that are lower by about 0.1% per month, with the effect being marginally significant at the 10% level (t -statistic of 1.76).

On the other hand, in Column 4, we see that the uniqueness of the ESG passage—as compared to the ESG passage in the PIS of other funds’ prospectuses filed in the same year—has no effect on fund flows. This could be because an ESG focus is sufficient to differentiate a fund’s investment strategy, hence draw investors’ attention, in the first place. Also, in Columns 5–6 we see that the positioning of the ESG mentions (measured as the proportion of text until the first ESG-related sentence or as a dummy indicating whether an ESG keyword appears in the first sentence) has no effect on fund flows. This is also not surprising given that the PIS section of the prospectus is quite small at an average of 400 words, so investors who read it are likely to read its entirety.

We note that, in the analysis discussed above, we also control for the text features of the entire PIS, and we report the estimated coefficients in the Internet Appendix. We find that higher readability and more positive tonality of the entire PIS have a weakly positive effect on fund flows. These results are consistent with those of Hillert, Niessen-Ruenzi and Ruenzi (2021), who find that investors react positively to a more positive tone and a personal writing style in the shareholder letters they receive from funds.

4.2 Greenwashing

In this section, we study how investors react to greenwashing, i.e., the practice of overstating a fund’s level of ESG investing. That is, do they recognize the discrepancy between words and actions, or do they respond by rewarding greenwashing funds with additional capital? The latter

would not only indicate investors' inattention or inability to process the available information but would also imply that funds have an incentive to engage in greenwashing. Indeed, we find that the increase in flows is similar across greenwashing and non-greenwashing funds, suggesting that investors are irrationally affected by cosmetic ESG talk. Subsequently, we study which funds engage in greenwashing and thus reap the benefits of investors' irrationality.

4.2.1 Do investor flows respond to greenwashing?

We modify our baseline fund flows regression in Equation 4 so that we estimate the effect of talking about ESG in the prospectus separately for greenwashers and for non-greenwashers. That is, we estimate the model

$$Flows_{i,t} = \alpha_{s,t} + \beta_T^{GW} ESG_{i,t-1}^T \times GW_{i,t-1} + \beta_T^{NGW} ESG_{i,t-1}^T \times (1 - GW_{i,t-1}) + \beta_H ESG_{i,t-1}^H + \beta' \mathbf{x}_{i,t-1} + \varepsilon_{i,t}, \quad (5)$$

where we interact our text-based ESG measure $ESG_{i,t-1}^T$ with the variables $GW_{i,t-1}$ and $1 - GW_{i,t-1}$ indicating whether fund i is greenwashing or not in month $t - 1$, hence β_T^{GW} and β_T^{NGW} are the effects on flows of talking about ESG in the PIS deceptively and truthfully, respectively. If investors are rational, i.e., they pay attention to and correctly process the information available to them, then β_T^{GW} should not be statistically different from zero.¹⁸

In Table 10, we present the results from estimating Equation 5 for three alternative greenwashing dummies indicating funds that talk about ESG in their PIS but (i) their holdings-based MSCI ESG score is below the median (Column 1), (ii) their holdings-based ESG score is in the bottom quartile based on MSCI (Column 2) or a combined ESG database (Column 4), and (iii) both their holdings- and returns-based MSCI ESG score are below the median (Column 3). Our results show that, contrary to what rationality would imply, investors direct their flows to funds that talk about

¹⁸We note that there may be a lag between when a fund changes its PIS to include ESG keywords and when these changes are reflected in its investment strategy hence its holdings. Nonetheless, this lag should be small, as the PIS is meant to describe the fund's actual investment strategy rather than its long-term aspirations. To account for this possible lag, and for the possibility that investors may invest in a fund in anticipation of impending changes in investment strategy that are not already reflected in a fund's holdings at the time of the change in the PIS, in a robustness check we exclude from our analysis the first quarter following a change in a fund's text-based ESG measure. Results from this robustness check are reported in the Internet Appendix and are similar to those reported below.

ESG in their prospectus regardless of whether these funds back their words with actions by actually investing in ESG stocks.¹⁹

Specifically, we see that our estimate of β_T^{GW} is consistently positive and statistically significant, with an average t -statistic of 2.61. We note that, while the coefficient β_T^{GW} on the greenwashing funds appears to be almost twice as high as the coefficient β_T^{NGW} on the non-greenwashing funds, it is estimated using a smaller sample (as fewer funds greenwash) and so has substantially wider confidence intervals. Indeed, while the point estimates differ, their difference is not statistically significant.

This evidence indicates that (at least some) investors do not uncover the discrepancy between what funds say about ESG investing in their prospectus and how much ESG investing they actually do. These results extend to the context of fund prospectuses and ESG investing the finding by Cooper, Dimitrov and Rau (2001), who study fund name changes during the dot com bubble and find that investors irrationally respond to cosmetic changes.

4.2.2 The determinants of greenwashing

Our results above show that flows respond more to the ESG content of fund prospectuses rather than that of fund portfolios, which implies that funds are rewarded for and therefore have an incentive to engage in greenwashing. In this section, we turn our attention to studying what leads funds to greenwash their prospectus and thus reap the benefits of investors' behavior. Our hypothesis is that funds are more likely to greenwash if they have a potential benefit from doing so, a strong incentive to do so, and/or the opportunity to go undetected.

First, greenwashing is beneficial as long as there is a demand for ESG investing, so we would expect that funds are more likely to greenwash in the more recent years, during which investors' interest in ESG investing is higher. Second, the incentive to attract capital flows by overstating a fund's ESG commitment is likely higher when the fund's past capital flows have been low. Finally, the opportunity to greenwash and go undetected exists when oversight is low; following Gil-Bazo

¹⁹To further alleviate concerns that our holdings-based ESG measure may be inaccurate, in the Internet Appendix we also report similar results from an analysis that excludes fund-months for which the portfolio coverage of the stock-level ESG scores that we use to construct our fund-level holdings-based ESG measure is below 50%.

and Ruiz-Verdú (2009) who show that funds with weaker governance charge higher expense ratios, we use a fund's expense ratio as a proxy for managerial oversight. To test these hypotheses, we estimate regressions of the form

$$GW_{i,t} = \alpha_s + \beta_1 PastFlows_{i,t-1} + \beta_2 ExpenseRatio_{i,t-1} + \beta_3 Post2016_t + \gamma' \mathbf{x}_{i,t-1} + \varepsilon_{i,t}, \quad (6)$$

where $GW_{i,t}$ is an indicator variable equal to 1 if fund i starts greenwashing in month t and 0 in all previous months,²⁰ $PastFlows_{i,t-1}$ is the mean fund flow of fund i over the 12-month period ending in month $t - 1$, $ExpenseRatio_{i,t-1}$ is the logarithm of the expense ratio charged by fund i in month $t - 1$ adjusted for amortized loans (see Roussanov, Ruan and Wei, 2021), $Post2016_t$ indicates the period after the publication of the Morningstar ratings in March 2016, and $\mathbf{x}_{i,t-1}$ is a vector of other lagged fund characteristics that are potential determinants of greenwashing such as fund age, size (the logarithm of a fund's total net assets), turnover ratio, marketing expenses (a fund's effective 12b-1 fee as in Roussanov, Ruan and Wei, 2021), skill (the average fund alpha over the past 12 months based on the 4-factor model of Carhart, 1997), and an institutional-fund dummy. We also include fixed effects for funds' investment styles based on the CRSP objective codes.

In Table 11 we report the results from estimating Equation 6 based on four alternative greenwashing dummies designed to check the robustness of our results.²¹ Across specifications, we see that the estimated coefficient on past fund flows (β_1) is significantly negative, the coefficient on fund expense ratio (β_2) is significantly positive, and the coefficient on the post-2016 dummy is significantly positive.²² Taken together, these findings support our hypotheses that, as investors' interest in ESG investing rises, funds that have experienced fund outflows over the preceding year and/or have potentially weaker oversight are more likely to exhibit greenwashing behavior.

²⁰As the decision to greenwash is likely rather sticky, for this analysis we do not use a fund's observations from months after it decides to greenwash for the first time.

²¹All OLS regression results we report in this table are qualitatively identical to those from a logistic regression model.

²²Specifically, a standard deviation (i.e., 0.1) decrease in the mean flows of the past 12 months is associated with a 0.01% increase in a fund's likelihood to start greenwashing in any given month (so a 0.12% increase in the likelihood to do so in any given year); a standard deviation (i.e., 0.5) increase in the log expense ratio is associated with a 0.02% increase in a fund's likelihood to start greenwashing in any given month (so a 0.24% increase in the likelihood to do so in any given year); and, after 2016, a fund is 0.06% more likely to start greenwashing in any given month (so 0.70% more likely to do so in any given year).

4.2.3 Greenwashing and fund performance

In the last part of the paper, we study whether greenwashing is a value-destroying or value-enhancing activity, i.e., whether it is associated with worse or better fund performance. In principle, since a fund that greenwashes inserts ESG keywords in its prospectus without altering its investment strategy, there should be no effect on its subsequent performance. As a result, as long as our measures of greenwashing do a good job of capturing funds that greenwash, we should find that greenwashing has no effect on subsequent performance. Essentially, this estimation serves as a further sanity check for our greenwashing measures.

Specifically, we estimate the model

$$\text{Performance}_{i,t} = \alpha_{s,t} + \gamma_T^{GW} \text{ESG}_{i,t-1}^T \times \text{GW}_{i,t-1} + \gamma_T^{NGW} \text{ESG}_{i,t-1}^T \times (1 - \text{GW}_{i,t-1}) + \gamma_H \text{ESG}_{i,t-1}^H + \gamma' \mathbf{x}_{i,t-1} + \varepsilon_{i,t}, \quad (7)$$

where $\text{Performance}_{i,t}$ is fund i 's performance in month t measured as net returns or as fund alpha based on the Carhart (1997) 4-factor model, and all other variables are as defined previously in Equation 5. If greenwashing has no effect on performance, then γ_T^{GW} should not be statistically different from zero.

