LEIBNIZ INSTITUTE
SAFE
Sustainable Architecture for Finance in Europe

Alexander Hillert | Alexandra Niessen-Ruenzi | Stefan Ruenzi

Mutual Fund Shareholder Letters: Flows, Performance, and Managerial Behavior

SAFE Working Paper No. 380 | February 2023

Leibniz Institute for Financial Research SAFE
Sustainable Architecture for Finance in Europe
info@safe-frankfurt.de | www.safe-frankfurt.de

Mutual Fund Shareholder Letters: Flows, Performance, and Managerial Behavior*

Alexander Hillert†, Alexandra Niessen-Ruenzi, and Stefan Ruenzi‡

January 2023

Abstract

Fund companies regularly send shareholder letters to their investors. We use textual analysis to investigate whether these letters' writing style influences fund flows and whether it predicts performance and investment styles. Fund investors react to the tone and content of shareholder letters: A less negative tone leads to higher net-flows. Thus, fund companies can use shareholder letters as a tactical instrument to influence flows. However, at the same time, a dishonest communication that is not consistent with the fund's actual performance decreases flows. A positive writing style predicts higher idiosyncratic risk as well as more style bets, while there is no consistent predictive power for future performance.

JEL-Classification Codes: G23, G11

Keywords: Fund Flows, Textual Analysis, Shareholder Letters, Investment Styles

*We would like to thank Darwin Choi (discussant), Tim Loughran, Nic Schaub (discussant), Laura T. Starks (discussant), and seminar participants at the 2014 European Finance Association (EFA) Meetings, the 2014 Asset Management Conference of Humboldt University and ESMT, the 2015 Colloquium on Financial Markets of the University of Cologne, the 2015 European Retail Investor Conference, the 2015 Financial Management Association (FMA) Meetings, 2016 Research in Behavioral Finance Conference Amsterdam, 2016 Annual Meeting of the German Finance Association, 2017 WU Gutmann Center Symposium, the DFG Network Textual Analysis in Economics and Finance, the Goethe University Frankfurt, the Technical University of Munich, the University of Hanover, the University of Mannheim, the University of Piraeus, the University of Zurich, and the WHU for helpful comments. Declarations of interest: none. We are grateful for financial support from the Julius Paul Stiegler Memorial Foundation. Alexander Hillert gratefully acknowledges research support from the Leibniz Institute for Financial Research SAFE. All errors are our own.

†Corresponding author. Goethe University Frankfurt and Leibniz Institute for Financial Research SAFE, Theodor-W.-Adorno-Platz 3, 60323 Frankfurt am Main, Germany, hillert@safe-frankfurt.de

‡University of Mannheim, L9, 1-2, 68131 Mannheim, Germany, alexandra.niessen-ruenzi@uni-mannheim.de; stefan.ruenzi@uni-mannheim.de.

Mutual Fund Shareholder Letters: Flows, Performance, and Managerial Behavior

Abstract

Fund companies regularly send shareholder letters to their investors. We use textual analysis to investigate whether these letters' writing style influences fund flows and whether it predicts performance and investment styles. Fund investors react to the tone and content of shareholder letters: A less negative tone leads to higher net-flows. Thus, fund companies can use shareholder letters as a tactical instrument to influence flows. However, at the same time, a dishonest communication that is not consistent with the fund's actual performance decreases flows. A positive writing style predicts higher idiosyncratic risk as well as more style bets, while there is no consistent predictive power for future performance.

JEL-Classification Codes: G23, G11

Keywords: Fund Flows, Textual Analysis, Shareholder Letters, Investment Styles

1 Introduction

Shareholder letters are part of the semi-annual shareholder reports (Form N-CSR and N-CSRS filings) that registered management investment companies file with the SEC and send to their investors. In these letters, fund managers discuss the general economic environment and outlook, as well as individual stocks that they hold and how these stocks performed. They also offer explanations for why certain stocks have been bought, sold or not been considered at all and comment on various other issues they consider relevant. The writing style of these letters varies greatly from very technical to almost literary.[1] In this paper, we examine whether fund investors are influenced by the writing style of these letters and whether the tone and content of these letters is predictive of a fund's future performance and investment style.

Although several papers have investigated the impact of soft information on financial markets based on textual analysis (see, e.g., Tetlock (2007), Tetlock, Saar-Tsechansky, and Macskassy (2008), Loughran and McDonald (2011)), the writing style of mutual funds' shareholder letters has not caught much attention in academic studies so far.[2] This is surprising, given that, according to the Investment Company Institute, in 2021 more than 48% of all households in the US own mutual funds and receive shareholder letters.[3]

Shareholder reports are sent out regularly to investors and are frequently quoted in the business press.[4] According to the ICI 2018 Shareholder Report Survey, 63% of mutual fund investors who recall having received a shareholder report read at least some part of the

[1] For example, Wintergreen Advisers' shareholder letters are written by a co-founder with a college minor in English language resulting in a letter full of "inspirational quotes". See http://online.barrons.com/article/SB50001424053111904009804579248280874966654.html

[2] Two notable exceptions are Hwang and Kim (2017), who look at shareholder reports of closed-end investment companies and show that firms issuing reports with low readability trade at significant discounts relative to the value of their fundamentals, and Chu and Kim (2019), who show that pessimism positively predicts closed-end fund performance.

[3] For a detailed view on the Investment Company Institute's annual statistics on households' mutual fund holdings, see http://www.ici.org/research/stats.

[4] For example, Morningstar offered a list of great shareholder letters to its readers: https://www.morningstar.com/articles/4793/great-shareholder-letters

report.[5] More than half of this group states that they have read at least 'a substantial part' of the report. As shareholder letters are typically at the beginning of funds' shareholder reports, this survey evidence suggests that a large number of investors is likely to pay attention to and may potentially react to funds' shareholder letters. Thus, it seems plausible that the writing style of these letters has an impact on investment decisions and capital flows.

If shareholder letters indeed influence mutual fund investors' investment decisions, they could be an important tactical tool for fund companies to influence mutual fund flows. This can be particularly important for fund companies that expect redemptions that would force them to engage in fire sales (e.g., Coval and Stafford (2007)). Shareholder letters can potentially help fund companies to influence investor behavior as they allow them to directly address a large number of investors and explain the fund's development in more detail. Furthermore, if shareholder letter writing style is predictive of fund managers' behavior and performance, this would be of foremost importance for fund investors interested in learning about the funds' future investment styles and performance.

In this paper, we use textual analysis methods to measure the writing style of shareholder letters of U.S. open-end equity mutual funds between 2006 and 2021. Specifically, we measure the tone of a letter based on its negativity. We find that mutual fund investors react strongly to the tone of shareholder letters. We document robust evidence that flows are significantly lower the more negative the tone of a shareholder letter is. Specifically, a one standard deviation increase in negativity predicts a $3.08 million average reduction of subsequent monthly fund flows. This effect amounts to more than 20% of the impact of a one standard deviation change in past performance (which cannot easily be influenced by the fund). These magnitudes are substantial and of relevance for fund companies, particularly given that the average monthly fund growth due to flows is negative during our sample period. Our findings are obtained after carefully controlling for variables that might influence letter tone and at

[5]The survey is available at https://www.ici.org/system/files/attachments/ppr_18_summary_shareholder.pdf.

the same time investor flows such as past fund and fund family performance and flows, and after including fund fixed effects.

Analyzing inflows and outflows–obtained from funds' N-SAR filings–separately shows that the effect of tone on flows is driven by both, lower outflows as well as additional inflows. Our main results also hold in a large battery of additional robustness tests and are stable across the first and second half of our sample and when we exclude the financial crises of 2008 and the Covid-19 crash in 2020.

We also analyze a row of potential mediators for the relationship between tone and flows like past performance, no-load status, or retail vs. institutional funds and find this relationship not to be significantly different between these groups. However, we find a strong impact of age: the impact of letter tone is much more important for younger funds, which can be explained by investors relying more on additional qualitative information like textual disclosures for funds with shorter and thus less reliable performance histories. In terms of economic significance, the impact of a one standard deviation change in negativity corresponds to more than a third of the impact of a one standard deviation change in past performance for younger funds.

For a subset of funds in our sample, we have daily flow data available. While we find no effect on flows before investors receive their shareholder letters, we observe a strong reaction of flows to tone that starts immediately after investors presumably receive their letters and that lasts for about two weeks. In this analysis, we also carefully control for other potentially confounding factors that might change at the disclosure date or that have happened during the reporting period and that the letter might be talking about, like manager, objective, fee, or name changes. The immediate flow reaction—combined with the fact that we carefully control for all other quantitative information contained in the filing as well as potentially confounding events—suggests a causal interpretation of the impact of letter tone on flows.

These findings might give rise to the question why not all funds then always write extremely positive letters. However, given that these letters are legally regulated to portray

a fair and truthful picture of the current economic situation of a fund, fund companies are restricted in the wording used in these letters. Thus, if they always write in an extremely optimistic tone they might face significant litigation risk. In addition, we find that overly optimistic letters after underperformance hurt fund flows, i.e., investors appreciate honest reports whose tone fits fund performance.

In a final step, we analyze whether the way in which letters are written is informative for future fund performance or behavior. While we do not find a stable relation between the tone of the shareholder letter and future performance, a negative tone does strongly predict less idiosyncratic risk-taking of fund managers, indicating that they make fewer benchmark-deviating bets after writing a negative shareholder letter. This finding is consistent with Scharfstein and Stein (1990), who show that more risk-averse managers tend to herd more. Consistently, we also provide evidence that a negative tone predicts less extreme style bets. Generally, these results indicate that letter tone allows investors to draw conclusions about the current attitude of fund managers and eventually their future investment behavior.

Our results have important implications for mutual fund companies and fund investors. They underline the importance of verbal information that fund companies provide to their investors: a significant number of investors seem to react strongly to the way in which shareholder letters are written.

Our paper contributes to several strands of the literature. We closely relate to the large literature on the determinants of mutual fund flows. Many papers examine the impact of past performance on fund flows (e.g., Sirri and Tufano (1998)). They find that investors (irrationally) chase funds that outperformed in the past although there is barely any skill-induced performance persistence of funds in the long-term (Carhart (1997)).[6] Solomon, Soltes, and Sosyura (2014) examine the impact of the media coverage of fund holdings on

[6]Other determinants of mutual fund flows that have been investigated in the literature comprise fund expenses (Barber, Odean, and Zheng (2005)), advertising of a fund (Jain and Wu (2000) and Gallaher, Kaniel, and Starks (2015)), a fund's media coverage (Kaniel, Starks, and Vasudevan (2007)), and fund manager characteristics (e.g., Wermers (2003), Kumar, Niessen-Ruenzi, and Spalt (2015), and Niessen-Ruenzi and Ruenzi (2019)).

investor flows. Also using textual analysis, an important recent paper by Kostovetsky and Warner (2020) analyses prospectuses of mutual funds to define a fund's uniqueness and finds that text-based uniqueness leads to more inflows for young funds. We contribute to this literature by showing that writing styles of mutual funds' shareholder letters, i.e., their general tone and honesty have a significant impact on mutual fund flows.

Furthermore, we complement the emerging literature exploring the information content on fund prospectuses. Sheng, Xu, and Zheng (2022) compare the risks disclosed in funds' summary prospectuses to funds' actual risk taking and find fund flows to be unrelated to risk coverage. Using machine learning, Abis (2022) classifies mutual funds into quantitative and discretionary funds and compares their investment approaches. Applying the k-means algorithm to funds' prospectuses, Abis and Lines (2021) identify 17 Strategy Peer Groups of funds and show that investors prefer less unique funds. DeHaan, Song, Xie, and Zhu (2021) examine the narrative complexity of the prospectuses of 38 S&P 500 index funds and find a positive association between fees and complexity. While these studies analyze funds' prospectuses, which show little variation in content over time, we study funds' shareholder letters whose content strongly varies over time (see Figure 1) and for whome we can identify time periods of increased attention towards them, namely around their respective filing date. This allows us to include fund fixed effects in our regressions so that our identification comes from within fund variation and results are not driven by unobserved heterogeneity across funds.

Altough several studies analyze the tone of financial texts and investor reactions in different contexts, the mutual fund setting offers several unique advantages: (1) investor behavior can be directly observed at the fund level because information on net flows is available, while existing studies on the impact of textual information on investor behavior have to rely on price reactions, (2) quantitative information relevant to investors when making investment decisions (e.g., fund performance, stock market performance) can exactly be controlled for, i.e., the soft- and hard information content can be separated, (3) the predictive power

of writing styles for managerial behavior (e.g., investment styles) and performance can be easily quantified, (4) the letters are contained in annual and semi-annual reports that are automatically sent to a large number of investors with known timing, and (5) the letters are prepared by individuals directly involved in the management of the fund (and not external parties as is the case when analyzing, e.g., newspaper article tone as in Tetlock (2007)), i.e., they are informative about the opinion and mood of these individuals and eventually their behavior.

2 Data and summary statistics

2.1 Shareholder letters

We obtain mutual funds' shareholder letters from annual N-CSR and semi-annual N-CSRS filings. According to section 30(e) of the Investment Company Act of 1940, every registered investment company has to transmit financial reports to its shareholders at least semiannually. These reports include information on the portfolio composition, a statement of income, and a balance sheet. Most importantly in our context, they usually include a letter directly addressing the fund's shareholders. These letters vary largely in terms of their content. For example, they can discuss the fund's performance relative to a benchmark, give reasons for why the fund outperformed or underperformed, describe economic and market conditions, highlight some securities of the portfolio (e.g., winners, losers, exposure to industries), or advertise the fund. In Section 1 of the Internet Appendix we present letters from two funds as examples. Negative words according to the Loughran and McDonald (2011) dictionary are printed in bold font. Inclusion of such a letter, classified as "narrative disclosure" by the SEC, is voluntary but must not contain any untrue statements and has to be certified by the mutual fund's principal executive and financial officers.

Since shareholder letters are not mandatory, there is no clear-cut section or item of the N-CSR filing which we can extract for our empirical analysis. Therefore, we identify

common phrases for the beginning and the ending of the letter to isolate it from the fund's financial report.[7] We also consider a fund's "Management's Discussion and Analysis of Fund Performance" to be a shareholder letter.

We are able to identify letters in more than 98% of all filings. Additionally, whether a filing contains a letter or not does not seem to have any impact on flows (see Table IA-1 in the Internet Appendix). Also, the results in Table IA-2 suggest that the decision to include a letter is strongly persistent and that funds do not change their decision to write a letter or not based on short-term outcomes.

We extract the text of these letters automatically and verify the precision of the letter extraction procedure by conducting manual checks. The manual checks show that the automatic extraction results in some implausibly short or long letters. We drop letters with less than 100 and more than 4,000 words. Furthermore, we exclude letters with less than 10 or more than 30 words per sentence on average. Additionally, we remove letter than contain more than 10% numbers. The details of the text extraction are described in Section 2 in the Internet Appendix. After these filters, we obtain a shareholder letter for 89% of the fund reports in our sample.

There are two separate dates included in N-CSR/N-CSRS filings that are relevant for our empirical analysis: the "report date" refers to the fiscal (half-)year end (i.e., reporting period end date), to which the filing refers, while the "filing date" marks the day on which the report is filed with the SEC. The distribution of financial reports over calendar months is displayed in Table IA-3 of the Internet Appendix. While most reports are filed in March (13.18%) and December (12.92%), report dates are fairly evenly distributed across the year.

According to SEC regulations, the maximum time span between the day on which the fund company sends off a report to investors and the day on which the report is filed with the SEC is 10 days. Thus, the earliest possible date at which we could expect a flow reaction is 10 days prior to the official filing date. However, as fund companies have no incentives to

[7] Common phrases include, e.g., "Dear Shareholders", "Dear Investors", "Sincerely", or "Yours Truly".

postpone the filing of their reports with the SEC after they sent them out to shareholders, it is likely that this typically happens soon afterwards. Thus, the filing date can be assumed to be relatively close to the day on which the fund's shareholders also receive the respective reports. Our later analysis of daily fund flows in Section 3.4 supports this assumption.

We use the Loughran and McDonald (2011) dictionary of negative words to measure the tone of shareholder letters.[8] The measure is labelled LMD^{-}.[9] We focus on dictionaries capturing negative tone as previous work has shown that positive statements often have an ambiguous meaning, which makes it much harder to precisely capture positive tone (see, e.g., Loughran and McDonald (2016)). Table IA-4 of our Internet Appendix lists the top 25 negative words that appear in our sample of letters.

Furthermore, we compute two refined negativity measures based on whether a negative word is used in forward or non-forward looking parts of the letter. Using the word list by Li (2010) to define forward and non-forward looking sentences, for each letter we calculate negativity separately for those two cases.[10] All negativity measures are standardized to mean zero and unit variance.

