# Internet Appendix for "Empty ESG Promises: A Text Analysis of Mutual Fund Prospectuses"

Angie Andrikogiannopoulou, Philipp Krueger, Shema Mitali, and Filippos Papakonstantinou*

*Angie Andrikogiannopoulou and Filippos Papakonstantinou are at King's College London, Philipp Krueger is at the University of Geneva, and Shema Mitali is at Skema Business School.

Here, we present additional information on the machine learning algorithm (in Section IA.A) and tables that provide additional information and support to the paper's results (in Section IA.B).

## IA.A Additional information on the machine learning algorithm

First, we summarize the main steps in implementing the ML algorithm, taking the opportunity to add some details that were omitted from the main text of the paper for the sake of brevity.

1. We pre-process the text of all PISs, i.e., turn all characters to lowercase, remove non-alphabetic characters, stop words, and uninformative words.

2. We use stratified sampling to select a subsample of PISs to be pre-classified, and then experts pre-classify these as E, S, G, and/or ESG or not.

3. We split the pre-classified subsample into a training and a testing subset. We only perform this split once, and we use the same training subset to train the separate ML algorithms that classify PISs as E, S, G, and/or ESG; this is so that the testing subset is untouched by (any variation of) the ML algorithm. To do this training/testing split, we use stratification with respect to the ESG classification. Regardless, the proportions of the other classifications of interest—E, S, and G—are also very similar across the two subsets.

4. We process the PISs using a bag-of-words approach that converts text into a vector of feature weights. When we compile the list of features, we use up to two-word combinations because we have not found higher-order combinations to be useful; exceptions are "environmental, social, and governance" and "socially responsible investing" which are replaced by "esg" and "responsible investing". As we explain below, we also exclude features that are likely to lead to overfitting.

5. To build the random forest, we need to build a number (or forest) of decision trees that classify PISs based on the presence or absence of certain features. To illustrate the process of building a decision tree, in Figure IA.A.1 (reproduced from the paper) we show a fictitious example of a decision tree for classifying PIS sections as ESG or non-ESG. For simplicity, we consider only three features: "ethic", "environment", and "water". We see that we first split the sample on "ethic", then on "environment", and finally on "water". It is clear that a first split on "ethic" works well because all PISs that contain it are classified as ESG, and a subsequent split on "environment" also works well because PISs that contain neither feature are all classified as non-ESG. The remaining PISs, i.e., those that do not contain "ethic" but contain "environment", can further—though imperfectly—be split on "water"; 80% of those that contain it are classified as ESG while 90% of those that do not contain it are classified as non-ESG.

6. We use cross-validation to optimally select the algorithm's hyper-parameters, specifically the number of decision trees in the forest, the size of the bootstrap subsample used to generate each

tree, the maximum depth of each tree, and the number of features considered at each node split. First, we construct a 4-dimensional grid of possible values, specifically we choose among number of trees in $\{500, 1000, 2000\}$, maximum depth in $\{10, 20, \infty\}$, bootstrap subsample size in $\{60\%, 70\%, 80\%\}$ of the training sample size, and number of features in $\{5\%, 10\%, 20\%\}$ of the total number of features. Then we use stratified $k$-fold cross-validation to pick the combination of hyper-parameters that optimizes performance. This scheme splits all the cases in the training sample into $k$ subsamples—called *folds*—that have similar size and similar proportions of each class; the training is then done in $k$ iterations, each time using $k-1$ folds to learn and leaving out one of the folds to test and, essentially, calculate accuracy metrics out-of-sample hence avoid overfitting. Figure IA.A.2 illustrates how the stratified $k-$fold cross-validation scheme works. We repeat this stratified $k$-fold scheme a number of times, with different randomization in each repetition, and then accuracy metrics (specifically, the $F_1$ score[1]) are averaged across repetitions. The fine-tuned hyper-parameters are the ones that yield the highest average accuracy metric across all iterations.

7. We remove features that may lead to overfitting. To do this, we identify features that are important in classifying the training sample but do not make much sense so are likely to be overfitted and lead to poor out-of-sample performance. To find the most important features, we use Shapley values; the Shapley value of a given feature for a given case is the weighted average—across all combinations of features that could be included in the model—of the difference in the model's prediction for this case with and without this feature. For each feature, we calculate the mean absolute Shapley value across cases, and then we sort features on this. We see that many important features make sense, but also that some do not; examples of the latter are "target", "equiti", and "research". While these features have a large average contribution in determining classification in the training sample, it is likely they will not work well in unseen cases, so it is preferable to exclude them from the model's training. To remove features, we iteratively train the model and each time we inspect the Shapley values for the 15 most important features and remove those for which there are cases with positive feature weight but zero Shapley value.[2]

8. With this finalized set of features, we fine-tune the algorithm's hyper-parameters again, and we train on the entire training subsample.

9. Finally, we use the trained algorithm model to classify *all* PISs in our data.

In Tables IA.A.1 and IA.A.2, we show how the algorithm performs in the testing subsample.