In Table 12 we present the results from estimating Equation 7 on fund net returns (Panel A) and fund alpha (Panel B, restricting the sample to domestic equity funds only). We see that the estimated coefficient for greenwashing funds (γ_T^{GW}) has a low t -statistic in almost all specifications, indicating that greenwashing does not have a significant effect on subsequent fund returns or fund alpha.²³ On the other hand, the estimated coefficient for non-greenwashing funds (γ_T^{NGW}) is positive and mostly significant, indicating that funds that truthfully reveal in their PIS their actual engagement with ESG investing have better performance.

²³It could be argued that if, as we have shown above, greenwashing increases subsequent fund inflows, hence increases fund size, then it may in fact decrease subsequent performance, especially when performance is measured as the alpha relative to benchmarks; see for example Pastor, Stambaugh and Taylor (2015). While it is not clear that greenwashing has a sufficiently large effect on fund size to reduce performance (especially during our sample period), in several of our estimations we do estimate a negative, though insignificant, effect of greenwashing on fund alpha.

5 Conclusion

In this paper, we develop a novel measure of a mutual fund's ESG commitments that is based on discretionary information provided by the mutual fund itself. This measure is based on text analysis of the principal investment strategy section of the mutual fund prospectus, a document that is contained in regulatory filings that mutual funds submit to the SEC. We observe that, over time, an increasing number of funds discuss ESG topics in their prospectuses. More specifically, at the end of our sample period in 2020, about ten percent of prospectuses contain ESG-related language. We find that funds that discuss their ESG investment strategy in their prospectus attract higher flows, especially if they use more readable text and a less ambiguous tone. Furthermore, fund flows appear to respond more strongly to text-based measures of ESG intensity than to fundamentals-based measures constructed from disclosed fund holdings or realized fund returns. Comparing our two sets of measures, we find that there is often a discrepancy between the discretionary ESG information provided in fund prospectuses and the objective ESG information contained in fund fundamentals. Moreover, we find that, as investors tend to pay more attention to text-based measures, they do not distinguish between funds that greenwash and those that truthfully reveal in the prospectus their commitment to ESG investing. We also show that greenwashing behavior is more likely to be observed in the last five years, during which investors' interest in ESG is higher, and that it is associated with lower past flows and potentially weaker fund oversight. Providing further support for our greenwashing measures, we find that greenwashing funds do not have better performance, while funds that follow through on their ESG commitments do perform better. Our results provide valuable insights for mutual fund managers wishing to communicate their ESG commitment to potential investors, as well as for asset allocators and regulators who seek to identify funds that do not follow through on their ESG commitments.

References

- Abis, S, and A Lines.** 2022. “Do mutual funds keep their promises?” Unpublished Paper.
- Ammann, M, C Bauer, S Fischer, and P Müller.** 2019. “The impact of the Morningstar Sustainability Rating on mutual fund flows.” *European Financial Management*, 25(3): 520–553.
- Barber, B. M, and T Odean.** 2013. “The behavior of individual investors.” In *Handbook of the economics of finance*, Vol. 2, ed. George M. Constantinides, Milton Harris and Rene M. Stulz, Chapter 22, 1533–1570. Amsterdam, the Netherlands:Elsevier.
- Barber, B. M, T Odean, and L Zheng.** 2005. “Out of sight, out of mind: The effects of expenses on mutual fund flows.” *Journal of Business*, 78(6): 2095–2120.
- Bauer, R, T Ruof, and P Smeets.** 2021. “Get real! Individuals prefer more sustainable investments.” *Review of Financial Studies*, 34(8): 3976–4043.
- Berg, F, J. F Koelbel, and R Rigobon.** 2022. “Aggregate confusion: The divergence of ESG ratings.” Unpublished Paper.
- Berg, F, J. F Kölbel, A Pavlova, and R Rigobon.** 2021. “ESG confusion and stock returns: Tackling the problem of noise.” Unpublished Paper.
- Bolton, P, and M Kacperczyk.** 2021. “Do investors care about carbon risk?” *Journal of Financial Economics*, 142(2): 517–549.
- Carhart, M. M.** 1997. “On persistence in mutual fund performance.” *Journal of Finance*, 52(1): 57–82.
- Chevalier, J, and G Ellison.** 1997. “Risk taking by mutual funds as a response to incentives.” *Journal of Political Economy*, 105(6): 1167–1200.
- Cohen, L, C Malloy, and Q Nguyen.** 2020. “Lazy prices.” *Journal of Finance*, 75(3): 1371–1415.
- Cooper, M. J, H Gulen, and P. R Rau.** 2005. “Changing names with style: Mutual fund name changes and their effects on fund flows.” *Journal of Finance*, 60(6): 2825–2858.
- Cooper, M. J, O Dimitrov, and P. R Rau.** 2001. “A rose.com by any other name.” *Journal of Finance*, 56(6): 2371–2388.
- Del Guercio, D, and P. A Tkac.** 2002. “The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds.” *Journal of Financial and Quantitative Analysis*,

37(4): 523–557.

- Evans, R. B, and R Fahlenbrach.** 2012. “Institutional investors and mutual fund governance: Evidence from retail-institutional fund twins.” *Review of Financial Studies*, 25(12): 3530–3571.
- Flesch, R.** 1948. “A new readability yardstick.” *Journal of Applied Psychology*, 32(3): 221–233.
- Gantchev, N, M Giannetti, and R Li.** 2021. “Sustainability or performance? Ratings and fund managers’ incentives.” Unpublished Paper.
- Gibson Brandon, R, P Krueger, and P. S Schmidt.** 2021. “ESG rating disagreement and stock returns.” *Financial Analysts Journal*, 77(4): 104–127.
- Gibson Brandon, R, S Glossner, P Krueger, P Matos, and T Steffen.** 2021. “Do responsible investors invest responsibly?” Unpublished Paper.
- Gil-Bazo, J, and P Ruiz-Verdú.** 2009. “The relation between price and performance in the mutual fund industry.” *Journal of Finance*, 64(5): 2153–2183.
- Gunning, R.** 1952. *Technique of clear writing*. McGraw-Hill.
- Hartzmark, S. M, and A. B Sussman.** 2019. “Do investors value sustainability? A natural experiment examining ranking and fund flows.” *The Journal of Finance*, 74(6): 2789–2837.
- Hillert, A, A Niessen-Ruenzi, and S Ruenzi.** 2021. “Mutual fund shareholder letters: Flows, performance, and Managerial Behavior.” Unpublished Paper.
- Hoberg, G, and G Phillips.** 2016. “Text-based network industries and endogenous product differentiation.” *Journal of Political Economy*, 124(5): 1423–1465.
- Hong, H, and L Kostovetsky.** 2012. “Red and blue investing: Values and finance.” *Journal of Financial Economics*, 103(1): 1–19.
- Jain, P. C, and J. S Wu.** 2000. “Truth in mutual fund advertising: Evidence on future performance and fund flows.” *Journal of Finance*, 55(2): 937–958.
- Kim, S, and A Yoon.** 2021. “Analyzing active managers’ commitment to ESG: Evidence from United Nations Principles for Responsible Investment.” *Management Science*, forthcoming.
- Kostovetsky, L, and J. B Warner.** 2020. “Measuring innovation and product differentiation: Evidence from mutual funds.” *Journal of Finance*, 75(2): 779–823.
- Krakow, J, and T Schäfer.** 2020. “Mutual funds and risk disclosure: Information content of fund

- prospectuses.” Unpublished Paper.
- Kumar, A, A Niessen-Ruenzi, and O. G Spalt.** 2015. “What’s in a name? Mutual fund flows when managers have foreign-sounding names.” *Review of Financial Studies*, 28(8): 2281–2321.
- Liang, H, L Sun, and M Teo.** 2020. “Greenwashing: Evidence from hedge funds.” *Available at SSRN 3610627*.
- Loughran, T, and B McDonald.** 2016. “Textual analysis in accounting and finance: A survey.” *Journal of Accounting Research*, 54(4): 1187–1230.
- Pastor, L, R. F Stambaugh, and L. A Taylor.** 2015. “Scale and skill in active management.” *Journal of Financial Economics*, 116(1): 23–45.
- Roussanov, N, H Ruan, and Y Wei.** 2021. “Marketing mutual funds.” *Review of Financial Studies*, 34(6): 3045–3094.
- Salganik-Shoshan, G.** 2016. “Investment flows: Retail versus institutional mutual funds.” *Journal of Asset Management*, 17(1): 34–44.
- Sharpe, W. F.** 1992. “Asset allocation: Management style and performance measurement.” *Journal of Portfolio Management*, 18(2): 7–19.
- Sheng, J, N Xu, and L Zheng.** 2022. “Do mutual funds walk the talk? A textual analysis of risk disclosure by mutual funds.” Unpublished Paper.
- Starks, L. T, P Venkat, and Q Zhu.** 2017. “Corporate ESG profiles and investor horizons.” Unpublished Paper.

Table 1: Summary statistics of fund characteristics

In Panel A, we present summary statistics for the entire sample of 398,572 fund-months (4,863 funds at an average of 81.96 months each). In Panel B, we present summary statistics for the sample of 17,325 fund-months (503 funds at an average of 34.44 months each) that have at least one ESG keyword in the Principal Investment Strategy (PIS) text block of their prospectus. Our sample period is from January 2011 to June 2020. Fund age is the number of years since the fund's inception. Total net asset value (TNAV) is measured in millions of U.S. dollars. Expense ratio is defined as total annual management, administrative, and 12b-1 fees and expenses divided by year-end TNAV. Turnover ratio is defined as the minimum of aggregate purchases and sales of securities divided by the average TNAV over the calendar year. Marketing expenses (effective 12b-1 fees) are defined similar to Roussanov, Ruan and Wei (2021) as the combination of 12b-1 fees and front loads. Fund inflows are the ratio of monthly fund flows to the beginning-of-month TNAV, winsorized at 1% and 99%. The institutional-fund dummy classifies each fund as targeted to institutional or retail investors using CRSP's institutional class indicator for its largest (by net asset value) share class. Net returns are the monthly fund returns net of fees, expenses, and transaction costs. Fund alpha is the monthly fund alpha estimated from the Carhart (1997) 4-factor model using daily fund and factor returns.