2.2 Mutual fund data

We use mutual fund data from the Center for Research on Security Prices (CRSP) survivorship bias free mutual fund database and aggregate all share classes at the fund level. To merge shareholder letters with the CRSP database, we establish a unique link between the Series ID obtained from the SEC filing and the WFICN (Wharton Financial Institution Code Number) of fund portfolios as provided in MFLinks. The matching procedure is based on the fund's ticker symbol.[11]

[8] Except for the shareholder letter, Form N-CSR/N-CSRS filings mostly provide quantitative information including a fund's expenses, its past performance, and its portfolio holdings. Consequently, we do not compute a tone measure for the rest of the document, as it would not provide any meaningful qualitative information.

[9] All variables are described in more detail in Section 3 of the Internet Appendix.

[10] Section 4 of the Internet Appendix shows five examples of negative forward-looking and non-forward-looking sentences.

[11] More details on the matching process are provided in Section 2 of the Internet Appendix.

To allow for easy comparability of performance across funds, we exclude balanced funds, money market funds, fixed income funds, and exchange traded funds. In addition, we drop all index funds and observations where a fund's total net assets in a given month are below one million dollar. Our final sample comprises 5,489 matched open-end equity funds from 2006 to 2021, a period that contains the bear market around the financial crises of 2008 and the Covid19-crash in 2020 but also the extended bull market from 2009 to 2017 and the strong recovery after the Covid19-crash till the end of 2021.

2.3 Summary statistics

Summary statistics are presented in Table 1. In Panel A, we present summary statistics for shareholder letters. The mean percentage of negative words in a given shareholder letter according to the LMD^{-} dictionary is 2.00%, which is similar to the average negativity of 1.51% for the MD&A section of 10-K filings documented in Loughran and McDonald (2011). The fraction of negative words varies substantially between virtually zero and more than 5%. The average shareholder letter includes about 875 words and there is a 63-day time difference between reporting period end date and the filing date.

In Panel B, we present summary statistics on fund characteristics. Average monthly flows amount to -0.28%. This number may seem small as compared to average flows from earlier studies focusing on samples from the 1980s and 1990s. However, it is consistent with the much lower aggregate growth rates of the active mutual fund industry during recent years that comprise our sample period (see, e.g., Barber, Huang, and Odean (2016), Berk and van Binsbergen (2016), Sialm, Starks, and Zhang (2015)). The average fund in our sample has total net assets of $2,363 million, is 17 years old, and has an annual expense ratio of about 1.16%.

Correlations between our main variables of interest are shown in Table IA-5 in the Internet Appendix. In the Internet Appendix, we also analyze the determinants of letter tone based on a regression analysis. Depending on the specification, we find a negative impact of past

fund and/or family performance and flows on negativity, i.e., better performance and higher flows correspond to less negative tone. A detailed description of this analysis is contained in Section 5 and the corresponding Table IA-6 of the Internet Appendix.

To get a better impression on how negativity varies during our sample period, Figure 1 plots average letter negativity as well as lagged 6-month S&P 500 returns. Visual inspection shows that negativity and lagged stock market returns are negatively related: After periods of low market returns, negativity increases substantially, while the opposite holds for periods after high market returns. These patterns show that our tone measure performs well in capturing the general tendency of market conditions. At the same time, the time-series correlation between average letter tone and market returns mandates to control for the general market environment and we do so in our later analysis by including various sets of time-fixed effects.

Besides this time-series variation in negativity, there is also large cross-sectional variation in writing styles. For example, some letters are written in a very technical manner with a formal discussion of the fund's financial outcome, while others are written in a very literary style with quotes and humorous comments. Section 1 in the Internet Appendix shows two excerpts of shareholder letters filed in December 2008 as examples. Both funds had significant negative returns over the six-month reporting period. The first fund, offered by American Century Investments, delivered a return of –34%, while the second fund, offered by Virtus Insight Trust, delivered a very similar return of –33%. Interestingly, the fund managers of these funds offer different views on how to interpret the financial outcome of the respective fund. The first shareholder letter makes several positive and encouraging statements, e.g., that "we are financially strong" and that the fund can "identify attractive investment opportunities regardless of market conditions". In contrast, the second letter contains more negative statements. For example, it talks about a "constant flow of negative news" and "destabilized financial markets". This difference is also reflected by the negativity measure which is 2.78% for the first letter, and 4.33% for the second letter, respectively. We con-

jecture that the first fund manager will, ceteris paribus, find it easier to avoid redemptions than the latter.

3 Letter tone and investor behavior

3.1 Main results based on monthly flow data

To investigate the impact of shareholder letters on mutual fund flows, correctly specifying the time structure in our empirical model is important. In our main specification, we relate fund flows in month t to the tone of a shareholder letter filed in month t only if the fund files the letter with the SEC until day 15 of a given month. If a fund files the letter after day 15, we relate its flows in the subsequent month $t+1$ to the tone of a shareholder letter filed in month t.[12] This specification helps us to capture the flow effects better if they occur in a relatively short time period after the letters are sent out. We think that this is likely to be the case since investors probably react to a shareholder letter immediately after they receive and read it, or not at all.[13]

The dependent variable in our main regressions is monthly fund flows winsorized at the top and bottom 1%. We relate fund flows to our negativity measure, LMD^-, and a set of controls.

Various papers show that past performance has a first-order impact on flows. Thus, we control for past fund performance in all regressions. Specifically, in our baseline specification, we include the short-term return between the reporting date and the filing date, the return of the fund over the six-month reporting period, and over the previous year (that ends with the end of the reporting period). The first measure is included to control for the performance

[12] In Table IA-7 in the Internet Appendix, we alternatively relate fund flows in month t (irrespective of the day the letter is filed) to the tone of a shareholder letter filed in the same month and our results remain statistically significant. The drawback of this approach is that funds filing their shareholder letters at the very end of month t are less likely to experience the complete flow effect within the same month.

[13] In our later analysis in Section 3.4, we investigate daily fund flows that circumvent the problem of properly defining a time structure based on monthly data as described above. However, daily flows are only available for a subsample of funds and a shorter time period.

of the fund after the end of the reporting period, as many investors might react to the short-term performance of a fund. The second measure is included to control for the return over the last reporting period that can also have an impact on both, the writing style and inflows (see Table IA-6 of the Internet Appendix). We additionally include the third measure, as the annual N-CSR reports typically discuss past performance over the whole year. We also add lagged fund family performance, lagged fund and fund family flows, lagged fund and fund family size, fund age, the fund's expense ratio, and fund risk. Furthermore, we control for the logarithm of the number of words a letter contains and the number of days between the shareholder letter's reporting and filing date. Standard errors are double-clustered at the fund and time (i.e., filing month) dimension.[14]

The regressions are estimated with fund, report month, and combined filing month and investment objective fixed effects. Investment objectives are taken from the CRSP database and are based on Lipper objective codes. Time fixed effects control for the impact of general market conditions on flows, while the inclusion of fund fixed effects takes care of the possibility that some funds' letters are always written in a certain style and combined filing month and investment objective effects control for anything that makes a segment attractive for investors at a certain point in time. Thus, our findings are driven by within fund time-series variation in the tone of the shareholder letter.

Results are shown in Table 2. In Column (1) we find that fund flows are significantly (at the 1% level) negatively related to the negativity of a shareholder letter, i.e., the more negative the tone of a letter, the lower the fund's subsequent flows. The impact of letter tone on flows is also economically meaningful: a one standard deviation increase in the fraction of negative words leads to flows in the subsequent month which are $3.08 million lower for a fund of average size, which is equivalent to a reduction in the growth rate of 0.13 percentage points. To put this into perspective, a one-standard deviation change in 12-month returns

[14]In Table IA-8 of the Internet Appendix, we alternatively double-cluster standard errors by fund and month of the flow reaction (Panel A), fund and report month (Panel B), and fund family and filing month (Panel C) and obtain very similar results.

is associated with a 63 basis point change in fund flows, i.e., the impact of a one standard deviation change in negativity corresponds to roughly 20% of the impact of a one-standard deviation change of past year's performance, a measure that is much harder to influence. Furthermore, given that the average fund in our sample receives even slightly negative net flows of -0.28 percent per month (see Table 1, Panel B), it does make sense for funds to devote attention to how letters are written.

Regarding our controls, we confirm the results from the previous literature like, e.g., the positive impact of past performance on fund flows, the persistence of flows, and lower percentage growth rates of larger funds.

Since letter tone is influenced by previous performance (see Table IA-6 in the Internet Appendix), and performance, at the same time, is a key driver of fund inflows (e.g., Sirri and Tufano (1998)), it is crucial to carefully control for the impact of past performance on flows. We therefore additionally control for a fund's CAPM-alpha over the previous year in Column (2), as Berk and van Binsbergen (2016) show that investors mainly react to this performance measure. Furthermore, to control for the possible convexity of the performance flow relationship we add previous year return ranks and squared return ranks in Column (3). Alternatively, in Column (4), we check the stability of our results when using three piecewise-linear regression coefficients which allow us to separately estimate the impact of performance on flows for the bottom, the three middle, and the top quintile of past performance (as in Sirri and Tufano (1998)).

We find a significant impact of the CAPM-alpha and a positive impact of return ranks, while squared return ranks show a positive albeit insignificant coefficient estimate. In Column (4), performance within all three segments has a signficantly positive impact on flows and the impact is weakest for the middle three quintiles and strongest in the top quintile.[15] Irrespective of this, our main result of a significantly negative impact of negativity on

[15]The convexity is less pronounced as compared to studies like Sirri and Tufano (1998) that focus on earlier sample periods, but very similar to the shape of the performance-flow relation documented in studies focusing on more recent sample periods (e.g., Sialm, Starks, and Zhang (2015)).

flows remains unaffected and is always significant at the 1%-level and of similar economic magnitude.

In Panel B of Table 2, we explore investors' reaction to shareholder letter tone in more detail by differentiating between outflows and inflows, respectively. Funds' outflows and inflows are obtained from funds' N-SAR filings, while our baseline analysis in Panel A is based on implicit flows computed based on CRSP data (see Section 2). If investors' reaction to the tone of shareholder letters is driven by new investors or existing investors increasing their investments, we would expect a negative relation between inflows and shareholder letter negativity. In contrast, if investors' reaction comes from existing investors withdrawing their money from the fund after a pessimistic letter, there will be a significant positive relation between letter negativity and outflows.

Results based on flows from N-SAR filings are presented in Panel B. First, for comparison, in columns (1) and (2) we re-do our baseline analysis for net-flows using N-SAR data based on the specifications in Columns (1) and (4) of Panel A. Results confirm our earlier findings of a highly significant negative impact of negativity on flows. When looking at the impact of negativity on inflows and outflows separately, we find a significant impact on both. While coefficient estimates and statistical significance are slightly larger for outflows than for inflows, effect sizes are of comparable magnitude.

3.2 Robustness and Interactions

In Panel A of Table 3 we show the results of several robustness tests we conduct to analyze the stability of our main result. For brevity, we only report the coefficients on negativity and on those control variables that are newly added to the baseline specification from column (1) of Panel A in Table 2. First, to assess the temporal stability of our results, we split our sample period in the middle and re-do the analysis separately for the years 2006 to 2013 and 2014 to 2021. Results are shown in columns (1) and (2) and confirm a significantly negative relationship between negativity and flows in both subperiods. To address the concern that

our results might be driven by some specific events during our sample period, in columns (3) and (4) we exclude letters that were filed during the financial crisis[16] (i.e., between September 2008 and March 2009) or during the Covid19-crash (i.e., between February 2020 and April 2020). We again find very similar results, alleviating concerns that our result hinge upon these crises periods. We also re-do this analysis, but exclude letters where the reporting period contains one of the crises months. Results in Table IA-9 in the Internet Appendix again show virtually identical results.

Columns (5) and (6) are corresponding to column (3) in Panel A of Table 2, but we replace the linear and squared performance ranks based on raw returns by ranks based on funds' 1-factor and 4-factor alphas, respectively. Additonally, we add five dummy variables for funds' Morningstar ratings (e.g., Evans and Sun (2021)). Our main result remains statistically significant and economically meaningful. Regarding the relation between fund flows and Morningstar ratings, we confirm the result of Del Guercio and Tkac (2008) that five-star funds receive disproportionally large flows.[17]

In column (7), we add a textual similarity measure to control for the amount of new information provided in the shareholder letter. More specifically, we compute the cosine similarity between the current year's and previous year's shareholder letter. The construction of the cosine similarity measure follows Cohen, Malloy, and Nguyen (2020). While we find no significant relation between fund flows and the amount of new information provided in the letter, our main result – the negative effect of negativity on fund flows – remains significant at the 1% level. Finally, in column (8), we replace the percentage fund flows in the filing month by the change in funds' market share in the filing month as in Spiegel and Zhang (2013). We still find a significant negative effect of shareholder letter tone on fund investors' investment decisions.

[16]'Crisis' is among the top-10 negative words, see Table IA-4 in the Internet Appendix.

[17]Note that, in this specification, we do not find a convex performance-flow relation (e.g., Sirri and Tufano (1998)) because we control for funds' Morningstar ratings. Without controlling for funds' Morningstar ratings, we confirm the convex performance-flow result from the previous literature.

In Panel B, we analyze potential mediators for the impact of negativity on flows. One might expect that fund investors respond more strongly to letter tone when their fund is under-performing. We test this conjecture in columns (1) to (4) by interacting negativity with the return over the last (two) reporting period(s) and with a dummy that takes on the value one if the return over the last (two) reporting period(s) was negative, respectively. In each case, we obtain an insignificant coefficient for the interaction term. Thus, we conclude that the strength of the investor reaction to shareholder letter tone does not depend on the performance of the fund. In column (5), we analyze whether there is a stronger investor reaction to the tone of shareholder letters in the annual N-CSR filings relative to the semi-annual N-CSRS filings, but do not find this to be the case.

In columns (6) and (7), we investigate whether a fund's distribution channel or investor type (retail/institutional) affect investors' reaction to shareholder letter tone, i.e., whether there is a difference in the tone-flow relation for direct-sold and broker-sold funds. We classify no-load funds as direct-sold and load funds as broker-sold funds. While the size of the coefficient on negativity is significant and close to the base specification, the no-load fund dummy and its interaction with negativity are close to zero and insignificant. This result is in line with the view that investors' reaction is mainly driven by existing investors withdrawing their money after a negative shareholder letter and that, thus, a fund's distribution channel—which is more likely to be important for the buying decisions of new investors—does not affect the tone-flow relation. In column (7), we analyze whether there is a stronger flow reaction to tone for retail relative to institutional funds. The coefficient for the interaction is not statistically significant and our baseline effect is still highly significant. Thus, we conclude that both retail and institutional investors respond to shareholder letters.[18]

Last, in column (8), we analyze the interplay between fund age and tone. As there is less past performance information available for young funds, investors may put more weight

[18]In our regressions, identification comes from within-fund time-series variation, i.e., in specifications (6) and (7) we analyze funds switching from load to no-load and from retail to institutional, respectively. When we re-estimate the regressions without fund fixed effects, i.e., focus on cross-sectional differences, the interaction effects are still insignificant.

on alternative sources of information like the fund's shareholder letters. In line with this idea, we find a significant negative interaction between negativity and a young fund dummy that takes the value of one if fund age is below the median, and zero otherwise. In terms of economic significance, the coefficient of the interaction term suggests that a one standard deviation increase in negativity is associated with 0.23 percentage points lower fund flows for young funds relative to old funds.[19] Note that the standard deviation of negativity for young and old funds is very similar (1.07% vs. 1.11%), i.e., the stronger negativity-flow relation for young funds is not driven by larger variation in tone.[20]

In the Internet Appendix we show (and describe in more detail) results from a large additional battery of robustness tests. First, we control for the impact of confounding events like manager, objective, fee, or name changes (Table IA-11). Second, we analyze flows always in the month of the filing day, irrespective of whether this was early or late in the monthly (Table IA-7). Third, we cluster standard errors in various ways (Table IA-8). Fourth, we add cubic terms for the impact of performance ranks (Table IA-12). Fifth, we add controls for the position of the person signing the letter (Tables IA-13 and 14). In all of these tests, we find very stable and consistent evidence of a negative impact of negativity on fund flows.

We also try out different alternative tone measures. We rerun our main regression using an orthogonalized negativity measure, look at changes in tone, add a positivity measure based on the respective Loughran and McDonald (2011) dictionary, reduce the effect of high-frequency (and potentially ambiguous) words on our negativity measure by using a term and inverse document frequency based negativity measure (see Loughran and McDonald (2011)), measure negativity based on the Harvard IV-4 psycho-social dictionary as in Tetlock (2007),

[19] Generally, the stronger impact of negativity for younger funds, which have less past data to reliably estimate performance, is consistent with a model of Bayesian updating where investors weight different sources of information according to their precision. Thus, the finding is consistent with Chevalier and Ellison (1997), who show that the most recent performance has a stronger impact on flows for younger funds, and Huang, Wei, and Yan (2012), who find that past performance has a weaker impact if it was more volatile.