[1] $F_1 := \frac{2}{1/\text{precision}+1/\text{recall}}$, where precision $:= \frac{\#\text{ true positives}}{\#\text{ true positives}+\#\text{ false positives}}$ and recall $:= \frac{\#\text{ true positives}}{\#\text{ true positives}+\#\text{ false negatives}}$. Intuitively, the precision (or positive predictive value) is the probability that a PIS classified as ESG is truly that, and recall (or sensitivity) is the probability that an ESG PIS will be classified as such; $F_1$ is their harmonic mean.

[2] We end up removing the following features: alloc, avoid, develop, distribut, equiti, equiti equiti, financi, focus, fundament, futur, global, growth, incom, opportun, region, research, resourc, risk, servic, target, valu, world.

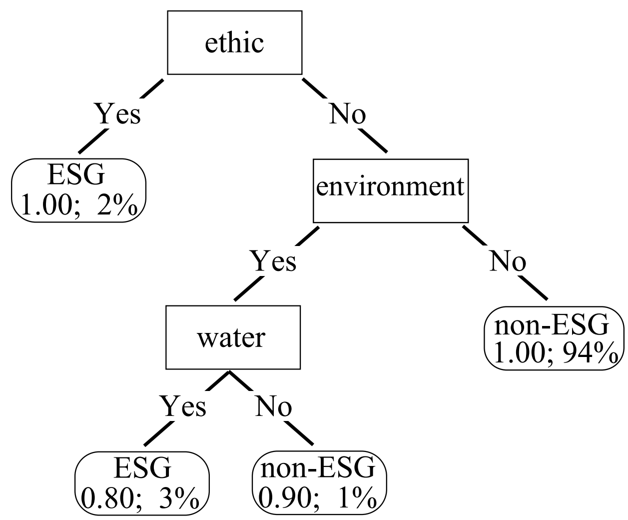


Figure IA.A.1: A (fictitious) example of a decision tree for the classification of PIS sections as ESG or non-ESG. Non-leaf nodes are indicated by regular rectangles and contain the feature on which cases are split. Leaf nodes are indicated by rounded rectangles and contain the majority classification (ESG or non-ESG) of the cases in the leaf, the proportion of these cases belonging to this classification, and the percentage of these cases in the sample. For example, the bottom left leaf node indicates that 3% of the cases in the sample do not contain the feature "ethic" but contain the features "environment" and "water"; of these, the majority (80%) are classified as ESG and the rest as non-ESG.

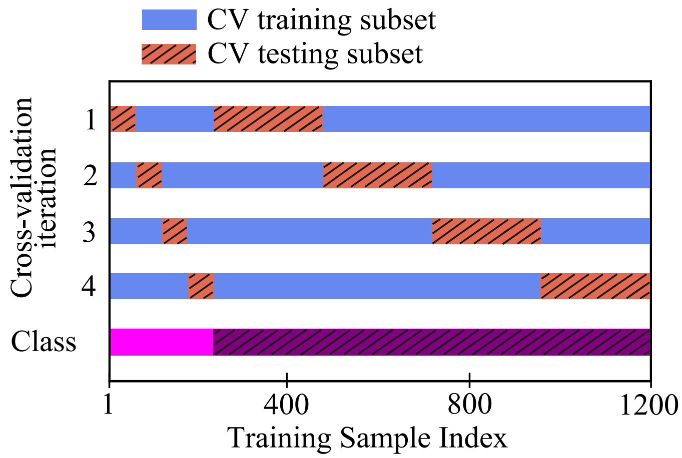


Figure IA.A.2: A (fictitious) example illustrating stratified $k$-fold cross-validation. The horizontal axis indicates the indices of the cases in the training sample, from case 1 to case 1,200, which are ordered so that those belonging to the first class (indicated, in the bottom row, with the pink-shaded area) appear before those belonging to the second class (indicated, in the bottom row, with the purple-shaded hatched area). For each iteration of the 4-fold cross-validation, the corresponding row shows which cases are used as a training subset (blue-shaded area) and which are used as a testing subset (orange-shaded hatched area).

Table IA.A.1: Performance for ESG classification on testing subsample — keyword-based vs. ML approach

This table shows the performance on the testing subsample of the keyword-based approach (in Panel A) and of the random forest approach (in Panel B) for classifying PISs as ESG or non-ESG. The performance metrics shown are: (i) accuracy, i.e., the proportion of correctly predicted classifications across both classes; (ii) recall, i.e., the accuracy within each class; (iii) precision, i.e., a classification's predictive value within each class; (iv) $F_1$, i.e., the harmonic mean of recall and precision within each class; and (v) $P_4$, i.e., the harmonic mean of the $F_1$ score across the classes.