Panel A: Entire sample

| | # Obs | Mean | Std. Dev. | Percentiles | | |
|--------------------------|---------|----------|-----------|------------------|------------------|------------------|
| | | | | 10 th | 50 th | 90 th |
| Fund age | 398,474 | 13.25 | 12.08 | 1 | 11 | 26 |
| Total net asset value | 395,029 | 1,604.37 | 6,491.72 | 10.40 | 223.70 | 3,112.90 |
| Expense ratio (%) | 386,751 | 1.08 | 0.49 | 0.46 | 1.08 | 1.59 |
| Turnover ratio (%) | 387,228 | 74.25 | 195.74 | 13.43 | 47.28 | 132.04 |
| Marketing expenses (%) | 388,231 | 0.48 | 0.46 | 0.00 | 0.25 | 1.00 |
| Fund inflows (%) | 391,768 | 0.39 | 5.84 | -3.19 | -0.33 | 4.15 |
| Institutional-fund dummy | 398,572 | 0.50 | 0.50 | 0 | 1 | 1 |
| Fund returns (net) (%) | 394,569 | 0.78 | 4.51 | -4.35 | 1.00 | 5.63 |
| Fund alpha (%) | 397,799 | -0.10 | 2.20 | -2.20 | -0.05 | 2.00 |

Panel B: Sample with ESG keyword in prospectus

| | # Obs | Mean | Std. Dev. | Percentiles | | |
|--------------------------|--------|--------|-----------|------------------|------------------|------------------|
| | | | | 10 th | 50 th | 90 th |
| Fund age | 17,325 | 11.48 | 10.73 | 1 | 9 | 24 |
| Total net asset value | 17,181 | 632.07 | 1,701.07 | 6.70 | 100.50 | 1,517.50 |
| Expense ratio (%) | 16,403 | 1.11 | 0.41 | 0.61 | 1.07 | 1.56 |
| Turnover ratio (%) | 16,403 | 56.61 | 53.18 | 14.15 | 42.23 | 113.39 |
| Marketing expenses (%) | 16,598 | 0.45 | 0.46 | 0.00 | 0.25 | 1.00 |
| Fund inflows (%) | 17,024 | 0.67 | 5.69 | -2.83 | -0.07 | 4.56 |
| Institutional-fund dummy | 17,325 | 0.50 | 0.50 | 0 | 1 | 1 |
| Fund returns (net) (%) | 17,101 | 1.11 | 5.36 | -5.08 | 1.42 | 6.91 |
| Fund alpha (%) | 17,290 | -0.07 | 2.24 | -2.47 | -0.05 | 2.31 |

Table 2: Summary statistics for ESG keywords

In this table, we present the ESG keywords we use in the text analysis of the PIS in funds' prospectuses, together with their summary statistics across the 37,399 prospectuses we analyze. In the first row we show statistics for all keywords together, and in subsequent rows we show statistics for each keyword separately, from the most frequent to the least frequent. In the first column, we show the number of instances of the ESG keyword across all prospectuses. In the second column, we show the number of prospectuses in which the keyword appears. In the third and fourth columns, we show the frequency counterparts (expressed as percentages) of the first and second columns, that is, the total number of instances of the ESG keyword divided by the total number of words in the PIS across all prospectuses, and the number of prospectuses containing the keyword divided by the number of all prospectuses. In the fifth column, we show the percentage of words corresponding to the keyword (similar to the third column), but here we divide by the total number of words in the PIS across all prospectuses in which at least one ESG keyword appears.

| ESG keyword | # of instances | # of prospectuses | % of words | % of prospectuses | % of words (ESG only) |
|-------------------------------|-----------------------|--------------------------|-------------------|--------------------------|------------------------------|
| All ESG keywords | 7,725 | 1,949 | 0.0508 | 5.21 | 0.7873 |
| ESG | 3,269 | 835 | 0.0215 | 2.23 | 0.3332 |
| environmental | 2,199 | 1,262 | 0.0145 | 3.37 | 0.2241 |
| ethic | 413 | 309 | 0.0027 | 0.83 | 0.0421 |
| carbon | 351 | 118 | 0.0023 | 0.32 | 0.0358 |
| SRI | 279 | 278 | 0.0018 | 0.74 | 0.0284 |
| responsible investing | 251 | 204 | 0.0017 | 0.55 | 0.0256 |
| human rights | 219 | 185 | 0.0014 | 0.49 | 0.0223 |
| green | 191 | 78 | 0.0013 | 0.21 | 0.0195 |
| climate change | 139 | 118 | 0.0009 | 0.32 | 0.0142 |
| renewable energy | 106 | 64 | 0.0007 | 0.17 | 0.0108 |
| social responsibility | 103 | 91 | 0.0007 | 0.24 | 0.0105 |
| pollution | 63 | 62 | 0.0004 | 0.17 | 0.0064 |
| sustainable business practice | 42 | 9 | 0.0003 | 0.02 | 0.0043 |
| sustainable development goals | 35 | 23 | 0.0002 | 0.06 | 0.0036 |
| biological | 25 | 25 | 0.0002 | 0.07 | 0.0026 |
| clean energy | 21 | 21 | 0.0001 | 0.06 | 0.0021 |
| SDG | 12 | 4 | 0.0001 | 0.01 | 0.0012 |
| toxic | 7 | 7 | 0.0000 | 0.02 | 0.0007 |

Table 3: Summary statistics for style of PIS text block in fund prospectus

Summary statistics for text style characteristics of the entire PIS (Panel A) and of the ESG portion of the PIS, i.e., all sentences containing at least one ESG keyword (Panel B). Word count is the number of words. Text readability is calculated using the Flesch Reading Ease (Flesch, 1948) and the (sign-flipped) Gunning Fog Index (Gunning, 1952) measures, with higher values indicating a passage that is easier to understand. Text uniqueness captures the text's average uniqueness relative to the corresponding text in other funds' prospectuses submitted in the same calendar year (see Section 3.1 for details on the definition). Text tonality is measured using the frequency (expressed as a percent) of positive/negative/uncertain words as defined in the Loughran-McDonald sentiment word list. ESG positioning is measured as the proportion of the text from the beginning of the PIS to the first sentence containing an ESG keyword (*Distance to ESG text*) and as a dummy indicating if an ESG keyword appears in the PIS's first sentence (*ESG in first sentence*). The percentiles presented in Panel B are conditional on fund-months whose PIS contains ESG keywords.

Panel A: Entire PIS

| | # Obs | Mean | Std. Dev. | Percentiles | | |
|---|---------|--------|-----------|------------------|------------------|------------------|
| | | | | 10 th | 50 th | 90 th |
| Total word count | 398,572 | 403.79 | 275.10 | 135 | 347 | 730 |
| Text readability (Flesch) | 398,572 | 19.13 | 9.92 | 7.58 | 19.58 | 30.49 |
| Text readability (Fog) | 398,572 | -22.43 | 2.94 | -25.78 | -22.22 | -19.19 |
| Text uniqueness | 398,558 | 0.00 | 0.99 | -1.17 | 0.09 | 1.13 |
| Text tonality (Uncertain word freq, as %) | 398,572 | 2.25 | 1.08 | 0.91 | 2.23 | 3.62 |
| Text tonality (Positive word freq, as %) | 398,572 | 1.09 | 0.91 | 0.00 | 0.92 | 2.27 |
| Text tonality (Negative word freq, as %) | 398,572 | 0.48 | 0.53 | 0.00 | 0.37 | 1.20 |

Panel B: ESG portion of PIS

| | # Obs | Mean | Std. Dev. | Conditional Percentiles | | |
|---|--------|--------|-----------|-------------------------|------------------|------------------|
| | | | | 10 th | 50 th | 90 th |
| ESG-portion word count | 17,325 | 107.19 | 119.59 | 17 | 65 | 240 |
| ESG positioning (Distance to ESG text) | 17,325 | 0.42 | 0.30 | 0.00 | 0.42 | 0.85 |
| ESG positioning (ESG in first sentence) | 17,325 | 0.13 | 0.33 | 0 | 0 | 1 |
| Text readability (Flesch) | 17,325 | -15.64 | 29.53 | -58.08 | -7.25 | 10.77 |
| Text readability (Fog) | 17,325 | -31.90 | 11.14 | -46.08 | -28.58 | -22.63 |
| Text uniqueness | 17,325 | 0.00 | 0.96 | -1.20 | -0.04 | 1.29 |
| Text tonality (Uncertain word freq, as %) | 17,325 | 1.59 | 1.94 | 0.00 | 0.98 | 4.55 |

Table 4: Summary statistics and correlation coefficients of fund ESG measures

In Panel A (B), we present summary statistics for the text-based (fundamentals-based) ESG measures, in Panel C we present correlations between these measures, and in Panel D we present summary statistics for the greenwashing measures. *ESG-keyword frequency* is the frequency of ESG keywords in the PIS text block. *ESG-text relative length* is the number of words in sentences containing ESG keywords divided by the PIS total word count. *ESG in prospectus* indicates if the PIS contains ESG keywords. *ESG-keyword frequency > p50* indicates if the ESG-keyword frequency exceeds the median conditional on the PIS containing ESG keywords. *Holdings ESG score (MSCI)* is the value-weighted mean of the fund investments' ESG scores from MSCI. *Holdings ESG score (combined)* is similar but it averages the standardized ESG scores across the MSCI, Sustainalytics, and Refinitiv databases. *Returns ESG score* is the fund's returns-based ESG score calculated using the Sharpe (1992) style analysis. *GW MSCI p50* (*GW MSCI p25*) indicates if the fund's PIS contains an ESG keyword and its *Holdings ESG score (MSCI)* is ranked below the median (below the 25th percentile) within its category-month. *GW Returns MSCI p50* indicates if the fund's PIS contains an ESG keyword and its *Holdings ESG score (MSCI)* and *Returns ESG score* are both ranked below the median within its category-month. *GW combined p25* indicates if the fund's PIS contains an ESG keyword and its *Holdings ESG score (combined)* is ranked below the 25th percentile within its category-month. Percentiles presented in Panels A and D are conditional on fund-months whose PIS contains ESG keywords. The number of fund-month observations drops from Panel A to Panel B due to missing portfolio holdings information and from Panel B to Panel D due to missing the holdings-based ESG ranking within a fund's investment category, as we require at least 100 observations within an investment category to calculate a ranking.