[20] In the Internet Appendix we show that the interaction effects from column (1) to (7) are also insignificant based on N-SAR net-flow data and for inflows as well as outflows separately, too (see Table IA-10). We also find a stronger flow reaction to tone for younger funds for net-flows, which is mainly driven by inflows.

and analyzing negativity based on a newly developed dictionary for negativity in a mutual fund context. All results are shown in Table IA-15. Section 6 of the Internet Appendix provides further information on the alternative tone measures. We find no impact of positivity, which confirms the argument of Loughran and McDonald (2016) that it is tricky to precisely capture positive tone. We always find a significant negative impact of negativity on tone, with the exception of negativity based on the Harvard IV-4 dictionary (this dictionary was not designed to capture negative tone in financial texts, but in a general context), which highlights the importance to use dictionaries that are constructed for the specific context analyzed (Loughran and McDonald (2011)). Confirming this line of reasoning, we do find the strongest effect based on negativity from a negativity dictionary we developed specifically for the mutual fund context.[21]

3.3 Speaking about the past or the future

Shareholder letters can (and typically do) talk about and explain the past performance of the fund, but they might also contain an outlook for the future. Thus, in the following analysis, we differentiate between forward-looking and non-forward-looking sentences to analyze whether investors react more strongly to a negative outlook or to a negative description of the past.[22] We compute our negativity measure separately for forward-looking and non-forward looking sentences and replace our overall negativity measure by these two measures. We also control for the fraction of forward-looking sentences in the letter. Furthermore, the same control variables as in Column (1) of Panel A in Table 2 are included but suppressed in results shown in Table 4.

Column (1) shows results based on our main sample. We find a negative and significant coefficient estimate for both, the impact of (standardized) negativity in forward and non-

[21] In Table IA-16 of the Internet Appendix, we control for the percentage of uncertainty and modal weak words according to the (Loughran and McDonald 2011) dictionaries. While negativity still has a strong negative effect on flows, uncertainty and modal weak words are unrelated to flows.

[22] Section 4 in the Internet Appendix shows five examples of negative forward-looking and non-forward-looking sentences.

forward looking sentences. However, the coefficient estimate for the impact of non-forward looking negativity is about twice as large as that for forward looking negativity and statistical significance is stronger (t-stat of -2.78 for non-forward looking negativity vs. -1.66 for forward looking negativity). These results might be partially explained by the fact that typically only a small part of the letters is forward looking (about 25% of all sentences are forward looking, see Table 1) and that investors mainly care about an evaluation of what happened in the past rather then speculations about the future. This pattern is also consistent with a significantly negative impact of the fraction of forward-looking sentences that we find in Column (1).

To analyze whether outflows (that can by definition only come from existing fund shareholders) react more strongly to non-forward looking tone than inflows, we run the same regression model as in column (1) separately for inflows and outflows. Results based on N-SAR data are reported in columns (2) to (4). In column (2), for comparison, we again first analyze net flows and find a strong negative impact of non-forward looking negativity, consistent with the results in column (1). Furthermore, this impact is again stronger than the impact of forward looking negativity, which now looses its significance. Looking at inflows and outflows separately, we find a very similar picture: there is a significant negative (positive) impact of negativity in non-forward looking sentences on inflows (outflows). In absolute terms, the coefficient estimate is larger for the impact on outflows, but the difference is small.

3.4 Short-term dynamics: daily fund flows

To further refine our main result of a negative relation between fund flows and tone, we now analyze how daily fund flows react to shareholder letters around the filing date. This short-term analysis allows us to better capture an immediate reaction of fund flows to shareholder letters and to investigate the time pattern of the flow reaction in more detail. This analysis also helps us to identify the causal impact of letter tone more clearly in an event-study-like setting.

While CRSP only contains monthly TNA data, a proxy for daily flows can be calculated based on daily TNA and return data provided in the Morningstar database. We have daily flow data from Morningstar (starting in July 2008) for 76.4% of the funds with letters in our sample For those funds for which daily flow information is available, we merge daily flow data from Morningstar with CRSP/MFLinks using the fund's 9-digit CUSIP.

Figure 2 shows cumulative daily flows from 10 days before to 20 days after the filing date of a shareholder letter. We plot the cumulative abnormal flows separately for funds with above- and below-median realizations of their orthogonalized negativity measures. Visual inspection shows a strong divergence of flows starting around the filing date. Funds with above-median negativity face strong outflows, while funds with below-median negativity exhibit positive net flows. After about two weeks, the difference in cumulative flows only slightly increases further, which suggests a relatively rapid flow reaction to letter tone. We observe no reversal of the flow effect during the event window.

To investigate the exact time pattern of the flow reaction more formally, we compute fund flows for different time windows around the filing date. Specifically, we compute non-overlapping 5-day flows from 15 days before to 15 days after the filing date and rerun our main regression specification (column (2) of Table 2). While the clear time-structure in Figure 2 suggests a causal impact of the shareholder report investors receive at that time, one concern is that investors do not react to the tone of the letter contained in the report, but to other recent changes in fund features that are communicated in the report or that happen around the filing date. While we already carefully control for past performance, there might be other confounding events that investors learn about for the first time when receiving their reports. Thus, we add dummy variables taking on the value one if there was a manager change, a fund name change, an investment objective change, a liquidation of a share class, a merge of share-classes, and an expense ratio change, respectively, and zero otherwise, as well as the (cumulative) change in expense ratios to all regressions. Results are reported in Table 5.

Looking at the pre-filing date period, only the last 5-day window prior to the filing date is significantly related to shareholder letter negativity. There is no significant reaction prior to day t-5. This result indicates that there is no systematic difference in the flows between funds with more positive and funds with more negative shareholder letters prior to the date when shareholders become aware of the letters.[23]

On the filing day t, we observe a negative coefficient on negativity which is, however, not significant at conventional levels (t-statistic -0.83). This insignificant result can be explained by the fact that we only look at one individual day here and some letters reach investors prior to day t=0 while others that are only sent out on the filing date by mail might reach investors with a delay, and by investors needing time to process the information.

For the post-filing date period, there is a strong reaction of daily fund flows to shareholder letter negativity. With a t-statistic of -2.83, the effect is strongest within the first five days after the shareholder letter is filed with the SEC. In economic terms, the coefficient implies that a one standard deviation increase in negativity is associated with a 2.6 bp decrease in flows from t+1 to t+5. The negative flow reaction becomes weaker during the subsequent 5-day window (t-statistic 2.30). For days 11 to 15 after the filing date, we still observe a negative sign but the relationship is no longer statistically significant.

Our results on the impact of potential confounding events show a significantly negative reaction to a fund manager and investment objective change, to a share class liquidation, and to an increase in expense ratios, while there is no reaction to any of the other variables.

Overall, our results confirm the general patterns from Figure 2 and show that most of the flow reaction occurs right after investors receive their letters, supporting a causal link between letter tone and flow reactions.

[23]Note that investment companies are allowed to send out shareholder letters up to 10 days before filing them with the SEC. Thus, some investors could receive a letter and react to it before the filing date t which explains the significant negative sign of the coefficients on negativity for the pre-filing period t-5 to t-1.

3.5 Honesty and Flows

Our results hitherto raise the question why fund letters are then not always written in an extremely positive style. In this context, it is important to note that all of our regressions include fund fixed effects, i.e., the effects we document are driven by within-fund variation in letter tone, and it seems very unlikely that a fund can write letters that continuously sound more positive over time. Furthermore, the N-CSR filing (of which the fund letter is a part) needs to be signed and the signee certifies that "this report does not contain any untrue statement of a material fact". Thus, fund companies might face litigation risk if they always write extremely positive letters. In a corporate context, Rogers, van Buskirk, and Zechman (2011) show that overly optimistic tone in earnings announcements can indeed increase litigation risk for companies.

Furthermore, if managers write extremely positive letters irrespective of how the fund actually performed, investors might perceive fund managers to be less trustworthy.[24] To test the importance of communicating honestly with shareholders more explicitly, we introduce two new proxies for how consistent and honest letters portray the fund. The first proxy is a dummy variable, that we label 'Inconsistency Dummy'. It takes on the value one if a fund's past performance (measured as return over the reporting period) is below (above) the median of all funds in the same segment over the same period and the negativity of the letter at the same time is also below (above) the median of all funds in the same segment at the same time, i.e., the proxy is one if the fund sends a more positive report despite its performance being subpar. We then regress flows on negativity, our inconsistency proxy, as well as an interaction term. Results are shown in Table 6.

Results in column (1) are based on the CRSP sample. The impact of the Inconsistency Dummy is negative and highly significant: a letter whose tone is inconsistent with its performance leads to netflows that are lower by 36 basis points. At the same time, the interaction term between inconsistency and tone is significantly positive, while tone itself remains neg-

[24] Guiso, Sapienza, and Zingales (2008) find trust to be an important determinant of investment decisions.

ative and highly significant. The coefficient estimates suggest that a one standard deviation less negative tone leads to an increase of flows by more than 21 basis points if letter tone and performance are consistent. This effect is reduced to less than 5 basis points for letters that are inconsistent.

In columns (2) to (4), we look at results based on N-SAR data. For net-flows we find very similar results as in column (1). Furthermore, the effects are largely consistent when looking at inflows and outflows separately. The only exception is the impact of the Inconsistency Dummy on outflows, which has the expected sign but is not significant.

While the Inconsistency Dummy captures whether the last letter is inconsistent or consistent with the fund's actual performance, we also develop a second proxy that focuses more on the honesty of the writing style over time: for each fund, we run univariate rolling regressions of negativity on the fund's reporting period return. The time window spans four years and we require at least five shareholder letters within this period.[25] The beta from these regressions, which we label 'Dishonesty Beta', is our proxy for dishonesty. Thus, a lower beta suggests a more honest writing style over time in the sense that LMD^- is higher when performance is lower. If investors punish dishonesty, we should see a negative impact on flows. The Dishonesty Beta is standardized to unit variance.

In Panel B we show regression results of long-term flows over the half year period after the filing month on Dishonesty Beta and controls.[26] Besides our standard set of controls, we additionally add the average return and the standard deviation of returns over the period over which our Dishonesty Beta is computed. We find a significantly negative influence of the Dishonesty Beta on flows for both, our CRSP sample as well as based on net-flows in the N-SAR sample. We also see a significantly negative effect on inflows and an (insignificant) positive coefficient estimate for the impact on outflows.

[25]The length of the time period for our rolling regressions results from a trade-off between reliably estimating managers' honesty (more letters) and minimizing overlap in the honesty measure (fewer letters) over time. Using our standard specification with double-clustered standard errors at the fund and time dimension, we account for overlapping variables.

[26]Specifically, we look at the period from month $t+1$ to $t+6$ ($t+7$) if the report was filed in the first (second) half of month t.

Overall, the results from Table 6 confirm the view that investors have a distaste for dishonest letters, whose writing style does not reflect the actual performance in the short-term (Panel A) and long-term (Panel B). Investors' appreciation for honest communication—combined with the potential legal risks for making too bold claims—are important reasons for why we do not observe that all funds always write extremely positive letters. Generally, our results hitherto show that it makes a lot of sense for fund companies to carefully draft their letters to portray a positive but realistic picture.

4 Letter tone and fund manager behavior

The results in the previous section show that mutual fund investors react to the way in which shareholder letters are written. We now investigate whether the writing styles of shareholder letters are informative for future performance and investment behavior of fund managers.

4.1 Future fund performance

We first investigate the predictive power of shareholder letters for future fund performance and relate various performance measures as dependent variable to letter tone as captured by LMD^{-} as well as control variables. As performance measures, we use the 6-month performance over the period starting with the month after the filing month, i.e., from t+1 to t+6, based on (1) raw returns, (2) CAPM 1-factor alphas, and (3) Carhart (1997) 4-factor alphas.[27] For example, if a fund files a shareholder letter in January 2010, we define performance over the period February 2010 to July 2010 (that is, until the next shareholder letter is sent out).

As control variables we include fund and fund family returns as well as fund and fund family flows over the twelve months before the fiscal (half-)year end. We also add fund size,

[27]Examining excess returns over the market is equivalent to using raw returns in regressions including month fixed effects. As all of our regressions include month fixed effects, we, thus, do not investigate market excess returns separately.

family size, fund age, and funds' expense ratios. To control for letter characteristics, we further include its length and the time difference between the letter's report and filing date. Finally, we include report month and filing month fixed effects. Including fund fixed effects in fund performance regressions that also include fund size can lead to biased estimates (Pástor, Stambaugh, and Taylor (2015)). Thus, in our performance regressions, we use the recursive demeaning estimator of Pástor, Stambaugh, and Taylor (2015) which recursively forward-demeans all variables. We instrument for forward-demeaned negativity, letter length, fund size, lagged 1-year fund returns and flows using the corresponding backward-demeaned variables. Standard errors are double-clustered at the fund and time dimension. Double-clustering standard errors at the fund family and time dimension to account for the fact that some letters cover multiple funds does not change our results (see Table IA-17 in the Internet Appendix).

Results are presented in Table 7. In Panel A, we focus on overall negativity and find insignificant coefficient estimates in columns (1) and (3), i.e., letter tone is not a predictor of future fund returns or Carhart (1997) 4-factor alphas. However, the impact of negativity on the future CAPM alpha is significantly negative, i.e., a more negative tone predicts worse CAPM alphas.

As we are interested in predicting managers' future performance and behavior, we next differentiate between their opinions about the future and their discussion of the present and past. For instance, a manager warning investors "that there will be many short-term disappointments and challenges" might have fewer good investment ideas and will follow a different investment strategy than a manager who looks optimistically to the future.[28] Thus, we rerun our regressions from Panel A but split negativity in its forward-looking and non-forward-looking component. We also control for the percentage of forward-looking sentences.

[28]Section 4 of our Internet Appendix shows examples of negative forward-looking and non-forward-looking sentences. There is large variation in the amount and tone of forward-looking information (see Table 1).

In all specifications, we find an insignificant impact of forward-looking negativity on future performance and of the fraction of forward-looking sentences. The same holds true for non-forward-looking negativity, with the exception of a significantly negative impact on the CAPM alpha in column (2), which also drives the negative impact of overall negativity on the CAPM alpha in Panel A. While the results from column (2) would be consistent with managers' more negative description of the past predicting lower future performance, we caution against over-interpreting this finding, as we only find significant results for one out of three performance measures.

4.2 Future risk taking

In the next step, we examine whether shareholder letters offer further insights regarding the future investment behavior of mutual fund managers by focusing on their risk taking behavior. A recent report by the ICI (Holden, Bogdan, and Schrass (2019)) highlights that more than 90% of all mutual fund investors care about the risk-level of their fund investment. Thus, predictive power of shareholder letters regarding future risk taking would be helpful for investors to learn about the way in which they can expect their funds to be managed.

We start by analyzing whether the tone of a shareholder letter predicts mutual fund managers' risk-taking behavior. We conjecture that the tone in which a letter is written proxies for the pessimism and eventually the risk-aversion of the manager. If a fund manager uses a more negative tone, we thus expect her to take less risk. Furthermore, Scharfstein and Stein (1990) show that risk averse managers herd more towards the market, i.e., we particularly expect low levels of idiosyncratic risk taking if letter tone is more negative.

In our empirical model, we relate managerial risk-taking to the tone of the preceding shareholder letter. As dependent variables we use three proxies for fund managers' risk taking measured over the six-month period starting one month after the filing month: (1) funds' total risk, defined as the standard deviation of daily fund returns, (2) systematic risk, computed as the market beta in the Carhart (1997) 4-factor model, and (3) idiosyncratic risk,

calculated as the standard deviation of the residuals from the same model. We include our standard set of controls including the one-year fund and fund family return. It is important to control for past performance, as the realized performance in the last period might lead to a change in managerial behavior because of strategic risk-taking incentives (see, e.g., Brown, Harlow, and Starks (1996), Kempf and Ruenzi (2008)). Results are presented in Panel A of Table 8.

While the estimates for the impact of negativity on total fund risk and systematic risk are not significant at conventional levels (columns (1) and (2)), there is a clearly significant negative relation between tone and idiosyncratic risk (column (3)).[29] That is, a more negative letter predicts significantly less idiosyncratic risk taking in the subsequent six months. This finding suggests that fund managers deviate from the market less and take fewer active bets if the tone of their last shareholder letter is more negative, which is consistent with risk-averse managers herding more towards the market (Scharfstein and Stein (1990)). Also, the result is in line with evidence from psychology suggesting that an individual's linguistic style provides a meaningful way to predict behavior (e.g., Pennebaker and King (1999)). For example, in a corporate setting, Chatterjee and Hambrick (2007) show that CEOs' language styles are associated with their risk taking.