Panel A: Performance metrics for keyword-based approach

| | accuracy | recall | precision | $F_1$ | $P_4$ |
|---|---|---|---|---|---|
| non-ESG | 98% | 100% | 98% | 99% | 91% |
| ESG | | 73% | 100% | 84% | |

Panel B: Performance metrics for ML approach

| | accuracy | recall | precision | $F_1$ | $P_4$ |
|---|---|---|---|---|---|
| non-ESG | 97% | 99% | 98% | 99% | 89% |
| ESG | | 77% | 85% | 81% | |

Table IA.A.2: Performance for E, S, and G classification on testing subsample

This table shows the performance on the testing subsample of the ML approach for classifying PISs as E or non-E (in Panel A), as S or non-S (in Panel B), and as G or non-G (in Panel C). The performance metrics are: (i) accuracy, i.e., the proportion of correctly predicted classifications across both classes; (ii) recall, i.e., the accuracy within each class; (iii) precision, i.e., a classification's predictive value within each class; (iv) $F_1$, i.e., the harmonic mean of recall and precision within each class; and (v) $P_4$, i.e., the harmonic mean of the $F_1$ score across the classes.

Panel A: Performance metrics for E classification

| | accuracy | recall | precision | $F_1$ | $P_4$ |
|---|---|---|---|---|---|
| non-E | 99% | 99% | 100% | 99% | 89% |
| E | | 86% | 75% | 80% | |

Panel B: Performance metrics for S classification

| | accuracy | recall | precision | $F_1$ | $P_4$ |
|---|---|---|---|---|---|
| non-S | 99% | 99% | 100% | 99% | 91% |
| S | | 89% | 80% | 84% | |

Panel C: Performance metrics for G classification

| | accuracy | recall | precision | $F_1$ | $P_4$ |
|---|---|---|---|---|---|
| non-G | 98% | 100% | 98% | 99% | 80% |
| G | | 55% | 86% | 67% | |

# IA.B Additional information on the data and results

In Table IA.B.1, we present summary statistics for the text characteristics of the entire PIS (in Panel A) and its ESG portion (in Panel B).

In Tables IA.B.2 and IA.B.3 we repeat the analyses of fund flows presented in Tables 4 and 9 of the paper, respectively, but, to allow for time-varying investment styles, fund category is inferred from the exposure of the past 24 months of fund returns to the Fama and French (1993) factors. Results are very similar to those in the corresponding tables presented in the paper, which utilize the commonly used CRSP investment styles.

In Table IA.B.4, we present an additional specification showing how fund flows respond to greenwashing versus truly green funds that include ESG keywords in the PIS text block of their prospectus. Specifically, the analysis presented in Table IA.B.4 is similar to that presented in Table 9 of the paper, with the difference that here we exclude from the analysis 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%. Again, results are very similar to those presented in the paper, which include in the analysis all fund-months regardless of portfolio coverage of stock-level ESG scores.

In Tables IA.B.5 through IA.B.9, we present analyses that differ from our baseline in the main paper in the following ways. First, we saturate the specifications with additional fund controls, specifically a dummy indicating if the fund name contains an ESG keyword, (log) assets under management by the fund family, 36-month return ranked within the investment category-by-month, dummies indicating if prior 36-month $\alpha$ is in the bottom or top 10% for the investment category-by-month, and dummies indicating if the Morningstar star rating is 1 or 5 stars. Second, we use a modified way of classifying funds as institutional vs. retail, specifically we follow Evans and Fahlenbrach (2012) and classify as institutional funds those that *only* have share classes open to institutional investors. Specifically, Table IA.B.5 is like Table 4 of the paper (fund flows and the presence of ESG keywords in the PIS); Table IA.B.6 is like Table 7 of the paper (fund characteristics related to greenwashing); Table IA.B.7 is like Table 8 of the paper (fund performance and greenwashing in the PIS); Table IA.B.8 is like Table 9 of the paper (fund flows and greenwashing in the PIS); and Table IA.B.9 is like Table C.2 of Appendix C of the paper (fund flows and the presence of ESG keywords in the PIS, but with interactions of ESG measures with institutional-fund dummies). Finally, in Table IA.B.10, we repeat the analysis shown in Table IA.B.5, separately for domestic-equity and for foreign-equity funds. For all these analyses, results are very similar to our baseline results with the exception that, moving from classifying institutional funds as those whose largest share class is targeted to institutional investors to those that only have share classes open to institutional investors, the estimated coefficient on the institutional fund dummy becomes negative and statistically significant (i) in all fund flow analyses (Tables IA.B.5, IA.B.8, IA.B.9, and IA.B.10) and (ii) in the analysis of fund characteristics related to greenwashing (Table IA.B.6). The former does not affect our message as the institutional fund dummy simply serves as a control in the fund flow analyses. We discuss the latter in detail in Section 5.1 the paper; essentially, the finding that institutional funds are less likely to greenwash is consistent with the idea that institutional investors may offer better oversight and therefore with our hypothesis that funds are more likely to greenwash if they are more likely to do so without being detected. Notably, moving to this alternative classification of institutional funds does not change the main result in Table C.2 of Appendix C of the paper, which is that fund flows do not respond to ESG keywords in the PIS differently for institutional vs. for retail funds.