Panel A: Text-based ESG measures

| | # Obs | Mean | Std. Dev. | Conditional Percentiles | | |
|---------------------------------|---------|------|-----------|-------------------------|------------------|------------------|
| | | | | 10 th | 50 th | 90 th |
| ESG-keyword frequency (as %) | 398,572 | 0.04 | 0.28 | 0.16 | 0.62 | 2.24 |
| ESG-text relative length (as %) | 398,572 | 0.95 | 5.86 | 4.57 | 15.82 | 49.61 |
| ESG in prospectus | 398,572 | 0.04 | 0.20 | 1 | 1 | 1 |
| ESG-keyword frequency > p50 | 398,572 | 0.02 | 0.15 | 0 | 1 | 1 |

Panel B: Fundamentals-based ESG measures

| | # Obs | Mean | Std. Dev. | Percentiles | | |
|-------------------------------|---------|------|-----------|------------------|------------------|------------------|
| | | | | 10 th | 50 th | 90 th |
| Holdings ESG score (MSCI) | 335,108 | 4.72 | 0.91 | 3.96 | 4.81 | 5.52 |
| Holdings ESG score (combined) | 337,623 | 0.04 | 0.28 | -0.30 | 0.04 | 0.40 |
| Returns ESG score | 377,166 | 2.24 | 0.98 | 0.98 | 2.20 | 3.59 |

Panel C: Correlations

| | (1) | (2) | (3) | (4) |
|-----------------------------------|------|------|------|-----|
| (1) ESG-keyword frequency | 1 | | | |
| (2) Holdings ESG score (MSCI) | 0.08 | 1 | | |
| (3) Holdings ESG score (combined) | 0.08 | 0.45 | 1 | |
| (4) Returns ESG score | 0.07 | 0.31 | 0.30 | 1 |

Panel D: Greenwashing measures

| | # Obs | Mean | Std. Dev. | Conditional Percentiles | | |
|---------------------|---------|-------|-----------|-------------------------|------------------|------------------|
| | | | | 10 th | 50 th | 90 th |
| GW MSCI p50 | 273,137 | 0.015 | 0.121 | 0 | 0 | 1 |
| GW MSCI p25 | 273,137 | 0.008 | 0.087 | 0 | 0 | 1 |
| GW Returns MSCI p50 | 264,326 | 0.007 | 0.085 | 0 | 0 | 1 |
| GW combined p25 | 275,258 | 0.008 | 0.090 | 0 | 0 | 1 |

Table 5: Fund flows and the presence of ESG keywords in the PIS

This table shows how fund flows respond to various definitions of text- and fundamentals-based ESG scores. A fund's text-based ESG score is: the ESG-keyword frequency in its prospectus's PIS text block (columns 1–3); the relative length of the part of the PIS containing ESG keywords (column 4); a dummy indicating if the PIS contains ESG keywords (columns 5–6); a dummy indicating if the ESG-keyword frequency in the PIS exceeds the median conditional on containing ESG keywords (column 7). A fund's fundamentals-based ESG score is: the value-weighted mean of its investments' ESG scores (columns 2 and 4); this score's ranking within the fund's investment category-by-month (column 3); a dummy indicating if this score is in the top 50% within the investment category-by-month (column 5) or in the top 90% (columns 6–7). All specifications include investment category-by-month fixed effects and fund controls for age, size, expense ratio, 12b-1 fees, prior 1-month raw return and 12-month return ranked within investment category-by-month, dummies indicating if prior 12-month α is in the bottom or top 10% for the investment category-by-month, a dummy indicating funds targeted to institutional investors, and the PIS total word count. t -statistics from standard errors clustered two-ways at the fund and year-by-month levels are reported. */**/*** indicate significance at the 10%/5%/1% levels.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|--------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| ESG keyword frequency | 0.208 *** 3.000 | 0.294 *** 4.261 | 0.297 *** 4.322 | | | | |
| ESG text relative length | | | | 0.012 *** 3.570 | | | |
| ESG in prospectus | | | | | 0.004 *** 3.084 | 0.003 *** 2.807 | |
| ESG keyword frequency > p50 | | | | | | | 0.006 *** 4.564 |
| Holdings ESG score (raw) | -0.000 -0.765 | | -0.000 -0.741 | | | | |
| Holdings ESG score (rank) | | -0.000 -0.610 | | | | | |
| Holdings ESG score > p50 | | | | -0.000 -0.753 | | | |
| Holdings ESG score > p90 | | | | | 0.002 *** 2.786 | 0.002 ** 2.527 | |
| log(Fund size) | -0.001 *** -4.899 | -0.000 *** -2.693 | -0.000 *** -2.651 | -0.000 *** -2.699 | -0.000 *** -2.673 | -0.000 *** -2.620 | -0.000 ** -2.591 |
| log(Expense ratio) | -0.459 *** -6.222 | -0.184 ** -2.343 | -0.177 ** -2.283 | -0.184 ** -2.350 | -0.179 ** -2.307 | -0.173 ** -2.247 | -0.173 ** -2.241 |
| log(Effective 12b-1 fee) | 0.311 *** 5.464 | 0.282 *** 4.827 | 0.276 *** 4.755 | 0.282 *** 4.837 | 0.277 *** 4.767 | 0.280 *** 4.817 | 0.276 *** 4.749 |
| Prior 1-month return (raw) | 0.049 *** 3.855 | 0.045 *** 3.106 | 0.041 *** 2.843 | 0.045 *** 3.100 | 0.041 *** 2.832 | 0.041 *** 2.834 | 0.040 *** 2.834 |
| Prior 12-month return (rank) | 0.000 *** 24.920 | 0.000 *** 25.658 | 0.000 *** 25.572 | 0.000 *** 25.642 | 0.000 *** 25.586 | 0.000 *** 25.804 | 0.000 *** 25.824 |
| Prior 12-month α < p10 | -0.005 *** -8.881 | -0.005 *** -7.818 | -0.004 *** -7.806 | -0.005 *** -7.828 | -0.005 *** -7.849 | -0.005 *** -7.892 | -0.005 *** -7.837 |
| Prior 12-month α > p90 | 0.011 *** 13.639 | 0.012 *** 14.281 | 0.012 *** 14.235 | 0.012 *** 14.284 | 0.011 *** 14.171 | 0.011 *** 14.126 | 0.011 *** 14.156 |
| log(Fund age) | -0.011 *** -24.384 | -0.010 *** -20.622 | -0.010 *** -20.594 | -0.010 *** -20.641 | -0.010 *** -20.670 | -0.010 *** -20.722 | -0.010 *** -20.667 |
| Fund for institutions | -0.000 -0.261 | 0.000 0.321 | 0.000 0.214 | 0.000 0.326 | 0.000 0.248 | 0.000 0.278 | 0.000 0.255 |
| log(Prospectus word count) | -0.000 -0.639 | -0.001 *** -2.766 | -0.001 ** -2.584 | -0.001 *** -2.793 | -0.001 *** -2.724 | -0.001 *** -2.804 | -0.001 *** -2.703 |
| Category-by-Time Fixed Effects | Yes |
| # of Observations | 307,170 | 261,269 | 258,212 | 261,269 | 258,212 | 258,212 | 258,212 |
| Adjusted R^2 | 0.084 | 0.081 | 0.080 | 0.081 | 0.080 | 0.080 | 0.080 |

Table 6: Fund flows and the presence of ESG keywords in the PIS – additional specifications

This table shows how fund flows respond to text- and fundamentals-based ESG measures, focusing on alternative definitions for the fundamentals-based ESG score. A fund's fundamentals-based ESG score is defined as follows: in column 1, it is a dummy indicating if its Morningstar globe rating is 1 or 5 globes, with 5 the top rating; in column 2, it is the value-weighted mean of its investments' standardized ESG scores, which are calculated as the mean of the standardized ESG scores across the MSCI, Sustainalytics, and Refinitiv databases; in column 3, it is the value-weighted mean of its investments' carbon emissions; in column 4, it is the returns-based ESG score calculated using the Sharpe (1992) style analysis (see Section 3.2). In all specifications, a fund's text-based ESG measure is the ESG-keyword frequency in its prospectus's PIS text block. All specifications include investment category-by-month fixed effects and fund controls for age, size, expense ratio, 12b-1 fees, prior 1-month raw return and 12-month return ranked within investment category-by-month, dummies indicating if prior 12-month α is in the bottom or top 10% for the investment category-by-month, a dummy indicating funds targeted to institutional investors, and the PIS total word count. t -statistics from standard errors clustered two-ways at the fund and year-by-month levels are reported. */**/* indicate significance at the 10%/5%/1% levels.

| | (1) | (2) | (3) | (4) |
|--------------------------------|--------------------|--------------------|--------------------|--------------------|
| ESG keyword frequency | 0.323 *** 3.769 | 0.294 *** 4.270 | 0.279 *** 4.085 | 0.214 *** 3.089 |
| Morningstar Globe 1 | 0.001 0.511 | | | |
| Morningstar Globe 5 | 0.002 * 1.997 | | | |
| Holdings ESG score (combined) | | -0.002 -1.291 | | |
| Holdings carbon emissions | | | -0.000 -1.136 | |
| Returns ESG score | | | | -0.000 -1.370 |
| Fund Controls | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 137,885 | 262,207 | 258,670 | 307,125 |
| Adjusted R^2 | 0.072 | 0.081 | 0.081 | 0.084 |

Table 7: Fund flows and the presence of ESG keywords in the PIS – 2016/03 onward and fund fixed effects