Funds with higher idiosyncratic risk may also follow more unconventional investment styles. Thus, we investigate whether letter tone predicts a fund's style extremity. We focus on the most important investment styles, i.e., size, book-to-market, and momentum, and follow Baer, Kempf, and Ruenzi (2011) by computing the deviations of a fund's factor-weightings with respect to the three style factors from the average across all funds in the same segment and year. A higher style extremity corresponds to a more extreme weighting of the respective style, i.e., to a more extreme style of this fund as compared to the average fund in its segment. We define the style extremity measure at the fund level as the average of the deviations across the three style factors. Results in column (4) show that the coefficient

[29] Double-clustering standard errors at the fund family and time dimension does not change our results (see Table IA-18 in the Internet Appendix).

on negativity is significantly negative at the 5% level, i.e., a more negative shareholder letter also predicts lower future style extremity.

Taken together, these results show that shareholder letters provide useful information for fund investors and allow them to learn more about the future risk profile of their fund investment.

5 Conclusion

We investigate whether the writing style of shareholder letters affects mutual fund investor behavior and whether shareholder letters are predictive for future performance and managerial behavior. Our results show that mutual fund investors indeed react to the writing style of shareholder letters. Specifically, we find strong evidence that a more negative tone of these letters leads to lower net flows. This effect is much stronger for younger funds than for older funds, where investors can rely on a longer performance history. We also find that investors value a honest writing style that is consistent with the actual performance of a fund and that the impact of letter tone is stronger for such funds.

Overall, these findings suggests that at least some mutual fund investors pay close attention to the tone and the content of shareholder letters when making their investment decisions. Consequently, fund companies can use shareholder letters as a tactical instrument to influence short-term flows.

Our paper also shows that shareholder letters can be informative for investors of actively managed mutual funds: while we find no stable predictive relationship between letter tone and future performance, we document that a negative writing style is predictive of less daring investment styles.

References

Abis, S., 2022, "Man vs. machine: Quantitative and discretionary equity management," Working Paper.

Abis, S., and A. Lines, 2021, "Do mutual funds keep their promises?," Working Paper.

Baer, M., A. Kempf, and S. Ruenzi, 2011, "Is a Team Different from the Sum of its Parts? Teams Management in the Mutual Fund Industry," *Review of Finance*, 15, 359–396.

Barber, B., T. Odean, and L. Zheng, 2005, "Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows," *Journal of Business*, 78, 2095–2119.

Barber, B. M., X. Huang, and T. Odean, 2016, "Which Factors Matter to Investors? Evidence from Mutual Fund Flows," *The Review of Financial Studies*, 29(10), 2600–2642.

Berk, J., and J. van Binsbergen, 2016, "Assessing Asset Pricing Models Using Revealed Preference," *Journal of Financial Economics*, 119, 1–23.

Brown, K. C., W. V. Harlow, and L. T. Starks, 1996, "Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry," *Journal of Finance*, 51, 85–110.

Carhart, M. M., 1997, "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57–82.

Chatterjee, A., and D. C. Hambrick, 2007, "It's all about me: Narcissistic Chief Executive Officers and Their Effects on Company Strategy and Performance," *Administrative Science Quarterly*, 52, 351–386.

Chevalier, J., and G. Ellison, 1997, "Risk Taking by Mutual Funds as a Response to Incentives," *Journal of Political Economy*, 105(6), 1167–1200.

Chu, Y., and H. H. Kim, 2019, "Pessimistic fund managers," Working Paper.

Cohen, L., C. Malloy, and Q. Nguyen, 2020, "Lazy prices," *The Journal of Finance*, 75(3), 1371–1415.

Coval, J., and E. Stafford, 2007, "Asset fire sales (and purchases) in equity markets," *Journal of Financial Economics*, 86, 479–512.

DeHaan, E., Y. Song, C. Xie, and C. Zhu, 2021, "Obfuscation in mutual funds," *Journal of Accounting and Economics*, 72(2-3), 101429.

Del Guercio, D., and P. A. Tkac, 2008, "Star Power: The Effect of Morningstar Ratings on Mutual Fund Flows," *Journal of Financial and Quantitative Analysis*, 43, 907–936.

Evans, R. B., and Y. Sun, 2021, "Models or stars: The role of asset pricing models and heuristics in investor risk adjustment," *The Review of Financial Studies*, 34(1), 67–107.

Gallaher, S. T., R. Kaniel, and L. Starks, 2015, "Advertising and Mutual Funds: From Families to Individual Funds," Working Paper.

Guiso, L., P. Sapienza, and L. Zingales, 2008, "Trusting the Stock Market," *Journal of Finance*, 63(6), 2557–2600.

Holden, S., M. Bogdan, and D. Schrass, 2019, "What US Households Consider When They Select Mutual Funds," *ICI Research Perspective*, 25(3), 1–12.

Huang, J. C., K. D. Wei, and H. Yan, 2012, "Investor learning and mutual fund flows," Working Paper.

Hwang, B.-H., and H. H. Kim, 2017, "It pays to write well," *Journal of Financial Economics*, 124(2), 373–394.

Jain, P. C., and J. S. Wu, 2000, "Truth in Mutual Fund Advertising: Evidence on Future Performance and Fund Flows," *Journal of Finance*, 55, 937–958.

Kaniel, R., L. T. Starks, and V. Vasudevan, 2007, "Headlines and Bottom Lines: Attention and Learning Effects from Media Coverage of Mutual Funds," Working Paper.

Kempf, A., and S. Ruenzi, 2008, "Tournaments in mutual-fund families," *The Review of Financial Studies*, 21(2), 1013–1036.

Kostovetsky, L., and J. B. Warner, 2020, "Measuring innovation and product differentiation: Evidence from mutual funds," *The Journal of Finance*, 75(2), 779–823.

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, 2281–2321.

Li, F., 2010, "The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach," *Journal of Accounting Research*, 48(5), 1049–1102.

Loughran, T., and B. McDonald, 2011, "When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *Journal of Finance*, 66, 35–65.

——— , 2016, "Textual analysis in accounting and finance: A survey," *Journal of Accounting Research*, 54(4), 1187–1230.

Niessen-Ruenzi, A., and S. Ruenzi, 2019, "Sex Matters: Gender Bias in the Mutual Fund Industry," *Management Science*, 65(7), 3001–3025.

Pástor, L., R. F. Stambaugh, and L. A. Taylor, 2015, "Scale and skill in active management," *Journal of Financial Economics*, 116(1), 23–45.

Pennebaker, J. W., and L. A. King, 1999, "Linguistic styles: Language use as an individual difference," *Journal of Personality and Social Psychology*, 77, 1296–1312.

Rogers, J. L., A. van Buskirk, and S. L. C. Zechman, 2011, "Disclosure Tone and Shareholder Litigation," *The Accounting Review*, 86, 2155–2183.

Scharfstein, D. S., and J. C. Stein, 1990, "Herd behavior and investment," *American Economic Review*, 80, 465–479.

Sheng, J., N. Xu, and L. Zheng, 2022, "Do Mutual Funds Walk the Talk? A Textual Analysis of Risk Disclosure by Mutual Funds," Working Paper.

Sialm, C., L. T. Starks, and H. Zhang, 2015, "Defined Contribution Pension Plans: Sticky or Discerning Money?," *Journal of Finance*, 70, 805–838.

Sirri, E. R., and P. Tufano, 1998, "Costly Search and Mutual Fund Flows," *Journal of Finance*, 53, 1589–1622.

Solomon, D. H., E. Soltes, and D. Sosyura, 2014, "Winners in the spotlight: Media coverage of fund holdings as a driver of flows," *Journal of Financial Economics*, 113, 53–72.

Spiegel, M., and H. Zhang, 2013, "Mutual fund risk and market share-adjusted fund flows," *Journal of Financial Economics*, 108(2), 506–528.

Tetlock, P., 2007, "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *Journal of Finance*, 62, 1139–1168.

Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy, 2008, "More Than Words: Quantifying Language to Measure Firms' Fundamentals," *Journal of Finance*, 63, 1437–1467.

Wermers, R., 2003, "Is Money Really 'Smart'? New Evidence on the Relation Between Mutual Fund Flows, Manager Behavior, and Performance Persistence," Working Paper.

Figure 1: Average negativity over time

This figure plots the average negativity of shareholder letters and the S&P500 Index return over time. The blue line shows the average shareholder letter negativity according to the Loughran and McDonald (2011) dictionary of negative words. The sample period corresponds to the reporting dates of the shareholder letters in our sample and ranges from November 2005 to October 2021. The red dashed line shows the S&P 500 returns over the six-month reporting period.

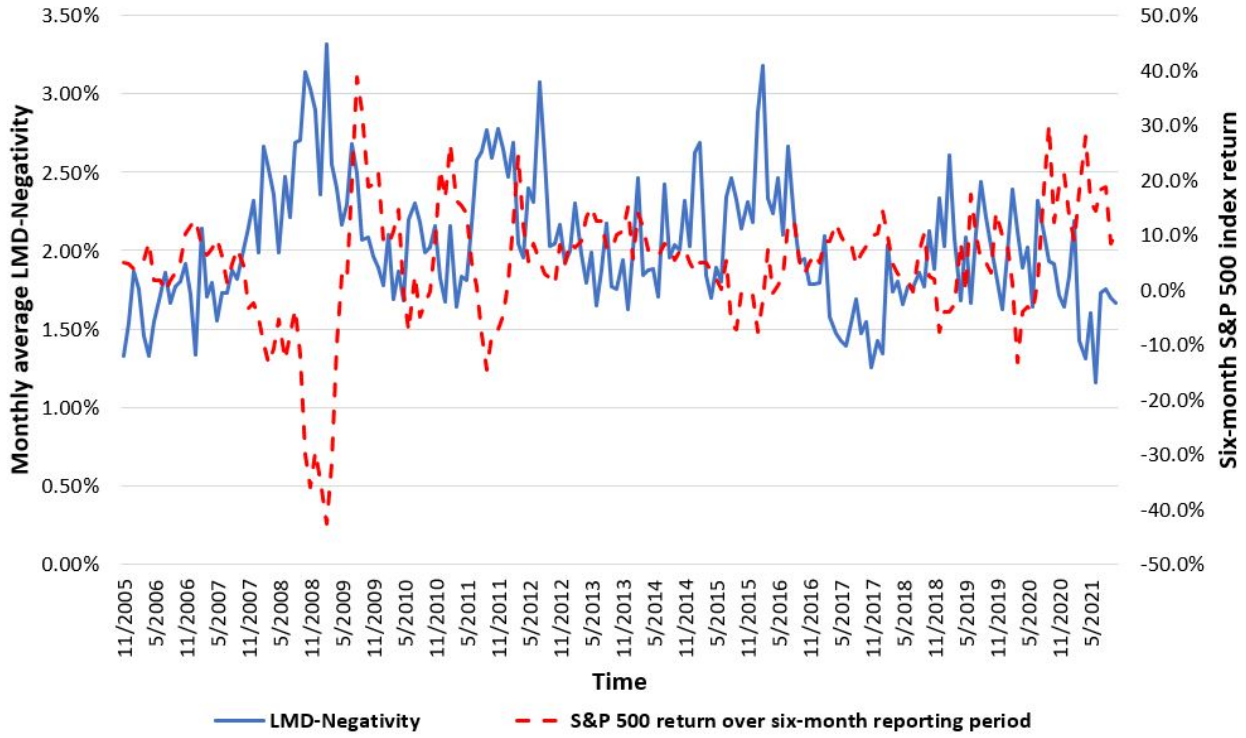

Figure 2: Cumulative daily flows after shareholder letter filing date

This figure shows cumulative flows from 10 days before up to 20 days after the filing date (day 0) of a shareholder letter separately for letters with positive (blue solid line) and negative (red dashed line) orthogonalized tone. Tone is measured by the fraction of negative words in the shareholder letter based on the Loughran and McDonald (2011) dictionary. We orthogonalize LMD^{-} in a regression and then use the residual from that regression as orthogonalized tone measure. The orthogonalized tone regression corresponds to column (1) in Table IA-6 of the Internet Appendix. The two groups are obtained by a median split of the letters based on orthogonalized $LMD^{-}_{adj.}$ tone. For better readability, we plot daily abnormal net fund flows defined as the difference between a fund's net flow and the average net flow of all funds on that event day.

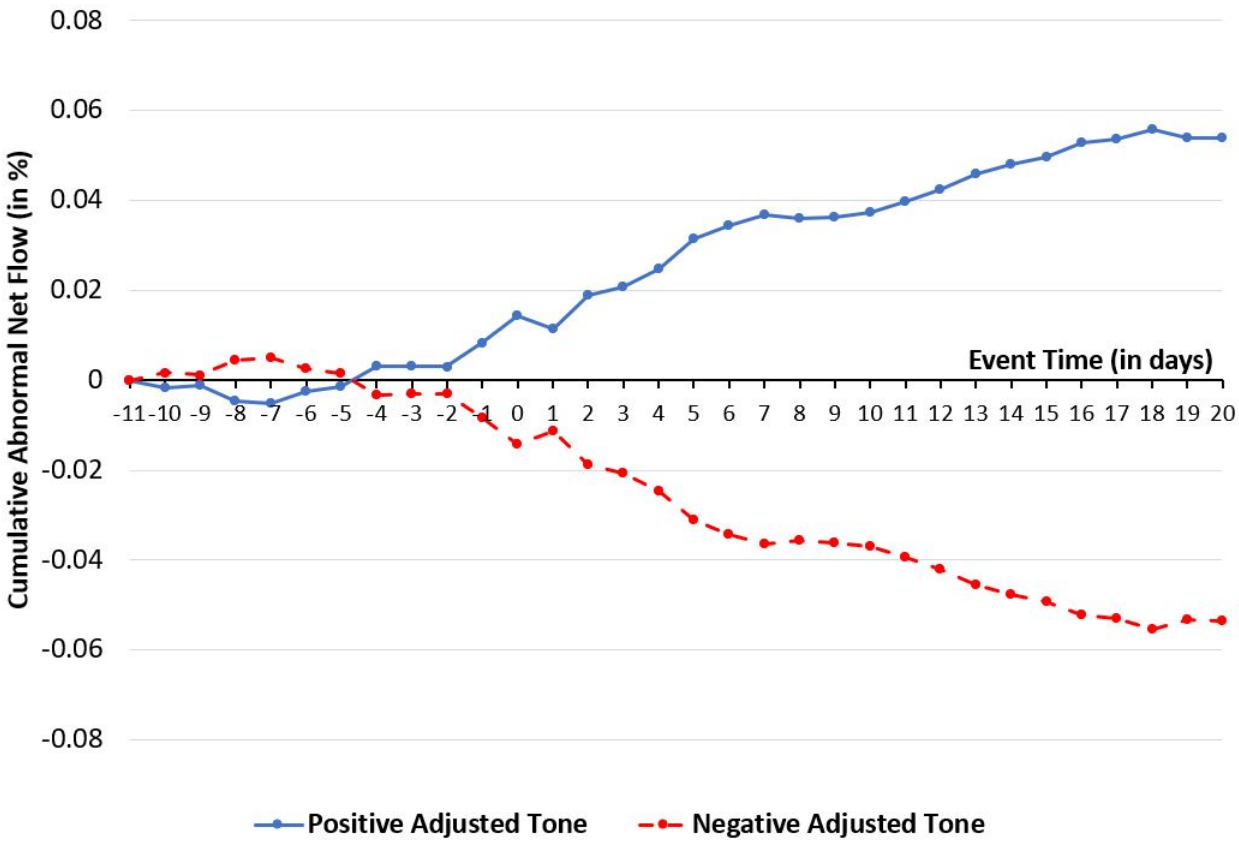

Table 1: Summary statistics

This table shows summary statistics (mean, standard deviation (sd), median (p50), 1st percentile (p1), 99th percentile (p99), and number of observations (N)) of shareholder letters (Panel A) and fund characteristics (Panel B). LMD^- is the fraction of negative words according to the Loughran and McDonald (2011) dictionary. In the second and third row, we distinguish between the negativity in forward-looking and non-forward-looking sentences. Forward-looking sentences are defined as in Li (2010). Number of words is the number of words of the shareholder letter. Time difference is the number of days between the fiscal (half-) year end date (report date) and the date when the document is filed with the SEC (filing date). All fund characteristics are defined in detail in Section 3 of the Internet Appendix. The sample includes all actively managed US open-end equity funds with CRSP and SEC data. The sample period is from 2006 to 2021.