Table IA.B.1: Summary statistics of text characteristics of PIS text block in fund prospectus

Summary statistics for text 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

| | | | | Percentiles | | |
|---|---|---|---|---|---|---|
| | # Obs | Mean | Std. Dev. | $10^{th}$ | $50^{th}$ | $90^{th}$ |
| 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

| | | | | Conditional Percentiles | | |
|---|---|---|---|---|---|---|
| | # Obs | Mean | Std. Dev. | $10^{th}$ | $50^{th}$ | $90^{th}$ |
| 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 IA.B.2: Fund flows and the presence of ESG keywords in the PIS – Alternative style categories

This table shows how fund flows respond to various definitions of text- and fundamentals-based ESG scores. The analysis presented in this table is almost identical to that presented in Table 4 of the paper, with the difference that here fund investment category is inferred from the exposure of the past 24 months of fund returns to the Fama and French (1993) factors. 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 $\alpha$ 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 | 0.270 *** | 0.329 *** | 0.344 *** | | | | |
| | 3.852 | 4.691 | 4.882 | | | | |
| ESG text relative length | | | | 0.014 *** | | | |
| | | | | 3.813 | | | |
| ESG in prospectus | | | | | 0.004 *** | 0.004 *** | |
| | | | | | 3.779 | 3.597 | |
| ESG keyword frequency > p50 | | | | | | | 0.007 *** |
| | | | | | | | 4.959 |
| Holdings ESG score (raw) | | 0.000 | | 0.000 | | | |
| | | 0.696 | | 0.724 | | | |
| Holdings ESG score (rank) | | | -0.000 | | | | |
| | | | -0.737 | | | | |
| Holdings ESG score > p50 | | | | | -0.000 | | |
| | | | | | -0.820 | | |
| Holdings ESG score > p90 | | | | | | 0.001 * | 0.001 * |
| | | | | | | 1.921 | 1.677 |
| Fund Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| # of Observations | 288,644 | 248,128 | 245,576 | 248,128 | 245,576 | 245,576 | 245,576 |
| Adjusted $R^2$ | 0.082 | 0.082 | 0.082 | 0.082 | 0.082 | 0.082 | 0.082 |

Table IA.B.3: Fund flows and greenwashing in the PIS – Alternative style categories

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 are truly green. The analysis presented in this table is almost identical to that presented in Table 9 of the paper, with the difference that here fund investment category is inferred from the exposure of the past 24 months of fund returns to the Fama and French (1993) factors. 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 truly green funds it is shown in the row presenting the interaction with the dummy indicating non-greenwashing funds (*non-GW* = 1 − *GW*). The specifications differ in the definition of the greenwashing dummy. In column 1, the greenwashing dummy is 1 for any fund that includes an ESG keyword in its PIS but whose MSCI holdings-based ESG score is below the $50^{\text{th}}$ 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 $50^{\text{th}}$ to the $25^{\text{th}}$ 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 $50^{\text{th}}$ 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 $25^{\text{th}}$ 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 $\alpha$ 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 | 0.510 *** | 0.696 ** | 0.658 ** | 0.763 ** |
| | 2.713 | 2.523 | 2.311 | 2.469 |
| ESG keyword frequency * Non-GW | 0.301 *** | 0.299 *** | 0.276 *** | 0.332 *** |
| | 3.998 | 4.168 | 3.800 | 4.640 |
| Holdings ESG score (raw) | 0.000 | 0.000 | 0.000 | -0.004 *** |
| | 0.950 | 0.960 | 0.680 | -3.893 |
| Fund Controls | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 245,468 | 245,468 | 238,773 | 246,952 |
| Adjusted $R^2$ | 0.082 | 0.082 | 0.082 | 0.083 |

Table IA.B.4: Fund flows and greenwashing in the PIS – Better ESG score coverage

This table presents additional specifications showing how fund flows respond to greenwashing versus truly green funds that include ESG keywords in the PIS text block of their prospectus. The analysis presented in this table is almost identical to that presented in Table 9 of the paper, with the difference that here we exclude from the analysis 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%. The specifications differ in the definition of the greenwashing (*GW*) and non-greenwashing (*non-GW*) dummy; for details on its definition, see the caption of Table 9. *ESG-keyword frequency* is the frequency of ESG keywords in the PIS text block. 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 $\alpha$ 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 | 0.480 * | 0.764 ** | 0.812 ** | 0.726 ** |
| | 1.726 | 1.997 | 2.326 | 2.125 |
| ESG keyword frequency * Non-GW | 0.191 ** | 0.183 ** | 0.145 * | 0.172 ** |
| | 2.371 | 2.359 | 1.888 | 2.356 |
| Holdings ESG score (raw) | -0.001 | -0.001 | -0.001 | -0.001 |
| | -1.232 | -1.195 | -1.312 | -0.803 |
| Fund Controls | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 166,914 | 166,914 | 165,320 | 172,495 |
| Adjusted $R^2$ | 0.082 | 0.082 | 0.083 | 0.082 |

Table IA.B.5: Fund flows and the presence of ESG keywords in the PIS – With additional controls