Like Table 5, this table shows how fund flows respond to text- and fundamentals-based ESG scores, but in Panel A it shows results for the period 2016/03 onward and in Panel B it shows results from a model that includes fund fixed effects. A fund's text-based ESG score is: the ESG-keyword frequency in its prospectus's PIS text block, in columns 1–2; a dummy indicating if the PIS contains ESG keywords, in column 3; a dummy indicating if the ESG-keyword frequency in the PIS exceeds the median conditional on containing ESG keywords, in column 4. A fund's fundamentals-based ESG score is: the value-weighted mean of its investments' ESG scores, in column 1; this score's ranking within the fund's investment category-by-month, in column 2; a dummy indicating if this score is in the top 90% within the investment category-by-month, in column 4. All specifications include investment category-by-month fixed effects and fund controls for age, size, expense ratio, 12b-1 fees, prior 1-month raw return and 12-month return ranked within investment category-by-month, dummies indicating if prior 12-month α is in the bottom or top 10% for the investment category-by-month, a dummy indicating funds targeted to institutional investors, and the PIS total word count. t -statistics from standard errors clustered two-ways at the fund and year-by-month levels are reported. * / ** / *** indicate significance at the 10% / 5% / 1% levels.

| | Panel A: For 2016/03 onward | | | |
|--------------------------------|-----------------------------|--------------------|--------------------|--------------------|
| | (1) | (2) | (3) | (4) |
| ESG keyword frequency | 0.345 *** 4.045 | 0.355 *** 4.182 | | |
| ESG in prospectus | | | 0.005 *** 3.733 | |
| ESG keyword frequency > p50 | | | | 0.007 *** 4.831 |
| Holdings ESG score (raw) | 0.000 0.047 | | | |
| Holdings ESG score (rank) | | -0.000 -0.855 | | |
| Holdings ESG score > p90 | | | 0.001 1.550 | 0.001 1.332 |
| Fund Controls | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 144,354 | 144,345 | 144,345 | 144,345 |
| Adjusted R^2 | 0.068 | 0.068 | 0.068 | 0.069 |

| | Panel B: With fund fixed effects | | | |
|--------------------------------|----------------------------------|------------------|----------------|--------------------|
| | (1) | (2) | (3) | (4) |
| ESG keyword frequency | 0.289 * 1.873 | 0.296 * 1.913 | | |
| ESG in prospectus | | | 0.003 1.449 | |
| ESG keyword frequency > p50 | | | | 0.008 *** 3.158 |
| Holdings ESG score (raw) | -0.000 -0.206 | | | |
| Holdings ESG score (rank) | | -0.000 -0.149 | | |
| Holdings ESG score > p90 | | | 0.001 0.707 | 0.000 0.653 |
| Fund Controls | Yes | Yes | Yes | Yes |
| Fund Fixed Effects | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 261,267 | 258,207 | 258,207 | 258,207 |
| Adjusted R^2 | 0.165 | 0.163 | 0.163 | 0.163 |

Table 8: Fund flows and the presence of ESG keywords in the PIS – institutional vs. retail funds

Like Table 5, this table shows how fund flows respond to text- and fundamentals-based ESG scores, but separately focuses on funds targeted to institutional investors (in Panel A) and to retail investors (in Panel B). A fund's text-based ESG score is: the ESG-keyword frequency in its prospectus's PIS text block, in columns 1–2; a dummy indicating if the PIS contains ESG keywords, in column 3; a dummy indicating if the ESG-keyword frequency in the PIS exceeds the median conditional on containing ESG keywords, in column 4. A fund's fundamentals-based ESG score is: the value-weighted mean of its investments' ESG scores, in column 1; this score's ranking within the fund's investment category-by-month, in column 2; a dummy indicating if this score is in the top 90% within the investment category-by-month, in columns 3–4. All specifications include investment category-by-month fixed effects and fund controls for age, size, expense ratio, 12b-1 fees, prior 1-month raw return and 12-month return ranked within investment category-by-month, dummies indicating if prior 12-month α is in the bottom or top 10% for the investment category-by-month, and the PIS total word count. t -statistics from standard errors clustered two-ways at the fund and year-by-month levels are reported. */**/*** indicate significance at the 10%/5%/1% levels.

| Panel A: Institutional | | | | |
|--------------------------------|--------------------|--------------------|-------------------|--------------------|
| | (1) | (2) | (3) | (4) |
| ESG keyword frequency | 0.345 *** 3.029 | 0.365 *** 3.200 | | |
| ESG in prospectus | | | 0.004 ** 2.569 | |
| ESG keyword frequency > p50 | | | | 0.007 *** 3.444 |
| Holdings ESG score (raw) | 0.000 0.332 | | | |
| Holdings ESG score (rank) | | -0.000 -1.184 | | |
| Holdings ESG score > p90 | | | 0.003 ** 2.309 | 0.002 ** 2.178 |
| Fund Controls | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 130,375 | 128,788 | 128,788 | 128,788 |
| Adjusted R^2 | 0.077 | 0.075 | 0.075 | 0.075 |

| Panel B: Retail | | | | |
|--------------------------------|--------------------|--------------------|----------------|--------------------|
| | (1) | (2) | (3) | (4) |
| ESG keyword frequency | 0.202 *** 2.860 | 0.189 *** 2.663 | | |
| ESG in prospectus | | | 0.001 1.066 | |
| ESG keyword frequency > p50 | | | | 0.004 *** 2.795 |
| Holdings ESG score (raw) | -0.001 -1.480 | | | |
| Holdings ESG score (rank) | | 0.000 0.370 | | |
| Holdings ESG score > p90 | | | 0.001 1.512 | 0.001 1.249 |
| Fund Controls | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 130,894 | 129,424 | 129,424 | 129,424 |
| Adjusted R^2 | 0.097 | 0.097 | 0.097 | 0.097 |

Table 9: Fund flows and the characteristics of the ESG-related text in the PIS

This table shows how fund flows respond to various characteristics of the ESG-related text in the PIS text block of a fund's prospectus. Specifically, the effects of the following ESG text characteristics are presented: readability in columns 1–2; tonality in column 3; uniqueness in column 4; and positioning in columns 5–6. *ESG in PIS* is a dummy indicating if the PIS contains any ESG keywords. *ESG readability* is calculated using the Flesch Reading Ease (Flesch, 1948) and the (sign-flipped) Gunning Fog Index (Gunning, 1952) measures, which are designed to measure how easy a passage in English is to understand. *ESG uncertain word freq* is the frequency of uncertain words as defined in the Loughran-McDonald sentiment word list. *ESG text uniqueness* captures the uniqueness of the ESG-related text in a fund's PIS relative to that of other funds, conditioning on prospectuses submitted in the same calendar year. *Distance to ESG text* is the proportion of the text from the beginning of the PIS to the first sentence containing an ESG keyword. *ESG in first sentence* is a dummy indicating if an ESG keyword appears in the PIS's first sentence. All specifications control for a fund's fundamentals-based ESG score, measured as the value-weighted mean of its investments' ESG scores. All specifications also include investment category-by-month fixed effects and fund controls for age, size, expense ratio, 12b-1 fees, prior 1-month raw return and 12-month return ranked within investment category-by-month, dummies indicating if prior 12-month α is in the bottom or top 10% for the investment category-by-month, a dummy indicating funds targeted to institutional investors, and the PIS total word count, as well as controls for the overall PIS's text style (i.e., readability, tonality, and text uniqueness). *t*-statistics from standard errors clustered two-ways at the fund and year-by-month levels are reported. * / ** / *** indicate significance at the 10%/5%/1% levels.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---------------------------------------|--------------------|--------------------|--------------------|-------------------|----------------|--------------------|
| ESG in PIS | 0.006 *** 4.863 | 0.020 *** 7.205 | 0.005 *** 3.536 | 0.003 *** 2.64 | 0.003 1.551 | 0.003 *** 2.783 |
| <i>ESG Text Readability</i> | | | | | | |
| ESG in PIS * ESG readability (Flesch) | 0.000 *** 6.940 | | | | | |
| ESG in PIS * ESG readability (Fog) | | 0.001 *** 7.383 | | | | |
| <i>ESG Text Tonality</i> | | | | | | |
| ESG in PIS * ESG uncertain word freq | | | -0.104 * -1.759 | | | |
| <i>ESG Text Uniqueness</i> | | | | | | |
| ESG in PIS * ESG text uniqueness | | | | -0.001 -1.228 | | |
| <i>ESG Positioning</i> | | | | | | |
| ESG in PIS * Distance to ESG text | | | | | 0.001 0.238 | |
| ESG in PIS * ESG In first sentence | | | | | | 0.001 0.312 |
| Fund Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Overall PIS style controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| # of Observations | 261,269 | 261,269 | 261,269 | 261,269 | 261,269 | 261,269 |
| Adjusted R^2 | 0.082 | 0.082 | 0.082 | 0.081 | 0.081 | 0.081 |

Table 10: Fund flows and ESG keywords in the PIS – greenwashing vs. non-greenwashing funds

This table shows how fund flows respond to the inclusion of ESG keywords in the PIS text block of a fund's prospectus, for funds that greenwash versus those that do not. The effect for greenwashing funds is shown in the row presenting the interaction of *ESG-keyword frequency* with the dummy indicating greenwashing funds (*GW*), while for non-greenwashing funds it is shown in the row presenting the interaction with the dummy indicating non-greenwashing funds (*non-GW*). The specifications differ in the definition of the greenwashing dummy, so also of the non-greenwashing dummy which is 1 minus the former. In column 1, the greenwashing dummy is 1 for any fund that includes an ESG keyword in its PIS but whose value-weighted mean of its investments' ESG scores from MSCI (i.e., whose holdings-based ESG score) is below the 50th percentile within the fund's investment category for that month. In column 2, the holdings-based ESG score cutoff below which the greenwashing dummy equals 1 changes from the 50th to the 25th percentile. In column 3, the greenwashing dummy is 1 for any fund that includes an ESG keyword in its PIS but whose MSCI holdings-based and returns-based ESG score (calculated using the style analysis of Sharpe, 1992) are both below the 50th percentile within the fund's investment category for that month. In column 4, the greenwashing dummy is defined as in column 2 (i.e., it equals 1 if the holdings-based ESG score is below the 25th percentile), but the fund's holdings-based ESG score is calculated as the fund's investments' standardized ESG scores averaged across multiple databases (MSCI, Sustainalytics, and Refinitiv). All specifications control for a fund's fundamentals-based ESG score, measured as the value-weighted mean of its investments' ESG scores. All specifications also include investment category-by-month fixed effects and fund controls for age, size, expense ratio, 12b-1 fees, prior 1-month raw return and 12-month return ranked within investment category-by-month, dummies indicating if prior 12-month α is in the bottom or top 10% for the investment category-by-month, a dummy indicating funds targeted to institutional investors, and the PIS total word count. *t*-statistics from standard errors clustered two-ways at the fund and year-by-month levels are reported. */**/* indicate significance at the 10%/5%/1% levels.