Variable	mean	sd	p50	p1	p99	N
Panel A: Shareholder letters						
LMD^- (%)	2.002	1.098	1.944	0.000	5.055	35,437
Forward-looking LMD^- (%)	1.574	1.548	1.316	0.000	7.407	35,437
Non-forward-looking LMD^- (%)	2.086	1.192	2.011	0.000	5.391	35,437
Forward-looking sentences (%)	20.342	12.410	19.643	0.000	55.738	35,437
Number of words	874.599	635.477	711.000	113.000	3039.000	35,437
Time difference	62.789	5.284	63.000	50.000	71.000	35,437
Panel B: Fund characteristics						
Fund Flow (%)	-0.283	4.804	-0.556	-15.713	17.430	35,437
Fund Size	2362.593	9549.371	356.400	4.000	41437.398	35,437
Fund Age	17.188	14.111	14.083	1.333	75.083	35,437
Expense Ratio (%)	1.157	0.399	1.128	0.224	2.336	35,437
Fund Risk (%)	4.518	2.175	4.069	1.131	11.327	35,437
6-month Return (%)	4.611	13.186	5.301	-35.301	37.754	35,437
1-year Return (%)	9.481	19.077	10.298	-41.736	58.783	35,437
$\text{Return}_{t+1,t+6}$ (%)	4.096	13.359	4.902	-36.392	38.399	34,528
4-Factor $\text{Alpha}_{t+1,t+6}$ (%)	-0.836	5.746	-0.750	-17.554	16.821	34,528
Total $\text{Risk}_{t+1,t+6}$ (%)	1.151	0.568	0.993	0.389	3.103	34,047
Systematic $\text{Risk}_{t+1,t+6}$	0.896	0.193	0.943	0.249	1.227	34,015
Idiosyncratic $\text{Risk}_{t+1,t+6}$ (%)	0.382	0.240	0.308	0.102	1.196	34,029
Style Extremity	1.039	0.642	0.894	0.171	3.373	32,944

Table 2: Shareholder letter tone and monthly fund flows

This table shows regressions of monthly fund flows on shareholder letter tone and various fund characteristics. In Panel A, the dependent variable is the net fund flow (in %) in the month of the SEC filing. In Panel B, the dependent variables are the net fund flows (columns (1) and (2)), inflows (columns (3) and (4)), and outflows (columns (5) and (6)) obtained from NSAR and NPORT filings. In both panels, we replace the flows of the filing month by the flows of the subsequent month whenever the filing date is after the 15th calendar day. Tone is the fraction of negative words in the shareholder letter based on the Loughran and McDonald (2011) (LMD^{-}) dictionary and is standardized to unit variance. In column (2) of Panel A, we add the fund's CAPM alpha. In columns (3) and (4) of Panel A, we include the fund's return rank and squared return rank in their investment objective and use piece-wise linear regressions like Sirri and Tufano (1998), respectively. All control variables are defined in detail in Section 3 of the Internet Appendix. All regressions include fund fixed effects, time fixed effects for the month of the fiscal (half-) year end (report month) and combined investment objective and filing month fixed effects. Investment objectives are based on the CRSP/Lipper objective codes. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 2: Shareholder letter tone and monthly fund flows (cont'd)

Panel A: Negativity and net flows Dependent Variable	Flow Filing Month			
	(1)	(2)	(3)	(4)
LMD^{-}	-0.134***	-0.131***	-0.111***	-0.111***
	(-3.62)	(-3.55)	(-3.07)	(-3.08)
$\text{Return}_{Report,Filing}$	12.936***	13.004***	12.903***	12.905***
	(11.98)	(12.08)	(12.01)	(12.02)
6-month Return	4.730***	4.470***	3.793***	3.714***
	(5.98)	(5.61)	(5.15)	(5.09)
1-year Return	3.288***	0.843		
	(5.44)	(0.79)		
1-year 1F-Alpha		42.258***		
		(3.29)		
1-year Return Rank			0.875**	
			(2.11)	
1-year Return Rank^2			0.630	
			(1.57)	
1-year Bottom Quintile				1.258*
				(1.85)
1-year Mid Quintiles				1.404***
				(8.17)
1-year Top Quintile				2.601***
				(3.43)
1-year Return Family	0.334	0.251	0.026	-0.006
	(0.54)	(0.41)	(0.04)	(-0.01)
1-year Flow	2.161***	2.139***	2.100***	2.096***
	(15.81)	(15.64)	(15.26)	(15.24)
1-year Flow Family	0.699***	0.681***	0.691***	0.694***
	(2.72)	(2.65)	(2.70)	(2.71)
ln(Fund size)	-0.931***	-0.922***	-0.906***	-0.906***
	(-12.89)	(-12.81)	(-12.61)	(-12.60)
ln(Family size)	0.024	0.021	0.023	0.023
	(0.33)	(0.30)	(0.32)	(0.31)
ln(fund age)	-0.593***	-0.611***	-0.623***	-0.624***
	(-3.54)	(-3.64)	(-3.75)	(-3.76)
Expense ratio	-92.739***	-93.072***	-95.016***	-95.104***
	(-3.19)	(-3.20)	(-3.29)	(-3.29)
Fund Risk	-7.119	-5.952	-6.361	-6.276
	(-1.64)	(-1.33)	(-1.43)	(-1.40)
ln(Words)	0.088	0.088	0.084	0.085
	(1.49)	(1.48)	(1.43)	(1.43)
Time Difference	-0.003	-0.003	-0.003	-0.003
	(-0.20)	(-0.26)	(-0.26)	(-0.25)
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.191	0.192	0.195	0.195
Observations	33,398	33,398	33,398	33,398

Table 2: Shareholder letter tone and monthly fund flows (cont'd)

Panel B: Negativity and inflows and outflows						
Dependent Variable	Net Flows		Inflows		Outflows	
	(1)	(2)	(3)	(4)	(5)	(6)
LMD^{-}	-0.110***	-0.085***	-0.057**	-0.045*	0.063***	0.051***
	(-3.89)	(-3.02)	(-2.26)	(-1.81)	(3.17)	(2.63)
Controls Panel A Column (1)	Y	N	Y	N	Y	N
Controls Panel A Column (4)	N	Y	N	Y	N	Y
Fund FE	Y	Y	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.267	0.274	0.386	0.391	0.330	0.335
Observations	22,881	22,881	23,098	23,098	22,885	22,885

Table 3: Robustness and Interactions

This table shows regressions of monthly fund flows on shareholder letter tone and various fund characteristics. The dependent variable is the net fund flow (in %) in the month of the SEC filing. We replace flows of the filing month by flows of the subsequent month whenever the filing date is after the 15th calendar day. In Panel A, we perform subsample tests by analyzing the period from 2006 to 2013 (column (1)) and from 2014 to 2021 (column (2)) and by excluding letters that are filed during the financial crisis from September 2008 until March 2009 (column (3)) and during the Covid-19 crisis from February 2020 until April 2020 (column (4)). In column (5) ((6)), we add the 1-factor alpha rank and squared rank (4-factor alpha rank and squared rank), dummies for the fund's Morningstar star rating, and the 6-month flow during the reporting period as additional controls. In column (7), we add the cosine similarity between the fund's current shareholder letter and its shareholder letter from the previous year. Cosine similarity is computed as in Cohen, Malloy, and Nguyen (2020). In column (8), we replace percentage flows by funds' market share (see Spiegel and Zhang (2013)). In Panel B, we analyze the interaction effects between the negativity of the shareholder letter (LMD^{-}) and 6-month returns (column (1)), 12-month returns (column (2)), a dummy for negative 6-month returns (column (3)), a dummy for negative 12-month returns (column (4)), a dummy that is one if the shareholder letter is part of the annual N-CSR filing (but not part of the semiannual N-CSRS filing; column (5)), a dummy for no-load funds (column (6)), retail funds (column (7)), and young funds (i.e., below median fund age; column (8)). In both panels, all control variables and fixed effects from column (1) in Panel A of Table 2 are always included. All variables are defined in detail in Section 3 of the Internet Appendix. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3: Robustness and Interactions (cont'd)

Panel A: Alternative specifications and robustness								
Dependent Variable	Flow Filing Month							Market Share
Specification	2006-2013	2014-2021	excluding Fin. Crisis	excluding Covid	Further controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LMD^{-}	-0.165***	-0.122**	-0.139***	-0.126***	-0.104***	-0.103***	-0.150***	-0.002**
	(-2.96)	(-2.02)	(-3.68)	(-3.42)	(-2.94)	(-2.90)	(-3.48)	(-2.24)
1-year 1F-Alpha Rank					0.517			
					(1.10)			
1-year 1F-Alpha Rank^2					0.317			
					(0.78)			
1-year 4F-Alpha Rank						0.428		
						(1.04)		
1-year 4F-Alpha Rank^2						0.410		
						(1.03)		
1 Star rating					-1.204***	-1.196***		
					(-8.14)	(-8.08)		
2 Star rating					-0.494***	-0.488***		
					(-5.43)	(-5.42)		
3 Star rating					-0.141*	-0.139*		
					(-1.70)	(-1.68)		
4 Star rating					0.640***	0.631***		
					(6.88)	(6.77)		
5 Star rating					1.624***	1.607***		
					(10.25)	(10.09)		
6-month Flow					4.574***	4.574***		
					(10.14)	(10.12)		
Textual similarity							0.557	
							(1.22)	
Controls Table 2 Panel A Column (1)	Y	Y	Y	Y	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.209	0.188	0.192	0.191	0.213	0.213	0.181	0.179
Observations	16,344	16,782	32,000	32,835	33,398	33,398	26,221	33,398

Table 3: Robustness and Interactions (cont'd)

Panel B: Interaction effects Dependent Variable	Flow Filing Month							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LMD$^-$	-0.133***	-0.136***	-0.140***	-0.137***	-0.127***	-0.129**	-0.167**	-0.079**
	(-3.59)	(-3.63)	(-3.21)	(-3.33)	(-2.61)	(-2.32)	(-2.35)	(-2.01)
LMD$^-$ x 6-month Return	-0.033							
	(-0.14)							
LMD$^-$ x 1-year Return		0.055						
		(0.30)						
LMD$^-$ x Negative Return Dummy$_{6month}$			0.017					
			(0.30)					
LMD$^-$ x Negative Return Dummy$_{12month}$				0.013				
				(0.23)				
LMD$^-$ x N-CSR					-0.012			
					(-0.22)			
LMD$^-$ x No-Load Fund						0.004		
						(0.04)		
LMD$^-$ x Retail Fund							0.051	
							(0.67)	
LMD$^-$ x Young fund								-0.152**
								(-2.00)
Negative Return Dummy$_{6month}$			-0.378**					
			(-2.30)					
Negative Return Dummy$_{12month}$				-0.304*				
				(-1.90)				
N-CSR					0.048			
					(0.37)			
No-Load Fund						-0.345		
						(-0.90)		
Retail Fund							0.149	
							(0.68)	
Young Fund								0.201
								(1.03)
Controls Table 2 Panel A Column (1)	Y	Y	Y	Y	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.191	0.191	0.192	0.191	0.191	0.184	0.192	0.191
Observations	33,398	33,398	33,398	33,398	33,398	20,393	33,237	33,398

Table 4: Investors' reaction to forward and non-forward-looking letter content

This table shows regressions of monthly fund flows on forward and non-forward-looking negativity of shareholder letters and various fund characteristics. In column (1), the dependent variable is the net fund flow (in %) in the month of the SEC filing from the CRSP database. In columns (2) to (4), we use the net fund flows, inflows, and outflows from NSAR and NPORT filings, respectively. We replace flows of the filing month by flows of the subsequent month whenever the filing date is after the 15th calendar day. We distinguish between negativity according to the Loughran and McDonald (2011) dictionary in forward-looking and non-forward-looking sentences. Forward-looking sentences are defined as in Li (2010). Control variables and fixed effects from column (1) of Panel A in Table 2 are always included in the regressions and defined in detail in Section 3 of the Internet Appendix. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	CRSP	NSAR/NPORT		
	net flows	net flows	inflows	outflows
	(1)	(2)	(3)	(4)
Forward-looking LMD^-	-0.055*	-0.011	-0.006	0.005
	(-1.66)	(-0.48)	(-0.34)	(0.35)
Non-forward-looking LMD^-	-0.105***	-0.107***	-0.051**	0.065***
	(-2.78)	(-3.77)	(-2.05)	(3.57)
Forward-looking sentences	-0.663**	-0.216	-0.000	0.172
	(-2.31)	(-1.22)	(-0.00)	(1.10)
Controls Table 2 Panel A Column (1)	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.191	0.267	0.386	0.330
Observations	33,398	22,881	23,098	22,885

Table 5: Letter tone and fund flows - daily fund flows

This table shows regressions of daily fund flows on shareholder letter tone and various fund characteristics. In columns (1) to (7), the dependent variables are the daily net fund flows (in %) from days t-15 to t-11, t-10 to t-6, t-5 to t-1, on day t (filing date), from days t+1 to t+5, t+6 to t+10, and t+11 to t+15, respectively. Tone is the fraction of negative words based on the LMD^- dictionary. LMD^- is standardized to unit variance. Control variables and fixed effects from column (4) of Panel A in Table 2 are always included in the regressions and defined in detail in Section 3 of the Internet Appendix. Additionally, we control for fund manager changes, fund name changes, changes in funds' investment objective, whether a share class of the fund has been liquidated, whether a share class has been merged into the fund, and whether there was an expense ratio change and the size of the change. These variables are based on changes from seven months before to one month after the filing month. We also include the flows to the fund's investment objective during the corresponding 1-day/5-day period. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Daily fund flows from	t-15;t-11	t-10;t-6	t-5;t-1	t	t+1;t+5	t+6;t+10	t+11;t+15
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LMD^-	-0.009	0.003	-0.024**	-0.002	-0.026***	-0.024**	-0.015
	(-0.90)	(0.30)	(-2.21)	(-0.83)	(-2.83)	(-2.30)	(-1.50)
Manager Change Dummy	-0.006	-0.028	-0.013	0.000	-0.039**	-0.035	-0.033*
	(-0.36)	(-1.48)	(-0.76)	(0.05)	(-2.27)	(-1.53)	(-1.84)
Fund Name Change Dummy	0.042	0.034	-0.017	-0.008	0.024	0.047	0.021
	(1.60)	(1.26)	(-0.62)	(-1.13)	(0.88)	(1.50)	(0.82)
Investment Obj. Change Dummy	-0.041	-0.035	0.048	0.003	-0.203**	-0.013	0.075
	(-0.48)	(-0.48)	(0.66)	(0.18)	(-2.45)	(-0.17)	(0.93)
Share Class Liquidated Dummy	-0.044	-0.075	-0.035	-0.018	-0.095*	-0.032	-0.045
	(-0.83)	(-1.63)	(-0.78)	(-1.47)	(-1.78)	(-0.70)	(-1.08)
Share Class Merged Dummy	-0.050	-0.002	-0.077**	0.014	-0.027	0.006	0.024
	(-1.33)	(-0.05)	(-2.40)	(1.47)	(-0.71)	(0.15)	(0.61)
Expense Ratio Change Dummy	0.035*	-0.000	0.036	-0.004	0.023	0.000	0.040
	(1.66)	(-0.00)	(1.59)	(-0.84)	(1.27)	(0.01)	(1.64)
Expense Ratio Change	-26.698**	-7.440	-31.294**	-6.740**	-26.981***	-26.629**	-26.671**
	(-2.29)	(-0.69)	(-2.60)	(-2.27)	(-3.01)	(-2.05)	(-2.37)
Controls Table 2 Column (4)	Y	Y	Y	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y	Y	Y	Y
Contemporaneous Flow to Inv. Obj.	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.171	0.185	0.183	0.122	0.176	0.169	0.174
Observations	18,357	18,358	18,371	18,282	18,365	18,315	18,285