This table presents additional specifications showing how fund flows respond to various definitions of text- and fundamentals-based ESG scores. The analysis presented here is almost identical to that presented in Table 4 of the paper, with the differences that here (1) we saturate the specifications with additional controls and (2) we use a modified way of classifying funds as institutional vs. retail. Specifically, the additional controls are, for each fund: a dummy indicating if the fund name contains an ESG keyword; (log) assets under management by the fund family; 36-month return ranked within the investment category-by-month; dummies indicating if prior 36-month $\alpha$ is in the bottom or top 10% for the investment category-by-month; and dummies indicating if the Morningstar star rating is 1 or 5 stars. Furthermore, as in Evans and Fahlenbrach (2012), here we classify as institutional funds those that *only* have share classes open to institutional investors. All other variable definitions and the differences across specifications are as in Table 4 of the paper and as described in its caption. For ease of comparison across the two tables, here we present estimation results for the common regressors on the first page and results for the new regressors on the second page of the table. $t$-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 | 0.384 *** | 0.397 *** | 0.414 *** | | | | |
| | 5.017 | 4.907 | 5.066 | | | | |
| ESG text relative length | | | | 0.016 *** | | | |
| | | | | 3.613 | | | |
| ESG in prospectus | | | | | 0.003 *** | 0.003 *** | |
| | | | | | 2.832 | 2.697 | |
| ESG keyword frequency > p50 | | | | | | | 0.007 *** |
| | | | | | | | 5.382 |
| Holdings ESG score (raw) | | 0.000 | | 0.000 | | | |
| | | 0.384 | | 0.414 | | | |
| Holdings ESG score (rank) | | | -0.000 * | | | | |
| | | | -1.822 | | | | |
| Holdings ESG score > p50 | | | | | -0.001 * | | |
| | | | | | -1.689 | | |
| Holdings ESG score > p90 | | | | | | 0.001 | 0.001 |
| | | | | | | 1.193 | 0.880 |
| log(Fund size) | -0.002 *** | -0.001 *** | -0.001 *** | -0.001 *** | -0.001 *** | -0.001 *** | -0.001 *** |
| | -9.162 | -7.541 | -7.458 | -7.527 | -7.461 | -7.455 | -7.454 |
| log(Expense ratio) | -0.419 *** | -0.241 ** | -0.243 ** | -0.243 ** | -0.242 ** | -0.232 ** | -0.231 ** |
| | -4.635 | -2.584 | -2.610 | -2.602 | -2.594 | -2.504 | -2.495 |
| log(Effective 12b-1 fee) | 0.219 *** | 0.258 *** | 0.248 *** | 0.259 *** | 0.249 *** | 0.251 *** | 0.244 *** |
| | 3.413 | 4.007 | 3.857 | 4.029 | 3.874 | 3.898 | 3.797 |
| Prior 1-month return (raw) | 0.030 ** | 0.020 | 0.017 | 0.020 | 0.017 | 0.016 | 0.016 |
| | 2.586 | 1.580 | 1.363 | 1.579 | 1.352 | 1.312 | 1.287 |
| Prior 12-month return (rank) | 0.000 *** | 0.000 *** | 0.000 *** | 0.000 *** | 0.000 *** | 0.000 *** | 0.000 *** |
| | 13.162 | 13.799 | 13.875 | 13.794 | 13.843 | 13.714 | 13.683 |
| Prior 12-month alpha < p10 | -0.004 *** | -0.003 *** | -0.003 *** | -0.003 *** | -0.003 *** | -0.003 *** | -0.003 *** |
| | -7.065 | -5.833 | -5.884 | -5.837 | -5.902 | -5.925 | -5.892 |
| Prior 12-month alpha > p90 | 0.006 *** | 0.007 *** | 0.007 *** | 0.007 *** | 0.007 *** | 0.007 *** | 0.007 *** |
| | 10.045 | 10.615 | 10.489 | 10.627 | 10.475 | 10.493 | 10.526 |
| log(Fund age) | -0.007 *** | -0.005 *** | -0.005 *** | -0.005 *** | -0.005 *** | -0.005 *** | -0.005 *** |
| | -15.287 | -11.725 | -11.611 | -11.751 | -11.693 | -11.754 | -11.731 |
| log(Prospectus word count) | -0.000 | -0.001 *** | -0.001 *** | -0.001 *** | -0.001 *** | -0.001 *** | -0.001 *** |
| | -0.791 | -3.436 | -3.268 | -3.478 | -3.378 | -3.395 | -3.284 |