| | (1) | (2) | (3) | (4) |
|--------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| ESG keyword frequency * GW | 0.468 ** 2.485 | 0.716 ** 2.555 | 0.698 *** 2.891 | 0.795 ** 2.522 |
| ESG keyword frequency * Non-GW | 0.268 *** 3.648 | 0.262 *** 3.755 | 0.235 *** 3.318 | 0.255 *** 3.730 |
| Holdings ESG score (raw) | -0.000 -0.687 | -0.000 -0.658 | -0.000 -0.805 | -0.001 -1.135 |
| log(Fund size) | -0.000 *** -2.666 | -0.000 *** -2.668 | -0.001 *** -4.596 | -0.000 *** -2.850 |
| log(Expense ratio) | -0.182 ** -2.335 | -0.183 ** -2.34 | -0.220 *** -2.842 | -0.222 *** -2.840 |
| log(Effective 12b-1 fee) | 0.278 *** 4.777 | 0.279 *** 4.784 | 0.273 *** 4.676 | 0.302 *** 5.143 |
| Prior 1-month return (raw) | 0.041 *** 2.854 | 0.041 *** 2.856 | 0.043 *** 3.098 | 0.042 *** 2.869 |
| Prior 12-month return (rank) | 0.000 *** 25.617 | 0.000 *** 25.629 | 0.000 *** 25.091 | 0.000 *** 25.497 |
| Prior 12-month $\alpha < p10$ | -0.004 *** -7.814 | -0.004 *** -7.811 | -0.004 *** -7.471 | -0.005 *** -8.069 |
| Prior 12-month $\alpha > p90$ | 0.012 *** 14.255 | 0.012 *** 14.265 | 0.012 *** 14.215 | 0.012 *** 14.301 |
| log(Fund age) | -0.010 *** -20.616 | -0.010 *** -20.618 | -0.009 *** -20.557 | -0.010 *** -20.745 |
| Fund for institutions | 0.000 0.203 | 0.000 0.188 | 0.000 0.182 | 0.000 0.079 |
| log(Prospectus word count) | -0.001 ** -2.606 | -0.001 ** -2.606 | -0.001 ** -2.539 | -0.001 ** -2.344 |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 258,099 | 258,099 | 251,150 | 259,863 |
| Adjusted R^2 | 0.080 | 0.080 | 0.081 | 0.081 |

Table 11: Explaining which funds greenwash

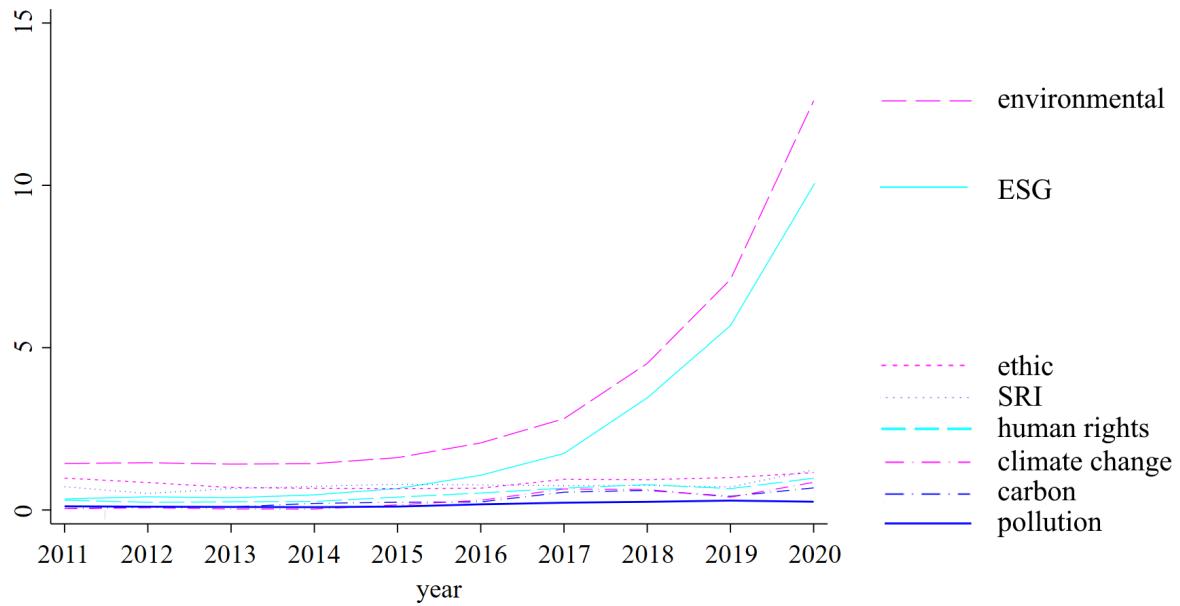
This table shows which fund characteristics are associated with a fund's choice to start greenwashing. The specifications differ in the definition of which funds greenwash. In column 1, a greenwashing fund is one that includes an ESG keyword in its prospectus's PIS text block but whose value-weighted mean of its investments' ESG scores from MSCI (i.e., whose holdings-based ESG score) is below the 50th percentile within the fund's investment category for that month. In column 2, the holdings-based ESG score cutoff below which a fund is deemed to be greenwashing changes from the 50th to the 25th percentile. In column 3, the holdings-based ESG score cutoff is back at the 50th percentile as in column 1, but for a fund to be deemed a greenwasher it is additionally required that the fund's *returns*-based ESG score calculated using the Sharpe (1992) style analysis is below the 50th percentile within the fund's investment category for that month. In column 4, greenwashing is defined as in column 2 (i.e., a fund greenwashes if the holdings-based ESG score is below the 25th percentile), but the fund's holdings-based ESG score is calculated as the fund's investments' standardized ESG scores averaged across the MSCI, Sustainalytics, and Refinitiv databases. All specifications include investment category fixed effects and the explanatory variables fund age, size, expense ratio, turnover ratio, 12b-1 fees, prior 12-month mean flows, and prior 12-month mean α , a dummy indicating funds targeted to institutional investors, and a dummy indicating months after 2016/03. The analyses include, for each fund, all months until the month the fund greenwashes for the first time. The dependent variable (the greenwashing dummy) is expressed as a percent, i.e., 0% or 100%, so the regression coefficients represent the percent change in a fund's likelihood to greenwash in a given month. *t*-statistics from standard errors clustered two-ways at the fund and year-by-month levels are reported. */**/*** indicate significance at the 10%/5%/1% levels.

| | (1) | (2) | (3) | (4) |
|------------------------------|------------|------------|------------|-----------|
| log(Fund size) | -0.007 * | -0.006 *** | -0.009 *** | -0.003 |
| | -1.772 | -2.681 | -3.011 | -1.494 |
| log(Expense ratio) | 0.039 ** | 0.027 ** | 0.030 ** | 0.029 *** |
| | 2.122 | 2.573 | 2.535 | 2.769 |
| Turnover ratio | 0.004 | 0.001 | 0.000 | -0.001 |
| | 0.745 | 0.194 | 0.038 | -0.361 |
| log(Effective 12b-1 fee) | 3.457 | 1.866 | -0.486 | 2.002 |
| | 1.049 | 0.994 | -0.290 | 1.225 |
| Prior 12-month mean flows | -0.100 *** | -0.069 *** | -0.060 ** | -0.052 * |
| | -2.768 | -2.885 | -2.542 | -1.783 |
| Prior 12-month mean α | 0.715 | 0.862 | -0.231 | 0.941 |
| | 0.743 | 1.154 | -0.226 | 1.064 |
| log(Fund age) | 0.001 | 0.001 | 0.015 * | -0.006 |
| | 0.064 | 0.091 | 1.788 | -1.075 |
| Fund for institutions | 0.017 | -0.009 | 0.020 | 0.002 |
| | 0.716 | -0.616 | 1.150 | 0.154 |
| After 2016 | 0.058 *** | 0.017 | 0.045 *** | 0.031 ** |
| | 2.736 | 1.360 | 3.000 | 2.575 |
| Category Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 159,725 | 161,059 | 156,112 | 162,328 |
| Adjusted R^2 | 0.0002 | 0.0001 | 0.0002 | 0.0001 |

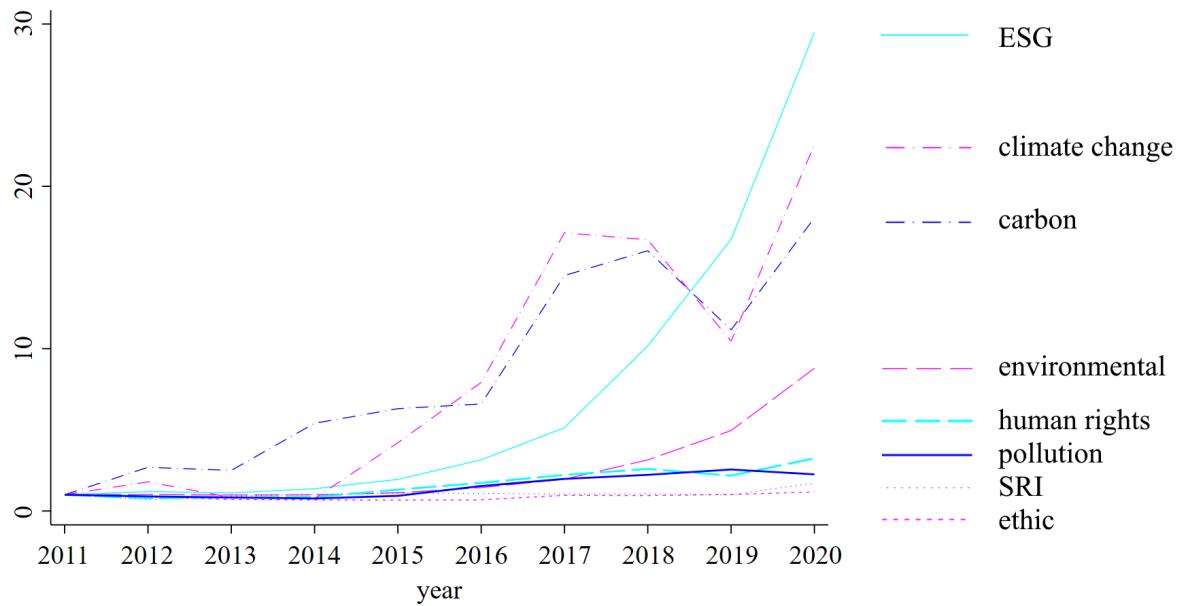
Table 12: Fund performance and ESG keywords in the PIS – greenwashing vs. non-greenwashing