Table 6: Letter tone consistency, honesty and fund flows

This table shows regressions of fund flows on shareholder letter tone inconsistency and dishonesty and various fund characteristics. In Panel A (B), the dependent variable is the fund flow (in %) in the filing month (from one to six months after the filing month). We replace the flows by the flows from the subsequent month whenever the filing date is after the 15th calendar day. In column (1), the dependent variable is the net fund flow from the CRSP database. In columns (2) to (4), we use the net fund flows, inflows, and outflows from NSAR and NPORT filings, respectively. In Panel A, inconsistency is a dummy equal to one if both negativity and the fund's 6-month return rank are above the median (good performance but high negativity) or if both negativity and the fund's 6-month return rank are below the median (poor performance but low negativity). The medians are computed for each point in time within the fund's investment objective. We include negativity (LMD^-) and its interaction with the inconsistency dummy. In Panel B, dishonest reporting is measured by the relation between tone and fund performance over a fund's eight most recent letters. More specifically, for each fund, we run rolling regressions of negativity (LMD^-) on the fund's 6-month return over the previous eight letters. We require at least five shareholder letters. The beta from these regressions is our proxy for dishonesty. The dishonesty beta is standardized to unit variance. Additionally, we control for the average and the standard deviation of 6-month returns over the previous eight letters used in the beta estimation. In both panels, all controls and fixed effects from column (1) of Panel A in Table 2 are included. All variables are defined in detail in Section 3 of the Internet Appendix. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Filing month flows and consistency				
Dependent Variable	CRSP	NSAR/NPORT		
	net flows	net flows	inflows	outflows
	(1)	(2)	(3)	(4)
LMD^-	-0.223***	-0.189***	-0.102***	0.095***
	(-4.68)	(-5.25)	(-3.44)	(3.47)
LMD^- x Inconsistency Dummy	0.191***	0.171***	0.096***	-0.069**
	(3.23)	(3.63)	(2.77)	(-2.06)
Inconsistency Dummy	-0.417***	-0.309***	-0.227***	0.091
	(-3.29)	(-3.46)	(-3.23)	(1.44)
Controls Table 2 Panel A Column (1)	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.192	0.267	0.386	0.330
Observations	33,398	22,881	23,098	22,885
Panel B: Long-term flows and honesty				
Dependent Variable	CRSP	NSAR/NPORT		
	net flows	net flows	inflows	outflows
	(1)	(2)	(3)	(4)
Dishonesty Beta	-0.314**	-0.286*	-0.325*	0.071
	(-2.15)	(-1.85)	(-1.80)	(0.50)
Average fund return	33.095***	29.319***	21.477***	-9.447*
	(6.29)	(5.54)	(3.35)	(-1.80)
Std. dev. fund return	3.165	3.876	13.073**	9.034*
	(0.65)	(0.72)	(2.16)	(1.81)
Controls Table 2 Panel A Column (1)	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.295	0.326	0.438	0.439
Observations	21,660	13,330	13,436	13,343

Table 7: Shareholder letter tone and future fund performance

This table shows regressions of fund performance (in %) on shareholder letter tone and various fund characteristics. The dependent variable is the fund's return (column (1)), CAPM 1-factor alpha (column (2)), and Carhart (1997) 4-factor alpha (column (3)). Performance is measured from one to six months after the filing date. Alphas are computed using beta coefficients obtained from a regression using daily fund returns over the previous twelve months. In Panel A, tone is the fraction of negative words based on the Loughran and McDonald (2011) dictionary. In Panel B, we split negativity into negativity in forward-looking and non-forward-looking sentences. Forward-looking sentences are defined as in Li (2010). All control variables are defined in detail in Section 3 of the Internet Appendix. LMD^- is standardized to unit variance. All regressions use the Pástor, Stambaugh, and Taylor (2015) recursive demeaning estimator which recursively forward-demeans all variables. We instrument for forward-demeaned negativity, letter length, fund size, lagged 1-year fund returns and flows using the corresponding backward-demeaned variables. Regressions include time fixed effects for the reporting and filing month. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Negativity and future fund performance			
Dependent variable	Raw Return	1-Factor Alpha	4-Factor Alpha
	(1)	(2)	(3)
LMD^-	0.007	-0.137**	-0.032
	(0.09)	(-2.09)	(-0.54)
ln(Fund size)	-1.925*	-0.529	-0.387
	(-1.80)	(-0.63)	(-0.55)
ln(Family size)	0.241	-0.058	-0.041
	(0.82)	(-0.25)	(-0.22)
ln(fund age)	1.399**	-0.465	-0.347
	(2.12)	(-1.13)	(-0.96)
Expense ratio	-255.183	31.894	48.176
	(-1.46)	(0.24)	(0.45)
1-year Return	0.003	-0.015	-0.027
	(0.09)	(-0.51)	(-1.12)
1-year Flow	-1.649**	-0.527	-0.351
	(-2.03)	(-0.88)	(-0.75)
1-year Return Family	-0.070*	-0.035	-0.009
	(-1.93)	(-1.38)	(-0.47)
1-year Flow Family	0.941	-0.058	-0.237
	(1.08)	(-0.09)	(-0.44)
ln(Words)	0.094	-0.007	0.018
	(0.41)	(-0.05)	(0.13)
Time Difference	-0.008	0.006	0.016
	(-0.29)	(0.32)	(1.02)
Recursive demeaning	Y	Y	Y
Fund FE	N	N	N
Reporting Month FE	Y	Y	Y
Filing Month FE	Y	Y	Y
Adj. R^2	0.009	0.014	0.010
Observations	34,539	34,539	34,539

Table 7: Shareholder letter tone and future fund performance (cont'd)

Panel B: Forward-looking negativity and future fund performance			
Dependent variable	Raw Return	1-Factor Alpha	4-Factor Alpha
	(1)	(2)	(3)
Forward-looking LMD^-	0.018	0.025	0.018
	(0.24)	(0.51)	(0.39)
Non-forward-looking LMD^-	-0.029	-0.159**	-0.036
	(-0.37)	(-2.54)	(-0.62)
Forward-looking sentences	-0.933	-1.218	-0.651
	(-0.90)	(-1.52)	(-0.95)
Controls from Panel A	Y	Y	Y
Recursive demeaning	Y	Y	Y
Fund FE	N	N	N
Reporting Month FE	Y	Y	Y
Filing Month FE	Y	Y	Y
Adj. R^2	0.009	0.014	0.010
Observations	34,539	34,539	34,539

Table 8: Letter tone and managerial risk-taking

This table shows regressions of fund risk and investment style variables on shareholder letter tone and various fund characteristics. Tone is the fraction of negative words based on the Loughran and McDonald (2011) (LMD^-) dictionary. LMD^- is standardized to unit variance. The dependent variable is the fund's total risk (column (1)), systematic risk (column (2)), idiosyncratic risk (column (3)), and style extremity (column (4)). Fund risk is the standard deviation of daily returns from one to six months after the filing month. Systematic (idiosyncratic) risk is the market beta (standard deviation of residuals) obtained from estimating a Carhart 4-Factor-Model using daily data from one to six months after the filing month. Total risk and idiosyncratic risk are expressed in %. Style extremity is the absolute difference of a fund's SMB-, HML-, and UMD-beta from the average beta in the fund's investment objective. Betas are calculated using daily returns over the twelve months after the filing month. For a detailed description of style extremity, see Baer, Kempf, and Ruenzi (2011). Control variables are defined in detail in Section 3 of the Internet Appendix. All regressions include fund fixed effects and time fixed effects for the month of the fiscal (half-) year end (report month) and for the month of the SEC filing (filing month). Standard errors are double-clustered at the fund and time dimension (filing month). t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable	Total Risk (1)	Systematic Risk (2)	Idiosyncratic Risk (3)	Style Extremity (4)
LMD^-	0.002	0.001	-0.002**	-0.010**
	(0.82)	(0.53)	(-2.23)	(-2.05)
ln(Fund size)	0.024***	0.016***	0.009***	0.022**
	(5.98)	(6.82)	(4.27)	(2.54)
ln(Family size)	0.003	0.003	-0.003*	-0.001
	(0.76)	(1.25)	(-1.68)	(-0.15)
ln(fund age)	-0.046***	-0.029***	-0.014**	-0.088***
	(-4.76)	(-4.88)	(-2.55)	(-3.51)
Expense ratio	5.033**	3.188**	1.928**	-0.895
	(2.30)	(2.36)	(2.22)	(-0.22)
ln(Words)	0.005	0.003	0.004**	0.007
	(1.48)	(1.42)	(2.31)	(0.86)
Time Difference	0.001	0.001*	-0.000	-0.001
	(1.45)	(1.97)	(-0.25)	(-0.39)
1-year Return	0.050	0.060**	0.047***	0.097*
	(1.44)	(2.52)	(3.11)	(1.90)
1-year Flow	-0.000	-0.006**	-0.003	-0.015
	(-0.01)	(-2.44)	(-1.28)	(-1.54)
1-year Return Family	-0.005	0.024	0.006	-0.176**
	(-0.15)	(1.13)	(0.35)	(-2.35)
1-year Flow Family	-0.019	-0.006	0.006	0.014
	(-1.41)	(-0.91)	(0.93)	(0.50)
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Filing Month FE	Y	Y	Y	Y
Adj. R^2	0.901	0.703	0.862	0.577
Observations	33,699	33,678	33,678	32,627

Internet Appendix for "Mutual Fund Shareholder Letters: Flows, Performance, and Managerial Behavior"

January 2023

Abstract

This Internet Appendix provides additional analyses, discussions, and results for the paper "Mutual Fund Shareholder Letters: Flows, Performance, and Managerial Behavior".

1 Two excerpts from letters to shareholders

This appendix shows two excerpts of shareholder letters filed in December 2008. Words that are included in the Loughran and McDonald (2011) dictionary of negative words are printed in bold font.

"Dear Investor: Thank you for taking time to review the following discussions, from our experienced portfolio management team, of the fund reporting period ended December 31, 2008. It was a time of enormous upheaval and change. We understand and appreciate the **challenges** you have faced during this historic period, and share your **concerns** about the economy, the markets, and fund holdings. To help address these issues, I'd like to provide my perspective on how we have managed–and continue to manage–your investments in these uncertain times. As a company, American Century Investments is well positioned to deal with market **turmoil**. We are financially strong and privately held, which allows us to align our resources with your long-term investment interests. In addition, our actively managed, team-based approach allows our portfolio teams to identify attractive investment opportunities regardless of market conditions. Our seasoned investment professionals have substantial experience and have successfully navigated previous market **crises**. These portfolio managers and analysts continue to use a team approach and follow disciplined investment processes designed to produce the best possible long-term results for you. For example, our equity investment teams are working closely with our fixed income group to monitor and assess credit **crisis** developments. The fixed income team anticipated dislocation in the credit markets and–through its disciplined processes and teamwork–helped reduce our exposure to investments that suffered substantial **losses**. How soon a sustainable recovery will occur is uncertain. But I am certain of this: Since 1958, we've demonstrated a consistent ability to execute solid, long-term investment strategies and the discipline to remain focused during times of **volatility** or shifts in the markets. We've stayed true to our principles, especially our belief that your success is the ultimate measure of our success. Thank you for your continued confidence in us." (**AMERICAN CENTURY QUANTITATIVE EQUITY FUNDS, INC. Small Company Fund, December 2008, percentage of negative words based on the Loughran and McDonald (2011) dictionary: 2.78%, previous 6-month return: –33.84%**)

"DEAR FELLOW SHAREHOLDERS OF VIRTUS MUTUAL FUNDS: The past year was unprecedented in the financial markets and a sobering period for most investors. And that may be the most flattering description we can give 2008. Economies across the globe were buffeted by the **severe** credit **contraction** that **destabilized** financial markets and led to bank **closures**, **failures** of financial services companies, and massive government bailouts. Corporations suffered from tightened commercial lending and a sharp drop in consumer demand, and responded with predictable **cutbacks** in employment and capital spending. The financial markets reflected the scope of these global economic **challenges**. The Dow Jones Industrial Average was down 31.9

percent in 2008, its **worst** year since 1931. The Standard & Poor's 500 index **dropped** 22 percent in the fourth quarter alone, and 37 percent for the full year - its **worst** performance since 1937. The NASDAQ market had its **worst** year ever. Investor confidence has been a major casualty of this financial **turmoil**. Many investors, paralyzed by the constant flow of **negative** news, have reacted to this extraordinary market **volatility** by **deviating** from their long-term financial plans. But just as it is **unrealistic** to base investment expectations on the market's supercharged returns from much of the 1980s and 1990s, it may be equally **misleading** to assume that future long-term results will track the market's recent **dismal** performance. While no one can predict the future, it is important to remember that the market has generally rewarded investors over the long term. Since 1927, stocks have returned 9.6 percent on average annually, and that includes the steep **decline** experienced through the end of last year. Although the near-term outlook continues to be filled with uncertainties, we believe that investors with long-term goals - such as saving for a child's college education or preparing for one's own comfortable retirement - are best served by structuring and modifying their investment program with an eye to the long-term, rather than giving **disproportionate** weight to the short-term fluctuations in the marketplace. (...) On behalf of the entire team at Virtus Investment Partners, and the investment professionals at our affiliated managers and subadvisers, I thank you for entrusting your assets to us." (**VIRTUS INSIGHT TRUST DISCIPLINED SMALL-CAP OPPORTUNITY FUND, December 2008, percentage of negative words based on the Loughran and McDonald (2011) dictionary: 4.33%, previous 6-month return: −33.22%**)

2 Details on extracting and merging shareholder letters

N-CSR/N-CSRS filings are available since 2003 but only in 2006 the SEC introduced unique share class-level (class ID) and portfolio-level (series ID) identifiers. Before 2006, there is only the Central Index Key (CIK) that can be linked to one or multiple portfolios and thus cannot be used as a unique portfolio identifier. Beginning February 6, 2006, all open-end mutual fund companies have been required to use these electronic IDs that allow identification of fund portfolios and share classes when making their filings with the SEC.[1]

There are two types of identifiers used by the SEC. Series ID is used as an identifier on the fund portfolio level, while Class ID is used as an identifier on the share class level. Since we use these portfolio identifiers as well as the ticker symbols to merge shareholder letters to the CRSP/MFLinks mutual fund data, our sample starts in 2006. Filings that do not include a Series ID are dropped from our sample.

We download the complete submission files of all N-CSR and N-CSRS filings filed with the SEC during the period from January 2006 to December 2021. From the header of the complete submission file, we identify the SEC's fund identifier (series ID), the fund name (series name), the class identifier (class ID), and the ticker symbols of all share classes. As we merge the SEC data with MFLinks/CRSP data by ticker, we can only match funds if at least one of their share classes has a ticker symbol.

After matching CRSP and SEC data via the ticker symbol and the filing date, we conduct several plausibility checks to make sure that the SEC's Series ID and the WFICN indeed correspond to the same fund portfolio. Since both identifiers are on the portfolio level, we drop observations where a single SEC Series ID is assigned to multiple WFICNs at the same point in time. Similarly, we exclude observations where a single WFICN is linked to multiple SEC Series IDs at the same point in time. During the period from 2006 to 2021, we can match 92.3% (=9,920/10,743) of all funds in the CRSP/MFLinks universe to at least one N-CSR/N-CSRS filing.

We restrict the sample to actively managed U.S. equity funds, i.e., we exclude bonds, balanced, and money market funds as well as index funds. These filters leaves us with 5,489 unique funds (WFICNs) and 91,734 fund-report observations that come from 29,530 unique filings.

A N-CSR/N-CSRS filing often covers multiple funds. In our sample, the average (median) filing contains the reports of 8.3 (6) unique funds. During our sample period from 2006 to

[1] See adopting release http://www.sec.gov/rules/final/33-8590.pdf.

2021, the average number of fund reports contained in a filing has increased from 7.5 in the first half (2006 to 2013) to 9.0 in the second half (2014 to 2021).

We extract the main document from the complete submission file that we obtain from EDGAR: To do so, we apply standard cleaning procedures as described in Loughran and McDonald (2011) and their internet appendix. For example, we remove html and xml code, re-encode characters such as   (blank space) back to their original ACSII form, and delete tables if the percentage of numbers is above 10%.

In the next step, we identify all potential shareholder letters using regular expressions for key words indicating the start and the end of shareholder letters. We also search for the titles of the sections that typically follow the shareholder letter and use them as an alternative end condition for shareholder letters.

We use the following 5 groups of key words for the start of a shareholder letter:

1. Greetings: e.g., "dear shareholder," "to our investors," "fellow shareholders"

2. Letter: e.g., "letter from the chairman," "president's message," "message from your fund manager"

3. MDA: e.g., "managers' discussion of fund performance," "management's discussion and analysis"

4. Review: e.g., "portfolio management commentary," "investment commentary," "fund performance and commentary"

5. Interviews: e.g., "Q&A with you fund manager," "how did the fund perform?"

We use the following 4 groups of key words for the end of a shareholder letter and for the beginning of a subsequent section in the filing:

1. Closings: e.g., "best regards," "sincerely," "truly yours"

2. Investment example: e.g., "hypothetical growth of a $10,000 investment," "total return based on a $10,000 investment," "comparison of change in value of $10,000 investment"

3. Expense example: e.g., "fund expense example," "understanding your fund's expenses"

4. Investment schedule: "schedule of investments," "portfolio of investments"

Then, we extract all shareholder letters of a filing going from the start match to the first subsequent end match.