Continued on next page

Table IA.B.5 – continued from previous page

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| ESG in name | -0.006 * | -0.005 | -0.005 | -0.004 | -0.001 | -0.001 | -0.004 |
| | -1.865 | -1.350 | -1.274 | -0.869 | -0.184 | -0.296 | -1.275 |
| log(Fund family size) | 0.000 * | -0.000 | -0.000 | -0.000 | -0.000 | -0.000 | 0.000 |
| | 1.951 | -0.049 | -0.089 | -0.090 | -0.126 | -0.040 | 0.022 |
| Prior 36-month return (rank) | 0.000 *** | 0.000 *** | 0.000 *** | 0.000 *** | 0.000 *** | 0.000 *** | 0.000 *** |
| | 13.857 | 14.031 | 14.274 | 14.019 | 14.270 | 14.303 | 14.333 |
| Prior 36-month alpha < p10 | -0.004 *** | -0.004 *** | -0.004 *** | -0.004 *** | -0.004 *** | -0.004 *** | -0.004 *** |
| | -5.544 | -5.707 | -5.599 | -5.720 | -5.613 | -5.610 | -5.585 |
| Prior 36-month alpha > p90 | 0.004 *** | 0.004 *** | 0.004 *** | 0.004 *** | 0.004 *** | 0.004 *** | 0.004 *** |
| | 4.751 | 4.596 | 4.518 | 4.590 | 4.510 | 4.522 | 4.532 |
| Morningstar Star 1 | -0.003 *** | -0.003 *** | -0.003 *** | -0.003 *** | -0.003 *** | -0.003 *** | -0.003 *** |
| | -4.286 | -4.348 | -4.174 | -4.341 | -4.169 | -4.130 | -4.124 |
| Morningstar Star 5 | 0.019 *** | 0.019 *** | 0.019 *** | 0.019 *** | 0.019 *** | 0.019 *** | 0.019 *** |
| | 21.558 | 20.329 | 20.188 | 20.311 | 20.129 | 20.057 | 20.073 |
| Fund for institutionals | -0.002 *** | -0.003 *** | -0.003 *** | -0.003 *** | -0.003 *** | -0.003 *** | -0.003 *** |
| | -3.089 | -3.205 | -3.334 | -3.222 | -3.307 | -3.241 | -3.280 |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| # of Observations | 259,690 | 225,667 | 223,569 | 225,667 | 223,569 | 223,569 | 223,569 |
| Adjusted $R^2$ | 0.093 | 0.097 | 0.097 | 0.097 | 0.097 | 0.097 | 0.097 |

Table IA.B.6: Fund characteristics related to greenwashing – With additional controls

This table shows which fund characteristics are associated with greenwashing. The analysis presented here is almost identical to that presented in Table 7 of the paper, with the differences that here (1) we saturate the specifications with additional controls and (2) we use a modified way of classifying funds as institutional vs. retail. Specifically, the additional controls are, for each fund: a dummy indicating if the fund name contains an ESG keyword; (log) assets under management by the fund family; 36-month return ranked within the investment category-by-month; dummies indicating if prior 36-month $\alpha$ is in the bottom or top 10% for the investment category-by-month; and dummies indicating if the Morningstar star rating is 1 or 5 stars. Furthermore, as in Evans and Fahlenbrach (2012), here we classify as institutional funds those that *only* have share classes open to institutional investors. All other variable definitions and the differences across specifications are as in Table 7 of the paper and as described in its caption. $t$-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: With category fixed effects | | | Panel B: With category-by-time fixed effects | | |
|---|---|---|---|---|---|---|
| | (1) holdings discrepancy | (2) higher holdings discrepancy | (3) holdings & returns discrepancy | (1) holdings discrepancy | (2) higher holdings discrepancy | (3) holdings & returns discrepancy |
| log(Fund family size) | 0.003 | 0.001 | 0.001 | 0.003 | 0.001 | 0.001 |
| | 0.751 | 0.400 | 0.282 | 0.812 | 0.484 | 0.271 |
| log(Fund size) | -0.007 | -0.008 ** | -0.010 ** | -0.006 | -0.007 * | -0.009 ** |
| | -1.332 | -2.070 | -2.474 | -1.212 | -1.945 | -2.347 |
| log(Expense ratio) | 0.024 | 0.016 | 0.023 * | 0.036 | 0.021 | 0.033 ** |
| | 1.080 | 1.034 | 1.724 | 1.601 | 1.293 | 2.445 |
| Turnover ratio | 0.007 | 0.003 | -0.003 | 0.006 | 0.002 | -0.004 |
| | 1.066 | 0.707 | -0.609 | 0.925 | 0.601 | -0.716 |
| log(Effective 12b-1 fee) | 2.631 | 2.239 | -1.461 | 2.445 | 2.200 | -1.426 |
| | 0.763 | 1.083 | -0.768 | 0.727 | 1.087 | -0.753 |
| Prior 12-month mean flows | -0.131 ** | -0.099 ** | -0.096 * | -0.080 | -0.084 * | -0.063 |
| | -2.149 | -2.287 | -1.806 | -1.351 | -1.884 | -1.280 |
| Prior 12-month mean alpha | -1.472 | 0.263 | -1.217 | -0.859 | 1.606 | -0.739 |
| | -0.793 | 0.187 | -0.809 | -0.434 | 0.955 | -0.503 |
| Prior 36-month mean alpha | 2.850 | -0.879 | 1.054 | -3.127 | -4.063 | -3.755 |
| | 0.769 | -0.318 | 0.390 | -1.000 | -1.426 | -1.179 |
| Morningstar Star 1 | -0.025 | -0.021 | -0.045 ** | -0.040 | -0.024 | -0.056 *** |
| | -0.997 | -0.827 | -2.511 | -1.564 | -0.955 | -3.021 |
| Morningstar Star 5 | 0.014 | 0.021 | 0.013 | 0.023 | 0.023 | 0.018 |
| | 0.637 | 0.968 | 0.651 | 1.013 | 1.063 | 0.963 |
| log(Fund age) | 0.005 | -0.004 | -0.002 | 0.000 | -0.007 | -0.007 |
| | 0.352 | -0.360 | -0.185 | 0.027 | -0.701 | -0.669 |
| Fund for institutionals | -0.031 * | -0.023 ** | -0.029 * | -0.026 * | -0.022 * | -0.026 * |
| | -1.861 | -2.115 | -1.864 | -1.661 | -1.972 | -1.722 |
| After 03/2016 | 0.065 *** | 0.025 ** | 0.051 *** | | | |
| | 3.161 | 1.999 | 2.920 | | | |
| Category-by-Time Fixed Effects | No | No | No | Yes | Yes | Yes |
| Category Fixed Effects | Yes | Yes | Yes | No | No | No |
| # of Observations | 129,642 | 130,783 | 126,526 | 129,642 | 130,783 | 126,526 |
| Adjusted $R^2$ | 0.0004 | 0.0003 | 0.0003 | 0.010 | 0.008 | 0.009 |