This table shows the effect on fund performance (net returns in Panel A and alpha in Panel B) of the inclusion of ESG keywords in the PIS text block of a fund's prospectus. The effect for greenwashing funds is shown in the row presenting the interaction of *ESG-keyword frequency* with the dummy indicating greenwashing funds (*GW*), while for non-greenwashing funds it is shown in the row presenting the interaction with the dummy indicating non-greenwashing funds (*non-GW*). The specifications differ in the definition of the greenwashing dummy, so also of the non-greenwashing dummy which is 1 minus the former. In column 1, the greenwashing dummy is 1 for any fund that includes an ESG keyword in its prospectus's PIS but whose value-weighted mean of its investments' ESG scores from MSCI (i.e., whose holdings-based ESG score) is below the 50th percentile within the fund's investment category for that month. In column 2, the holdings-based ESG score cutoff below which the greenwashing dummy equals 1 changes from the 50th to the 25th percentile. In column 3, the holdings-based ESG score cutoff is back at the 50th percentile as in column 1, but for the greenwashing dummy to equal 1 it is additionally required that the fund's *returns*-based ESG score calculated using the Sharpe (1992) style analysis is below the 50th percentile within the fund's investment category for that month. In column 4, the greenwashing dummy is defined as in column 2 (i.e., it equals 1 if the holdings-based ESG score is below the 25th percentile), but the fund's holdings-based ESG score is calculated as the fund's investments' standardized ESG scores averaged across the MSCI, Sustainalytics, and Refinitiv databases. All specifications control for a fund's fundamentals-based ESG score, measured as the value-weighted mean of its investments' ESG scores. All specifications also include investment category-by-month fixed effects and fund controls for age, size, expense ratio, 12b-1 fees, prior 1-month raw return and 12-month return ranked within investment category-by-month, dummies indicating if prior 12-month α is in the bottom or top 10% for the investment category-by-month, a dummy indicating funds targeted to institutional investors, and the PIS total word count. *t*-statistics from standard errors clustered two-ways at the fund and year-by-month levels are reported. * / ** / *** indicate significance at the 10%/5%/1% levels.

| | Panel A: Net Returns | | | | Panel B: Alpha | | | |
|--------------------------------|----------------------|------------|------------|------------|----------------|------------|------------|------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| ESG keyword frequency * GW | 0.153 * | 0.088 | 0.023 | 0.109 | 0.016 | -0.001 | 0.044 | -0.016 |
| | 1.960 | 0.926 | 0.202 | 1.128 | 0.355 | -0.017 | 0.589 | -0.289 |
| ESG keyword frequency * Non-GW | 0.050 ** | 0.063 *** | 0.062 *** | 0.073 *** | 0.036 ** | 0.036 ** | 0.026 | 0.041 ** |
| | 2.328 | 2.795 | 2.696 | 3.121 | 2.157 | 2.188 | 1.510 | 2.560 |
| Holdings ESG score (raw) | 0.001 *** | 0.001 *** | 0.001 *** | 0.001 | 0.000 | 0.000 | 0.000 | -0.000 |
| | 3.394 | 3.363 | 3.318 | 0.622 | 1.099 | 1.095 | 0.721 | -0.289 |
| log(Fund size) | 0.000 * | 0.000 * | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 1.709 | 1.694 | 1.326 | 1.546 | 0.944 | 0.944 | 0.521 | 0.977 |
| log(Expense ratio) | -0.064 ** | -0.064 ** | -0.060 ** | -0.076 *** | -0.072 *** | -0.072 *** | -0.069 *** | -0.075 *** |
| | -2.195 | -2.191 | -2.065 | -2.712 | -3.469 | -3.466 | -3.395 | -3.713 |
| log(Effective 12b-1 fee) | 0.008 | 0.008 | 0.007 | 0.010 | -0.010 | -0.010 | -0.009 | -0.010 |
| | 0.600 | 0.606 | 0.492 | 0.704 | -0.799 | -0.804 | -0.731 | -0.809 |
| Prior 1-month return (raw) | 0.022 | 0.022 | 0.024 | 0.022 | 0.029 ** | 0.029 ** | 0.033 ** | 0.028 ** |
| | 0.507 | 0.507 | 0.562 | 0.515 | 2.121 | 2.121 | 2.345 | 2.110 |
| Prior 12-month return (rank) | 0.000 ** | 0.000 ** | 0.000 ** | 0.000 *** | 0.000 | 0.000 | 0.000 | 0.000 |
| | 2.581 | 2.582 | 2.614 | 2.662 | 1.102 | 1.101 | 1.131 | 1.186 |
| Prior 12-month $\alpha < p10$ | 0.000 | 0.000 | 0.001 | 0.000 | -0.000 | -0.000 | -0.000 | -0.000 |
| | 0.541 | 0.542 | 0.771 | 0.627 | -1.070 | -1.071 | -0.873 | -1.028 |
| Prior 12-month $\alpha > p90$ | -0.000 | -0.000 | -0.001 | -0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | -0.711 | -0.706 | -0.871 | -0.748 | 0.481 | 0.481 | 0.607 | 0.434 |
| log(Fund age) | 0.000 | 0.000 | 0.000 | 0.000 | -0.000 | -0.000 | -0.000 | -0.000 |
| | 0.011 | 0.008 | 0.035 | 0.061 | -0.791 | -0.792 | -0.723 | -0.847 |
| Fund for institutions | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 0.931 | 0.934 | 0.691 | 0.784 | 0.672 | 0.677 | 1.029 | 0.512 |
| log(Prospectus word count) | -0.000 *** | -0.000 *** | -0.000 *** | -0.000 *** | 0.000 | 0.000 | 0.000 | 0.000 |
| | -3.014 | -3.010 | -3.141 | -3.072 | 0.150 | 0.150 | 0.230 | 0.006 |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| # of Observations | 258,571 | 258,571 | 251,589 | 260,337 | 185,058 | 185,058 | 180,152 | 186,685 |
| Adjusted R^2 | 0.849 | 0.849 | 0.858 | 0.848 | 0.083 | 0.083 | 0.093 | 0.083 |



(a) Proportion of prospectuses, as a percentage.



(b) Proportion of prospectuses, as a multiple of the year-2011 value.

Figure 1: The evolution of ESG keywords' prevalence in funds' prospectuses over time. This figure plots, for our most common ESG keywords, the proportion of prospectuses in which the keyword appears. In Panel (a), the proportion of each keyword is plotted using raw numbers, expressed as percentages. In Panel (b), the proportion of each keyword is plotted as a multiple of the corresponding year-2011 value. The keywords and the line style representing the evolution of each keyword are shown in the legend to the right of each panel, sorted by prevalence in year 2020.

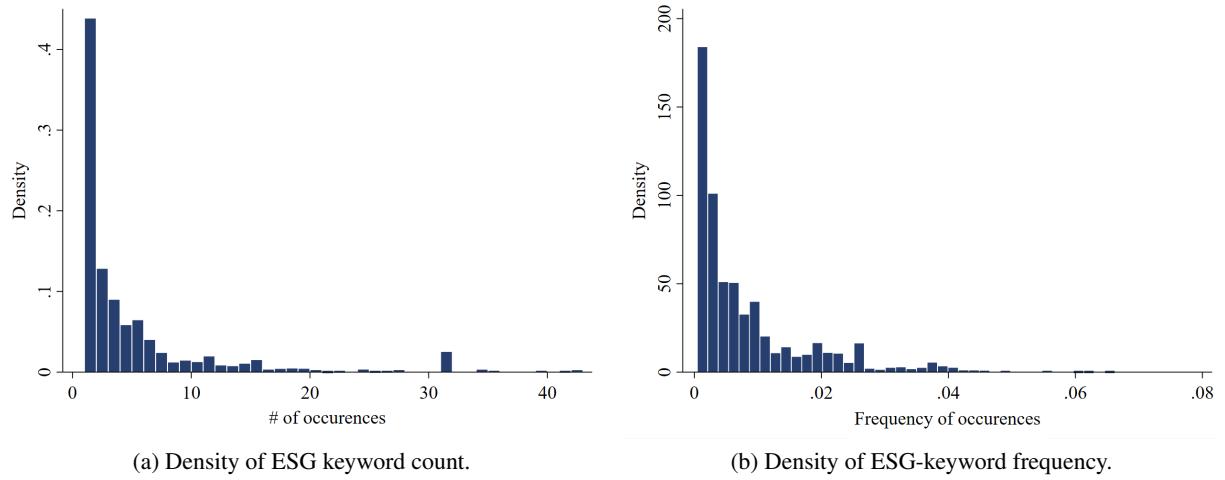


Figure 2: Density of ESG keywords for fund-months during which the filed prospectus contains at least one ESG keyword in the PIS. Panel (a) plots the density of the keyword count, i.e., of the number of ESG keywords in the PIS. Panel (b) plots the density of the keyword frequency, i.e., of the number of ESG keywords divided by the total number of words in the PIS.