To avoid missing a shareholder letter we use regular expressions that are not restricted to a few very specific phrases but allow for some flexibility in the wording. As a result, some extracted texts are not actual letters. Such mismatches happen, for example, if a reference to the shareholder letter somewhere in the filing text is incorrectly identified as the start of another shareholder letter. We exclude those mismatches by applying three filters.

First, we exclude potential letters that are implausibly short (less than 100 words) or very long (more than 4,000 words). Second, we drop potential letters where the number of words per sentence is below 10 or above 30. Loughran and McDonald (2014) report that the average number of words per sentence in Form 10-K filings is 23. Third, we exclude letters that have more than 10% numbers in the text.

After applying the cleaning procedures, we have 81,736 fund-report observations with filings that contain at least one letter that passes our filters. 9,998 fund-report observations are from filings without a valid letter. For 1,778 of these observations, we are unable to identify a start and/or an end condition of a letter using our regular expressions. For the remaining 8,220 observations, we can extract at least one letter candidate (i.e., we find a start and an end), but the extracted text does not pass our filters. 7,906 (79.1%) of observations without a letter are from the semi-annual N-CSRS filings and 2,092 (20.9%) are from the annual N-CSR filings, suggesting that annual reports are more likely to include an explicit letter than semi-annual reports.

In Panel A of Table IA-2, we analyze which fund and filing characteristics explain whether a filing includes a shareholder letter. This analysis is important to alleviate concerns about a potential selection bias of our sample. Such a bias could arise if fund managers would decide strategically to (not) include a letter depending on their past performance.

The dependent variable is a dummy that is equal to one if a filing includes at least one shareholder letter that passes our filters (number of words between 100 and 4,000, average number of words per sentence between 10 and 30, and percentage of numbers below 10%). We regress this dummy on various fund characteristics including fund performance at various horizons and fund flows, fund family performance, flows and size, as well as filing characteristics. Columns (1) and (2) include combined investment objective x time fixed effects. In columns (3) and (4), we add fund fixed effects.

In none of our specifications, any of the fund performance measures is significantly related to the letter dummy. In columns (3) and (4) the t-statistics of the return between the report and filing date, the 6-month reporting period return, and the 1-year return are all smaller than 1 (in absolute size). Thus, we do not find any evidence for fund managers strategically writing or not writing shareholder letters depending on their performance.

The 1-year fund flows are positively related to the availability of a shareholder letter indicating that – holding fund size constant – an increase in the number of fund investors may make the fund start including shareholder letters in its filings. We also find that filings that cover more funds are more likely to contain a shareholder letter. This result suggests that an investment management company may write letters if the number of shareholders receiving the report is larger, because investors of more funds are addressed.

Regarding the filing characteristics, we find that the annual N-CSR filings are much more likely to contain a letter, which is consistent with the semi-annual NCSR filings accounting for about 80% of filings without a letter in our sample.

The strongest predictor for whether a filing includes a shareholder letter is whether there was a letter in the previous year's filing. Adding this previous year letter dummy in specification (2) increases the adjusted R^2 from 0.11 to 0.49 which shows the strong persistence in including a letter in the filing. This high persistence further alleviates concerns that funds may strategically include letters based on their short-term performance.

Having identified all letters in a filing, we match them to individual funds. We therefore search for the fund names (series names) from the header of the filings in the letter texts. We start by using the full series names which usually include the investment management company (e.g., "Fidelity Small Cap Value Fund"). For all letters for which we have not found a match after searching for the full series names, we repeat the search using short series names without the investment management company name (e.g., "Small Cap Value Fund"). While this second search helps to link more letters to funds, the fund names in the header of the filling and the fund names in the text still vary substantially in many cases, for instance, due to the use of abbreviations (e.g., "SC Value Fund"). To avoid incorrect matches, we restrict our analysis to letters that we can uniquely link to a fund.

The approach results in 38,168 fund reports out of the 81,736 fund-report observations from filings with at least one shareholder letter. The high number of letters that cannot be clearly linked to a fund arises from the use of generic fund names in the text (e.g., "value fund") and from investment management companies not using the full series name in the letter which would help to handle such names. Also, some letters contain general content and do not mention any specific fund at all.

As the successful matching of letters to funds solely depends on the string matching of fund names, we expect the matching not to result in a systematic selection bias. We test this hypothesis empirically by running regressions similar to the ones in Panel A of Table IA-2. More specifically, we regress a matched letter dummy that is one if we can uniquely match a fund with a letter and zero otherwise on the same set of controls and fixed effects as in Panel A.

In our cross-sectional analysis in columns (1) and (2), we find that letters from larger funds, older funds, larger fund families, and more expensive funds are more likely to be uniquely matched with a fund. This results seems plausible as larger investment management companies may spend more time and effort in clearly structuring their shareholder reports. Large flagship funds may be presented more prominently in the filing which helps to identify the respective letter text and avoid mismatches from preceding or subsequent letters from other funds. After adding fund fixed effect in columns (3) and (4), these variables are no longer significantly related to the letter matching success.

With respect to the filing characteristics, the analysis shows that more filings in a letter are associated with a lower probability of matching the filing's letters clearly to individual funds. This result makes sense as the number of letters that are matched to multiple funds and that are, thus, no longer uniquely linked, increases with the number of potential fund names. This is especially true for generic fund names. Letters in the annual N-CSR filing are also more likely to be matched.

The strongest predictor for whether we can match a shareholder letter uniquely to a fund is again a dummy for whether the fund was matched to a letter in the previous year's filing. Adding this dummy for a matched letter in the previous year's filing increases the adjusted R^2 substantially from 0.20 to 0.70 in the specification without fund fixed effects and from 0.60 to 0.74 in the specification with fund fixed effects.

In all specifications, fund performance and flows show no significant relation to the matched letter dummy alleviating concerns about a selection bias.

As a last test, we add filing fixed effects in columns (5) and (6), i.e., we compare a fund with a matched letter to a fund without a matched letter within the same filing. This analysis helps to alleviate concerns that letters of funds with good performance might be presented in a way that makes the unique fund-letter matching more likely. In this specification, we drop all variables that have no within filing variation, i.e., variables on the investment company level and filing characteristics.

In this analysis, none of the fund characteristics shows a robust and significant relation with the matched shareholder letter dummy. The only variable that remains significant in this specification is the dummy for whether the letter in the previous year's filing was successfully matched which highlights the persistence in our ability to match a given fund with its letter.

Taken together, the results in panels A and B show that our approach to identify and extract shareholder letters and to match them with funds is unrelated to fund characteristics - most importantly to fund performance - and is, thus, unlikely to introduce a selection bias in our sample.

3 Variable description

This table defines the variables used in the empirical analysis. The data sources are:
(i) CRSP: CRSP Survivorship Bias Free Mutual Fund Database
(ii) SEC: Securities and Exchange Commission EDGAR Database
(iii) EST: Estimated by the authors
(iv) KF: Kenneth French Data Library
(v) MS: Morningstar Direct Database

Panel A: Main dependent variables

Variable name	Description	Source
Fund Flow	Computed as $(TNA_{i,t} - TNA_{i,t-1})/TNA_{i,t-1} - r_{i,t}$ where $TNA_{i,t}$ denotes fund i's total net assets (TNA) in month t and r_t denotes fund i's return in month t as reported in CRSP. The merger correction proposed in Lou (2012)) is applied. The variable is winsorized at the 1st and 99th percentile.	CRSP, EST
Inflow	The dollar inflows from funds' Form N-SAR (until 2018) and N-PORT (starting 2019) filings divided by funds' TNA (from CRSP) at the end of the previous month. The variable is winsorized at the 95th percentile.	CRSP, SEC, EST
Outflow	The dollar outflows from funds' Form N-SAR (until 2018) and N-PORT (starting 2019) filings divided by funds' TNA (from CRSP) at the end of the previous month. The variable is winsorized at the 95th percentile.	CRSP, SEC, EST
Netflow	The difference between N-SAR/N-PORT inflows and outflows. The variable is winsorized at the 2.5th and 97.5 percentile.	CRSP, SEC, EST
Daily Fund Flow	Computed as the dollar flow on day t (MS variable "Estimated Fund-Level Net Flow – aggregated from share classes (daily)") divided by the total net assets on day *t-1* (MS variable "Fund Size – aggregated from share classes (daily)"). The variable is winsorized at the 1st and 99th percentile.	MS, EST
Market Share	The change in the fund's market share from month t-1 to month t as defined in Spiegel and Zhang (2013). The variable is winsorized at the 1st and 99th percentile.	MS, EST
Long-term Flows	Long-term fund flow is the compounded monthly fund flow over the six months after the filing months. Long-term flows are truncated at the 1st and 99th percentile.	CRSP, EST
Raw Return	Funds' raw return from one month after the SEC filing to six months after the SEC filing. The variable is winsorized at the 1st and 99th percentile.	CRSP, EST
1-Factor Alpha	Performance alpha from a market model. The beta is estimated using daily fund returns over the previous twelve months. Market returns are from the Kenneth French data library. The variable is winsorized at the 1st and 99th percentile.	CRSP, KF, EST
4-Factor Alpha	Performance alpha from the Carhart (1997) 4-factor model. Betas are estimated using daily fund returns over the previous twelve months. The variable is winsorized at the 1st and 99th percentile.	CRSP, KF, EST
Fund Risk	Standard deviation of daily returns. The variable is calculated from one month to six months after the filing month. The variable is winsorized at the 1st and 99th percentile.	CRSP, EST

Panel A: Main dependent variables (cont'd)

Variable name	Description	Source
Systematic Risk	Loading on the excess return of the market in the Carhart (1997) 4-factor model. The estimation is based on daily returns from one to six months after the filing month. The variable is winsorized at the 1st and 99th percentile.	CRSP, KF, EST
Idiosyncratic Risk	Estimated as the standard deviation of the residual in the Carhart (1997) 4-factor model. The estimation is based on daily returns from one month to six months after the filing month. The variable is winsorized at the 1st and 99th percentile.	CRSP, KF, EST
Style Extremity	Style extremity is the average absolute difference of a fund's SMB-, HML-, and UMD-betas from the average betas in its investment objective. We follow the construction details of Baer, Kempf, and Ruenzi (2011). Betas are calculated using daily returns over the twelve months after the filing month. The variable is winsorized at the 1st and 99th percentile.	CRSP, KF, EST

Panel B: Main independent variables

Variable name	Description	Source
LMD^-	Fraction of negative words in a letter according to the Loughran and McDonald (2011) dictionary. Standardized to unit variance.	SEC, EST
LMD^+	Fraction of positive words (controlling for negations) in a letter according to the Loughran and McDonald (2011) dictionary. Standardized to unit variance.	SEC, EST
HVD^-	Fraction of negative words in a letter according to the Harvard IV-4 psychosocial dictionary. Standardized to unit variance.	SEC, EST
MF^-	Fraction of negative words in a letter according to our mutual fund dictionary of negative words. The construction of the dictionary is explained in Section 5 of the Internet Appendix. Table IA-16 shows the top 50 MF^- dictionary words. The variable is standardized to unit variance.	SEC, EST
$\text{LMD}^-_{tf.idf}$	Term frequency and inverse document frequency weighted negativity of a letter according to the Loughran and McDonald (2011) dictionary. The weight of negative word i is its term frequency times the log of the number of letters in the sample divided by the number of letters containing word i. The variable is standardized to unit variance.	SEC, EST
$\text{HVD}^-_{tf.idf}$	Negativity of a letter according to the Harvard IV-4 psychosocial dictionary using the term frequency and inverse document frequency weighting scheme (see above). Standardized to unit variance.	SEC, EST
$\text{MF}^-_{tf.idf}$	Negativity of a letter according to our mutual fund dictionary using the term frequency and inverse document frequency weighting scheme (see above). Standardized to unit variance.	SEC, EST
$\text{LMD}^-_{adj.}$	Residual of a regression of LMD^- on the variables from Specification (1) in Panel A of Table IA-5 in the Internet Appendix. Standardized to unit variance.	SEC, EST
Forward-looking LMD^-	The percentage of negative words according to the Loughran and McDonald (2011) dictionary in forward-looking sentences defined as in Li (2010). Standardized to unit variance.	SEC, EST

Panel B: Main independent variables (cont'd)

Variable name	Description	Source
Non-forward-looking LMD^-	The percentage of negative words according to the Loughran and McDonald (2011) dictionary in non-forward-looking sentences. Standardized to unit variance.	SEC, EST
Forward-looking Sentences	The percentage of sentences that contain forward-looking key words. Forward-looking keywords are defined as in Li (2010).	SEC, EST
Market/Economy LMD^-	The percentage of negative words according to the Loughran and McDonald (2011) dictionary in market and economy sentences. Market and economy sentences contain at least one of the key words "market' , "markets", "economy" , "economies", and "economic". The variable is standardized to unit variance.	SEC, EST
Non-market/Economy LMD^-	The percentage of negative words according to the Loughran and McDonald (2011) dictionary in non-market and economy sentences. Non-market and economy sentences contain none of the key words "market' , "markets", "economy" , "economies", and "economic". The variable is standardized to unit variance.	SEC, EST
Market/Economy Sentences	The percentage of sentences that contain at least one market/economy key word: "market' , "markets", "economy" , "economies", and "economic".	SEC, EST
Dishonesty Beta	Beta from a rolling regression for every fund of negativity (LMD^-) on the fund's 6-month reporting return using the fund's most recent eight letters. We require a minimum of five letters. The variable is standardized to unit variance.	CRSP, SEC, EST
Inconsistency Dummy	Dummy equal to one if both negativity and the fund's 6-month return rank are above the median (good performance but high negativity) or if both negativity and the fund's 6-month return rank are below the median (poor performance but low negativity). The medians are computed for each point in time within the fund's investment objective.	CRSP, SEC, EST

Panel C: Other control variables

Variable name	Description	Source
6-month (1-year) Return	Fund return over the six (twelve) months before the fiscal (half-)year end.	CRSP
6-month (1-year) Family Return	Average return of all funds in the fund family over the six (twelve) months before the fiscal (half-)year end.	CRSP, EST
1-year 1F-α (4F-α)	The fund's 1-factor (4-factor) alpha estimated using daily returns over the twelve months before the fiscal (half-) year end.	CRSP, KF, EST
1-year Return Rank	The fund's performance rank based on its return over the twelve months before the fiscal (half-) year end in its investment objective. Investment objectives are based on the CRSP/Lipper codes. This variable is normalized between zero and one.	CRSP, EST
1-year Return Rank^2	The fund's squared 1-year return rank	CRSP, EST
1-year Bottom Quintile	Computed as min(1-year Return Rank; 0.2).	CRSP, EST
1-year Mid Quintiles	Computed as min(1-year Return Rank - 1-year Bottom Quintile; 0.6).	CRSP, EST

Panel C: Other control variables (cont'd)

Variable name	Description	Source
1-year Top Quintile	Computed as 1-year Return Rank - (1-year Bottom Quintile + 1-year Mid Quintiles).	CRSP, EST
1-year 1F-α (4F-α) Rank	The fund's performance rank based on its 1-factor (4-factor) alpha estimated using daily returns over the twelve months before the fiscal (half-) year end in its investment objective. Investment objectives are based on the CRSP/Lipper codes. This variable is normalized between zero and one.	CRSP, KF, EST
1-year 1F-α (4F-α) Rank^2	The fund's squared 1-year 1-factor (4-factor) rank.	CRSP, KF, EST
Return$_{report,filing}$	The fund's return between the reporting and filing date.	CRSP, EST
6-month (1-year) Flow	Funds' net flow over the six (twelve) months before the fiscal (half-) year end.	CRSP, EST
6-month (1-year) Family Flow	Average net flow of all funds in the fund family over the six (twelve) months before the fiscal (half-)year end.	CRSP, EST
Fund Size	Logarithm of a fund's total net assets.	CRSP
Family Size	Logarithm of the fund family's total net assets.	CRSP, EST
Expense Ratio	A fund's annual expense ratio in percent. The variable is winsorized at the 1st and 99th percentile.	CRSP
Fund Age	Logarithm of a fund's age computed from the date a fund was first offered (CRSP variable "first_offer_dt").	CRSP, EST
Flow Segment	Aggregate percentage fund flow to a segment. Segments are based on the CRSP/Lipper investment objective code.	CRSP, EST
Investment Objective	The fund's investment objective based on CRSP/Lipper objective codes (variable "$crsp_obj_cd$").	CRSP
ln(Words)	Logarithm of a shareholder letter's total number of words.	SEC, EST
Time Difference	Number of days between the reporting period end date (report date) and the filing date of the shareholder letter.	SEC, EST
Textual Similarity	Cosine similarity between the fund's current and previous year's shareholder letter. The construction follows Cohen, Malloy, and Nguyen (2020).	SEC,EST
X Star Rating	Dummy variable equal to one if the fund's Morningstar rating is X stars and zero otherwise.	MS
N-CSR	Dummy variable equal to one if the fund's shareholder letter is part of the annual Form N-CSR filing and zero if the letter is part of the semi-annual Form N-CSRS filing.	SEC, EST
Young Fund	Dummy variable equal to one if fund age is below the median.	CRSP, EST
No-Load Fund	Dummy variable equal to one if the fund offers at least one share class with no front-end load and zero otherwise.	CRSP, EST
Retail Fund	Dummy variable equal to one if the fund offers only retail share classes and zero otherwise.	CRSP, EST

4 Examples of negative forward-looking and non-forward-looking sentences

This section shows five examples of negative forward-looking and non-forward-looking sentences. Negative words according to the (Loughran and McDonald 2011) dictionary of negative words are printed in bold font. Forward-looking words according to (Li 2010) are italicized.