Table IA.B.7: Fund performance and greenwashing in the PIS – With additional controls

This table shows the effect on fund performance (alpha) of the inclusion of ESG keywords in the PIS text block of a fund's prospectus, restricting the sample to domestic equity funds. The analysis presented here is almost identical to that presented in Table 8 of the paper, with the differences that here (1) we saturate the specifications with additional controls and (2) we use a modified way of classifying funds as institutional vs. retail. Specifically, the additional controls are, for each fund: a dummy indicating if the fund name contains an ESG keyword; (log) assets under management by the fund family; 36-month return ranked within the investment category-by-month; dummies indicating if prior 36-month $\alpha$ is in the bottom or top 10% for the investment category-by-month; and dummies indicating if the Morningstar star rating is 1 or 5 stars. Furthermore, as in Evans and Fahlenbrach (2012), here we classify as institutional funds those that *only* have share classes open to institutional investors. All other variable definitions and the differences across specifications are as in Table 8 of the paper and as described in its caption. $t$-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 | 0.021 | 0.024 | 0.076 | 0.001 |
| | 0.354 | 0.366 | 0.812 | 0.014 |
| ESG keyword frequency * Non-GW | 0.057 *** | 0.053 *** | 0.045 ** | 0.061 *** |
| | 3.018 | 2.763 | 2.203 | 3.294 |
| Holdings ESG score (raw) | 0.000 | 0.000 | 0.000 | -0.000 |
| | 1.556 | 1.569 | 1.320 | -0.339 |
| Fund Controls | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 163,696 | 163,696 | 159,405 | 164,769 |
| Adjusted $R^2$ | 0.084 | 0.084 | 0.095 | 0.084 |

Table IA.B.8: Fund flows and greenwashing in the PIS – With additional controls

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 are truly green. The analysis presented here is almost identical to that presented in Table 9 of the paper, with the differences that here (1) we saturate the specifications with additional controls and (2) we use a modified way of classifying funds as institutional vs. retail. Specifically, the additional controls are, for each fund: a dummy indicating if the fund name contains an ESG keyword; (log) assets under management by the fund family; 36-month return ranked within the investment category-by-month; dummies indicating if prior 36-month $\alpha$ is in the bottom or top 10% for the investment category-by-month; and dummies indicating if the Morningstar star rating is 1 or 5 stars. Furthermore, as in Evans and Fahlenbrach (2012), here we classify as institutional funds those that *only* have share classes open to institutional investors. All other variable definitions and the differences across specifications are as in Table 9 of the paper and as described in its caption. $t$-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 | 0.411 ** | 0.591 ** | 0.578 * | 0.735 ** |
| | 2.150 | 2.226 | 1.956 | 2.604 |
| ESG keyword frequency * Non-GW | 0.402 *** | 0.383 *** | 0.359 *** | 0.381 *** |
| | 4.629 | 4.618 | 4.467 | 4.672 |
| Holdings ESG score (raw) | 0.000 | 0.000 | 0.000 | -0.002 |
| | 0.128 | 0.171 | 0.181 | -1.437 |
| Fund Controls | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 223,468 | 223,468 | 217,408 | 224,587 |
| Adjusted $R^2$ | 0.097 | 0.097 | 0.099 | 0.097 |

Table IA.B.9: Fund flows and ESG keywords in the PIS – With institutional-fund interactions & additional controls