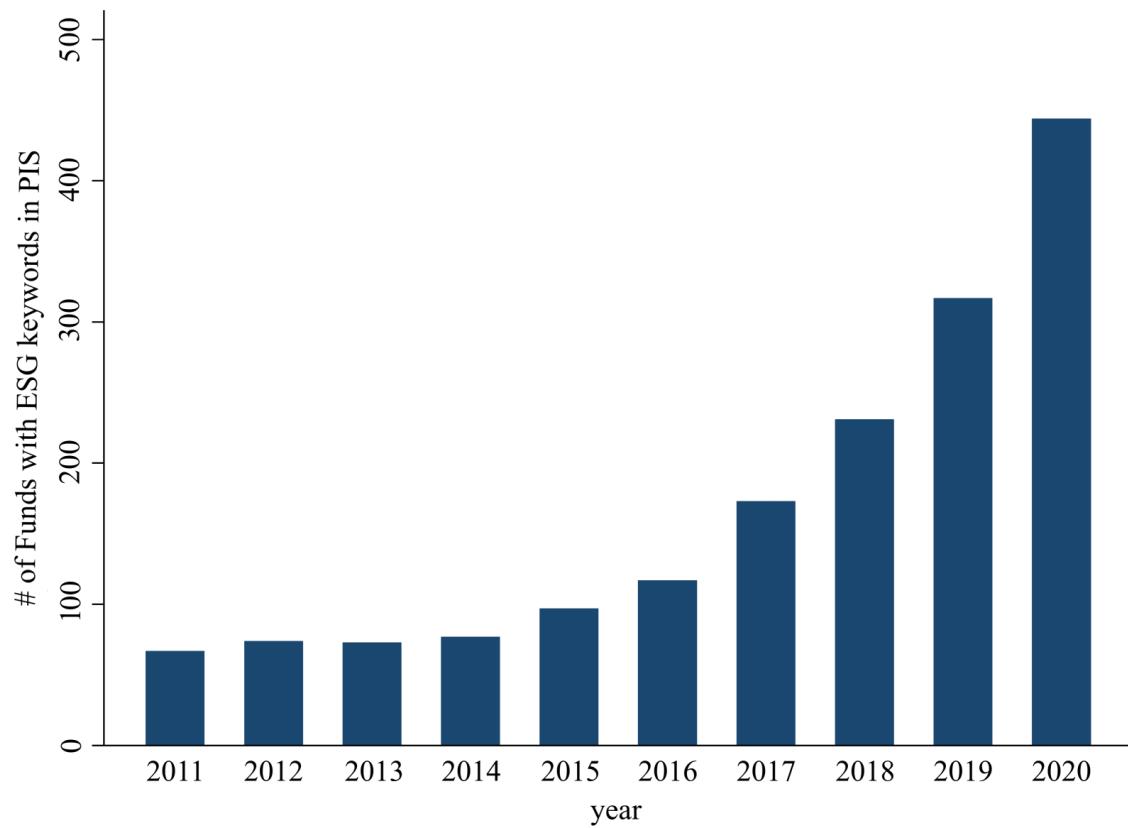


Figure 3: The number of ESG funds, where a fund is classified as ESG if the PIS from the prospectus contains at least one of our ESG keywords.

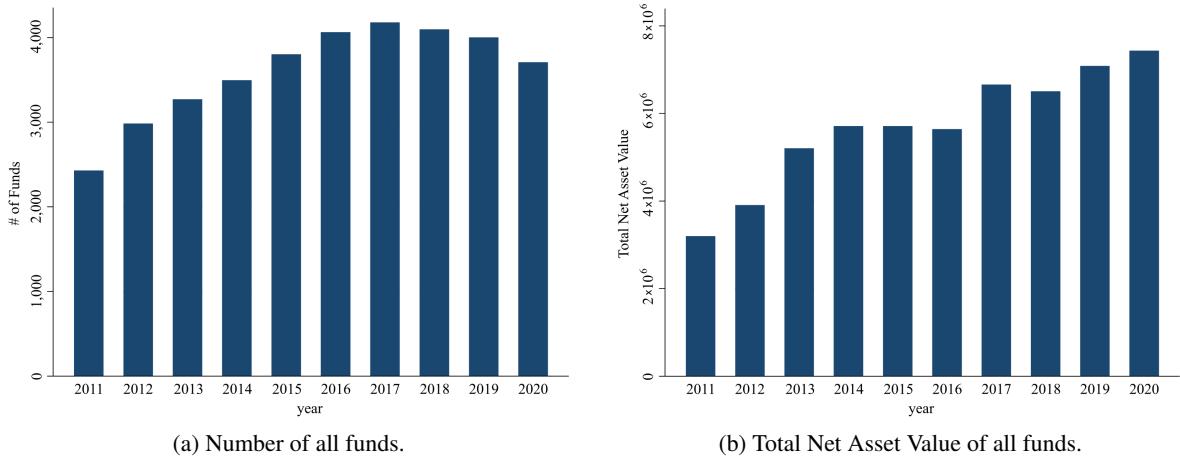


Figure 4: The evolution over time of the number of funds (in panel *a*) and their total net asset value, in million U.S. dollars (in panel *b*).

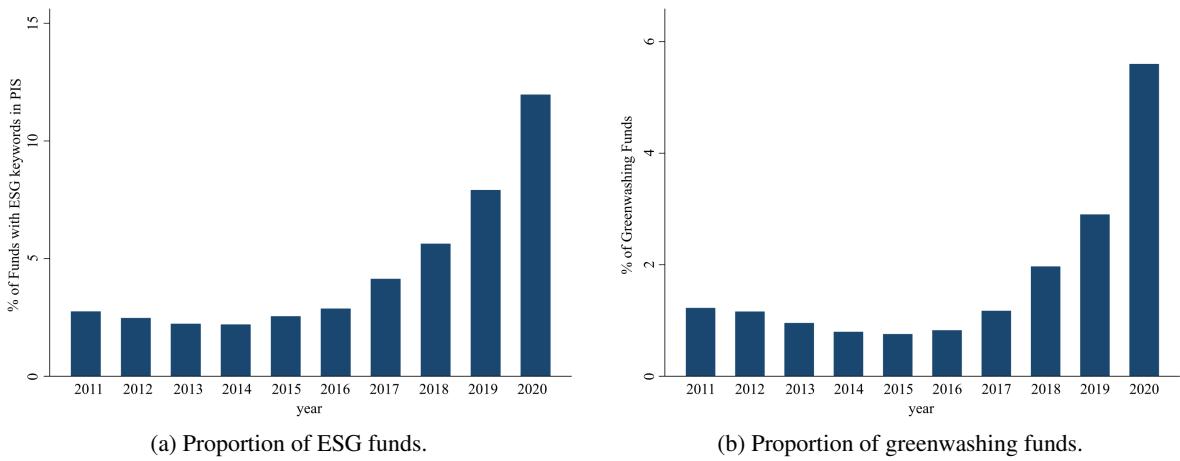


Figure 5: The evolution over time of the proportion (expressed as a percent) of funds with an ESG keyword in the PIS (in panel *a*) and of funds that greenwash (in panel *b*).

Appendix

Sample PIS text blocks

We show examples of funds' PIS text blocks containing ESG keywords, and our text-based measures. ESG keywords are in red italic and sentences containing them in bold typeface.

1. Fund Name: BMO Small-Cap Growth Fund, Prospectus Date: December 2019

The Fund invests at least 80% of its assets in growth-oriented common stocks of small-sized U.S. companies similar in size, at the time of purchase, to those within the Russell 2000 Growth Index. The largest company by market capitalization in the Russell 2000 Growth Index was approximately \$7.0 billion as of October 31, 2019 and the median market capitalization of companies in the Index as of the same period was \$869 million. The Fund may at times focus its investments in one or more sectors.

The Adviser selects stocks using a unique, growth-oriented approach focusing on high quality companies with sustainable earnings growth that are available at reasonable prices, which combines the use of proprietary analytical tools and the qualitative judgments of the investment team. In general, the Adviser believes companies that are undervalued relative to their fundamentals and exhibit improving investor interest outperform the market over full market cycles. As a result, the Adviser's investment process begins by using tools to rank stocks based on expected returns, construct preliminary portfolios with the use of fundamental factors, and manage risk. **The Adviser also integrates *environmental, social, and governance (ESG) considerations* into its security selection, portfolio construction, and monitoring processes.** All purchases and sales of portfolio securities, however, are subjected ultimately to the investment team's qualitative judgments developed from their cumulative investment experience. The entire process is designed to focus on company fundamentals through both quantitative and qualitative analysis to balance return generation with risk management.

From time to time, the Fund maintains a portion of its assets in cash. The Fund may increase its cash holdings in response to market conditions or in the event attractive investment opportunities are not available.

Total word count: 281, ESG-keyword frequency: 1%, ESG-related word frequency: 7%.

2. Fund Name: Putnam Sustainable Leaders Fund, Prospectus Date: March 2018

We invest mainly in common stocks of U.S. companies of any size, with a focus on companies that we believe exhibit a commitment to *sustainable business practices*. Stocks of companies that exhibit a commitment to *sustainable business practices* are typically, but not always, considered to be growth stocks. Growth stocks are stocks of companies whose earnings are expected to grow faster than those of similar firms, and whose business growth and other characteristics may lead to an increase in stock price. **We may consider, among other factors, a company's *sustainable business practices* (as described below), valuation, financial strength, growth potential, competitive position in its**

industry, projected future earnings, cash flows and dividends when deciding whether to buy or sell investments. We may also invest in non-U.S. companies.

Sustainable investing. **We believe that companies that exhibit leadership in *sustainable business practices* also often exhibit more profitable, durable financial returns with lower risk profiles.** Accordingly, in selecting investments, we focus on companies that we believe have a demonstrated commitment to *sustainable business practices*. This commitment may be reflected through *environmental*, social and/or corporate governance (*ESG*) policies, practices or outcomes.

Total word count: 187, ESG-keyword frequency: 9%, ESG-related word frequency: 78%.