This table shows how fund flows respond to various definitions of text- and fundamentals-based ESG scores. The analysis presented here is almost identical to that presented in Table IA.B.5, with the difference that it estimates a different effect for funds targeted to institutional investors and to retail investors by including terms that interact each ESG measure with a dummy indicating institutional-targeted funds. As in Evans and Fahlenbrach (2012), here we classify as institutional funds those that *only* have share classes open to institutional investors. All variable definitions and the differences across specifications are as in Table IA.B.5. For the sake of brevity, we only report estimation results for the coefficients on text- and fundamentals-based ESG scores, and interactions with the institutional-fund dummy. $t$-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 | 0.387 *** | 0.402 *** | | |
| | 4.562 | 4.699 | | |
| ESG in prospectus | | | 0.003 ** | |
| | | | 2.506 | |
| ESG keyword frequency > p50 | | | | 0.007 *** |
| | | | | 4.927 |
| Holdings ESG score (raw) | 0.000 | | | |
| | 0.376 | | | |
| Holdings ESG score (rank) | | -0.000 | | |
| | | -1.424 | | |
| Holdings ESG score > p90 | | | 0.001 | 0.000 |
| | | | 0.933 | 0.624 |
| ESG keyword frequency * Institutional Fund | 0.118 | 0.139 | | |
| | 0.777 | 0.866 | | |
| ESG in prospectus * Institutional Fund | | | 0.001 | |
| | | | 0.258 | |
| ESG keyword frequency > p50 * Institutional Fund | | | | 0.004 |
| | | | | 1.056 |
| Holdings ESG score (raw) * Institutional Fund | -0.000 | | | |
| | -0.056 | | | |
| Holdings ESG score (rank) * Institutional Fund | | -0.000 | | |
| | | -0.894 | | |
| Holdings ESG score > p90 * Institutional Fund | | | 0.001 | 0.002 |
| | | | 0.566 | 0.590 |
| Fund Controls | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes |
| # of Observations | 225,667 | 223,569 | 223,569 | 223,569 |
| Adjusted $R^2$ | 0.097 | 0.097 | 0.097 | 0.097 |

Table IA.B.10: Fund flows and ESG keywords in the PIS – Domestic vs. foreign, with additional controls

This table shows how fund flows respond to various definitions of text- and fundamentals-based ESG scores. The analysis presented here is almost identical to that presented in Table IA.B.5, with the difference that we show results separately for domestic-equity funds (in Panel A) and for foreign-equity funds (in Panel B). All variable definitions and the differences across specifications are as in Table IA.B.5. For the sake of brevity, we only report estimation results for the coefficients on text- and fundamentals-based ESG scores. $t$-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: Domestic-equity funds

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| ESG keyword frequency | 0.238 *** | 0.299 *** | 0.324 *** | | | | |
| | 2.761 | 3.212 | 3.505 | | | | |
| ESG text relative length | | | | 0.011 ** | | | |
| | | | | 2.450 | | | |
| ESG in prospectus | | | | | 0.003 ** | 0.003 ** | |
| | | | | | 2.172 | 2.130 | |
| ESG keyword frequency > p50 | | | | | | | 0.005 *** |
| | | | | | | | 3.431 |
| Holdings ESG score (raw) | | -0.000 | | -0.000 | | | |
| | | -0.060 | | -0.039 | | | |
| Holdings ESG score (rank) | | | -0.000 ** | | | | |
| | | | -2.480 | | | | |
| Holdings ESG score > p50 | | | | | -0.001 * | | |
| | | | | | -1.931 | | |
| Holdings ESG score > p90 | | | | | | -0.000 | -0.000 |
| | | | | | | -0.341 | -0.517 |
| Fund Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| # of Observations | 195,780 | 165,477 | 163,492 | 165,477 | 163,492 | 163,492 | 163,492 |
| Adjusted $R^2$ | 0.092 | 0.095 | 0.095 | 0.095 | 0.095 | 0.095 | 0.095 |

Panel B: Foreign-equity funds

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| ESG keyword frequency | 0.546 *** | 0.495 *** | 0.499 *** | | | | |
| | 3.227 | 3.169 | 3.162 | | | | |
| ESG text relative length | | | | 0.023 *** | | | |
| | | | | 2.993 | | | |
| ESG in prospectus | | | | | 0.005 * | 0.004 * | |
| | | | | | 1.975 | 1.797 | |
| ESG keyword frequency > p50 | | | | | | | 0.010 *** |
| | | | | | | | 4.325 |
| Holdings ESG score (raw) | | 0.001 | | 0.001 | | | |
| | | 0.800 | | 0.802 | | | |
| Holdings ESG score (rank) | | | 0.000 | | | | |
| | | | 0.394 | | | | |
| Holdings ESG score > p50 | | | | | -0.000 | | |
| | | | | | -0.202 | | |
| Holdings ESG score > p90 | | | | | | 0.003 ** | 0.003 * |
| | | | | | | 2.133 | 1.856 |
| Fund Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Category-by-Time Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| # of Observations | 63,910 | 60,190 | 60,077 | 60,190 | 60,077 | 60,077 | 60,077 |
| Adjusted $R^2$ | 0.096 | 0.098 | 0.098 | 0.098 | 0.098 | 0.098 | 0.099 |

# References

**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.

**Fama, E. F, and K. R French.** 1993. "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics*, 33(1): 3–56.

**Flesch, R.** 1948. "A new readability yardstick." *Journal of Applied Psychology*, 32(3): 221–233.

**Gunning, R.** 1952. *Technique of clear writing.* McGraw-Hill.

**Sharpe, W. F.** 1992. "Asset allocation: Management style and performance measurement." *Journal of Portfolio Management*, 18(2): 7–19.