Forward-looking sentences

1. **Defaults** are expected to remain below their historical average, but we are of the opinion that the trend of **deteriorating** new-issue quality *may* continue, planting the seeds of future **failures**. Filing date: March 9, 2006. Fund: Credit Suisse Asset Management Income Fund Inc.

2. Not to scare you, but with this **excessive** pessimism, I *expect* that there *will* be many short-term **disappointments** and **challenges**. Filing Date: December 12, 2007. Fund: Giordano Fund.

3. **Declining** stock prices in the first few weeks of the year seem to reflect a growing consensus that the consumer *will* **break** and the economy *will* slip into **recession** this year. Filing Date: February 25, 2008. Fund: T. Rowe Price New America Growth Portfolio.

4. However, a number of **challenges lie** ahead, and we **caution** our investors that the recovery *could* be **slow** and uneven. Filing Date: February 26, 2010. Fund: International Growth and Income Fund, Inc.

5. We remain **concerned**, however, that an extended period of **negative** real (inflation-adjusted) interest rates *could* have significant inflationary and other **unintended** consequences in the years ahead. Filing Date: February 24, 2011. Fund: T. Rowe Price Diversified Mid-Cap Growth Fund, Inc.

Non-forward-looking sentences

1. According to the U.S. Department of Commerce, third quarter 2008 GDP **declined** 0.5% and its advance estimate for fourth quarter GDP **decline** was 3.8%, the latter being the **worst** quarterly reading since 1982. Filing date: March 5, 2009. Fund: Western Asset High Income Fund Inc.

2. Regardless of how one defines a **recession**, it certainly has felt like we are in the midst of an economic **contraction**. Filing date: February 5, 2009. Fund: Legg Mason Partners Large Cap Growth Fund.

3. Since December 2007, approximately six million jobs have been shed and we have experienced seventeen consecutive months of job **losses**, matching the record that occurred during the 1981-82 **recession**. Filing date: August 4, 2009. Fund: Legg Mason Partners Mid Cap Core Fund.

4. It was **difficult** to find any positive financial news in 2008 as the U.S. economy led the global economy into the most **severe recession** since the 1930s. Filing Date: March 9, 2009. Fund: James Advantage Funds.

5. The historic **downgrade** followed a political stalemate in which Congress struggled to address the debt ceiling issue before an early-August deadline, resulting in heightened investor anxiety and **volatility** across major financial markets. Filing Date: October 27, 2011. Fund: Fidelity Four-in-One Index Fund.

5 Determinants of shareholder letters' writing style

In this section, we discuss the results of the formal analysis of the determinants of letter tone in Table IA-6 that we shortly refer to in Section 2.3 in the paper in more detail. We use our tone measures, LMD^- and forward-looking and non-forward-looking LMD^- as dependent variables and relate them to past performance as well as several fund and fund company characteristics.

To take into account that some fund managers might write in a generally more positive and optimistic tone, we include fund fixed effects in all regressions. Thus, any effect we find is purely driven by within-fund variation of letter writing style.

In addition, we also include time fixed effects (columns (1), (3), and (5)) to control for the overall performance of equity markets and general time trends in writing style. Specifically, because the time lag between the "report month" and the "filing month" differs between funds, we include both, report month and filing month fixed effects.

Report month fixed effects capture the impact that average performance and market conditions might have on the writing style of all funds whose reporting period ends in the same month as the fund under consideration. Filing month fixed effects capture any potential impact of general economic conditions at the moment the report is actually filed. In columns (2), (4), and (6), we replace the filing month fixed effects by combined investment objective x filing month fixed effects to capture any segment-specific market conditions and trends.

The distinction between report and filing date is important as, for example, a very negative market return after the reporting period end but before the filing date, i.e., in the period during which the letter is actually written, might have an impact on the tone of the letter. Thus, we include individual fund returns between the report and filing date ($\text{Return}_{report,filing}$. Standard errors are double-clustered at the fund and time dimension.

In all specifications in Panel A of Table IA-6, we find a significant and negative relation between the fund's performance in the six and twelve months before the report date and letter tone. Thus, funds with high (low) performance use fewer (more) negative words in their letters.

While the relation between six-month returns and tone is not different for forward- and non-forward-looking sentences, the 1-year returns are much more strongly related to the negativity in sentences about the present and past than to the negativity in sentences about the future. This result suggests that managers adjust their tone to their fund's past performance. Also, the adjusted R^2 for non-forward-looking (columns (5) and (6)) is about 70% higher (46.9% vs. 26.9% and 48.1% vs. 29.0%) than the one for forward-looking tone (columns (3) and (4)) suggesting that our approach of identifying future-oriented sentences works well.

The return between the report and filing date is not significantly related to the overall negativity of shareholder letters indicating that managers focus on the reporting period and do rarely comment on the most recent events. Still, throughout our analyses of fund flows we always control for the performance between the report and filing date.

While there is no significant relation between 1-year fund flows and the tone of shareholder letters, 6-month flows show a significant relation across all specifications. Higher fund flows are associated with a more positive tone.

All other fund characteristics show no robust and significant relation to letter tone.

To summarize, the main impact on negativity comes from the fund's return before the reporting date and from flows over the previous six months. Also, some specifications indicate that there is a weak relation between fund family flows and tone. We account for these relations in our empirical analyses by carefully controlling for these variables in our regressions.

In Panel B of Table IA-6, we conduct a separate analysis for the annual N-CSR and the semi-annual N-CSRS filings. For instance, one might expect that the 6-month performance matters for funds' semi-annual N-CSRS filings, while their 12-month performance affects the negativity in the annual N-CSR filings. However, we do not find support for this idea, as for both filing types negativity is strongly related to 1-year returns. Also, the strong relation between 6-month flows and tone from Panel A becomes much weaker and is even insignificant for the semi-annual filings.Taken together, these analyses show that it is important to control for the 12-month fund return, which we do in all our analyses.

In Panel C of Table IA-6, we also test whether non-linear fund performance measures and fund alphas explain shareholder letter negativity. We find evidence for such a non-linear

relation with the 1-year squared return rank being highly significantly related to negativity. This result indicates that funds that are among the best performers in their investment objective use more positive language in their letters. In our analysis of fund flows, we control for non-linearities in the flow-performance relation using two specifications: (1) return rank and squared return rank and (2) piece-wise linear regressions as in (Sirri and Tufano 1998).

Adding alphas as additional performance controls (columns (3) to (6)) shows that managers adjust their tone to both raw returns and risk-adjusted performance.

Taken together, these findings illustrate that funds do not simply always write extremely positive letters irrespective of how the fund actually performed. Otherwise, fund investors might perceive fund managers to be less trustworthy and Guiso, Sapienza, and Zingales (2008) show that trust generally is an important determinant of investment decisions.

6 Alternative tone measures and details on the mutual fund dictionary of negative words

6.1 Mutual fund dictionary

The Loughran and McDonald (2011) dictionary is designed to specifically capture negative tone in financial text. However, it is constructed based on Form 10-K filings. Therefore, we also construct our own dictionary of negative mutual fund words (MF^-) based on the shareholder letters in our sample as these letters might have different writing styles than Form 10-K filings. Specifically, we randomly draw 200 letters from the sample of shareholder letters and have two individuals read them independent of each other. While reading the letters, they mark all words that sound negative to them. If both readers have marked a word as negative, we include it in our dictionary. This results in a dictionary of 637 words out of which 329 are also included in the Loughran and McDonald (2011) dictionary.[2] Table IA-19 shows the top 50 negative words of our mutual fund dictionary. While many of the top 50 MF^- words overlap with the Loughran and McDonald (2011) dictionary, words like "inflation", "uncertainty", "fluctuate", and "detractors' are less common in Form 10-K filings but are used frequently to describe unfavorable market conditions and underperformance of mutual funds.

6.2 Alternative tone measures

In Table IA-15, we analyze alternative measures for the tone of shareholder letters. All regressions are based on column (1) of Panel A in Table 2 of our paper.

In column (1), we orthogonalize our standard LMD^- tone measure using the regression in column (1) of Table 6. We take the residuals from this regression as our characteristic-adjusted negativity measure. We still find a highly significant negative relation between tone and fund flows.

[2]The mutual fund dictionary of negative words in shareholder letters, MF^-, is available from the authors upon request.

In column (2), we use the change in negativity from the current letter to the previous year's letter. As we analyze changes in tone, we remove the fund fixed effects but include the lagged dependent variable, i.e., the flow in the month of the previous filing. Also, for tone changes we obtain a significantly negative relation between negativity and flows.

As the Loughran and McDonald (2011) dictionary is constructed based on Form 10-K filings, not all of its words may have a negative meaning in the context of mutual funds. For example, "volatility", the most frequent negative word from the dictionary in our sample, is likely to regularly have a neutral meaning in shareholder letters, as volatility is not necessarily harmful for fund performance. Similarly, a manager discussing "unemployment" (among the top 25 most frequent negative words) might talk about positive economic news like "low unemployment" or a "decline in unemployment". To reduce the effect of such high-frequency (and potentially ambiguous) words on our negativity measure, we follow Loughran and McDonald (2011) and also use a term and inverse document frequency (tf.idf) based negativity measure in column (3) of Table IA-15. With a t-statistic of -3.19 the $LMD^{-}_{tf.idf}$ measure is also highly significant alleviating concerns that some of the top negative words drive our results.

In column (4), we include both the percentage of positive and negative words according to the Loughran and McDonald (2011) dictionary. While we find a strong and significant effect for negativity, positive is not significant. This result is inline with Loughran and McDonald (2016), who caution against using positive words to measure sentiment due to the ambiguous meaning of positive words.

We do not find any significant relation between negativity according to the Harvard IV psycho-social dictionary and fund flows (column (5)). This result consistent with Loughran and McDonald (2011), who show that about 75% of the words in the Harvard dictionary are misclassified in a business context. Applying the tf.idf-weight slightly increases the absolute size of the coefficient on Harvard negativity but it still remains insignificant (column (6)).

In columns (7) and (8) we use our mutual fund dictionary. With a t-statistic of -4.31 MF^{-} is highly significant. A one standard deviation increase in MF^{-} is associated with a 15.6 basis point decrease in mutual fund flows which is slightly larger than the effect for LMD^{-} (13.4 basis points (see Column (1) of Panel A in Table 2 of the paper). $\mathrm{MF}^{-}_{tf.idf}$ is

also highly significant with a t-statistic of -3.77 showing that the relation between MF^- and flows is not driven by a few high-frequency words.

Taken together, the results in Table IA-15 show that our results hold for different approaches of quantifying shareholder letter tone. Like Loughran and McDonald (2011) we conclude that it is essential to use a domain-specific dictionary that is able to capture business/mutual fund language.

Besides negativity and positivity, we also analyze the whether fund investors respond to words expressing uncertainty or vague language. In Table IA-16, we regress monthly fund flows on negativity and the percentage of modal weak and uncertainty words according to the Loughran and McDonald (2011) dictionaries.

While negativity still has a significant negative effect on the flows in the filing month, neither uncertain nor vague language is related to fund flows (see columns (1) and (3) of Table IA-16). In columns (2) and (4), we test whether there is an interaction effect between negativity and uncertainty or vague language. The interaction terms are, however, not statistically significant.

In the next step, we analyze whether it matters if letters focus more on general market and economic conditions or on fund-specific issues. We classify a sentence as market/economy related if it contains at least one of the key words "market' , "markets", "economy" , "economies", and "economic." Sentences with no key words are considered to contain fund-specific information. We compute the percentage of negative words separately for market/economy-related sentences and for non-market/economy-related sentences.

In Table IA-20 we regress the fund flow in the filing month on the two negativity measures. We also include for the percentage of sentences that discuss the general market and economic conditions to analyze whether investors prefer fund-specific or more general economic information. All controls and fixed effects from Column (1) in Panel A of Table 2 in the paper are included.

When using CRSP net fund flows (column (1)), we obtain a significantly negative relation between the negativity in non-market, i.e., fund-specific sentences and fund flows but no significant relation between negativity in market- and economy-related sentences and flows.

When analyzing the N-SAR flows in columns (2) to (4), both the negativity in market/economy related sentences and the negativity in fund-specific sentences have a significant effect on net flows and outflows.

From this analysis, we conclude that investors respond to both negative sentiment about the overall economy and negative information about the fund.

References

Baer, M., A. Kempf, and S. Ruenzi, 2011, "Is a Team Different from the Sum of its Parts? Teams Management in the Mutual Fund Industry," *Review of Finance*, 15, 359–396.

Carhart, M. M., 1997, "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57–82.

Cohen, L., C. Malloy, and Q. Nguyen, 2020, "Lazy prices," *The Journal of Finance*, 75(3), 1371–1415.

Guiso, L., P. Sapienza, and L. Zingales, 2008, "Trusting the Stock Market," *Journal of Finance*, 63(6), 2557–2600.

Li, F., 2010, "The information content of forward-looking statements in corporate filings—A naïve Bayesian machine learning approach," *Journal of Accounting Research*, 48(5), 1049–1102.

Lou, D., 2012, "A Flow-Based Explanation for Return Predictability," *Review of Financial Studies*, 25, 3457–3489.

Loughran, T., and B. McDonald, 2011, "When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *Journal of Finance*, 66, 35–65.

——— , 2014, "Measuring Readability in Financial Disclosures," *Journal of Finance*, 69, 1643–1671.

——— , 2016, "Textual analysis in accounting and finance: A survey," *Journal of Accounting Research*, 54(4), 1187–1230.

Pástor, L., R. F. Stambaugh, and L. A. Taylor, 2015, "Scale and skill in active management," *Journal of Financial Economics*, 116(1), 23–45.

Sirri, E. R., and P. Tufano, 1998, "Costly Search and Mutual Fund Flows," *Journal of Finance*, 53, 1589–1622.

Spiegel, M., and H. Zhang, 2013, "Mutual fund risk and market share-adjusted fund flows," *Journal of Financial Economics*, 108(2), 506–528.

Table IA-1: Availability of shareholder letters and monthly fund flows

This table shows regressions of monthly fund flows on a shareholder letter dummy and various fund characteristics. The dependent variables are the net fund flow (columns (1) to (3)) and the absolute fund flow (columns (4) to (6)) in the filing month. We replace flows of the filing month by flows of the subsequent month whenever the filing date is after the 15th calendar day. "Letter Dummy" is equal to one if we can identify a shareholder letter in the fund's N-CSR or N-CSRS filing, and zero otherwise. All specifications include the control variables (except for the number of words) from column (1) of Panel A in Table 2 of our paper. In columns (1) and (4), we control for time fixed effects. In columns (2), (3), (5), and (6) we include reporting month and combined investment objective x filing month fixed effects. In columns (3) and (6), we add fund fixed effects. In Panel B (C), we rerun the regressions including only the annual N-CSR (semi-annual N-CSRS) filings. All control variables are defined in detail in Section 3. Standard errors are double-clustered by fund and time (filing month). t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table IA-1: Availability of shareholder letters and monthly fund flows (cont'd)

Panel A: annual N-CSR and semi-annual N-CSRS filings						
Dependent Variable	Flow Filing Month			Absolute Flow Filing Month		
	(1)	(2)	(3)	(4)	(5)	(6)
Letter Dummy	-0.001	-0.001	-0.001	-0.000	-0.000	0.000
	(-0.67)	(-0.71)	(-0.73)	(-0.58)	(-0.60)	(0.38)
6-month Return	0.050***	0.055***	0.047***	0.011***	0.013**	0.010**
	(8.25)	(7.66)	(6.85)	(2.75)	(2.48)	(2.19)
1-year Return	0.020***	0.028***	0.039***	-0.008***	-0.006*	0.003
	(5.61)	(6.81)	(9.23)	(-2.78)	(-1.72)	(0.83)
1-year Flow	0.027***	0.027***	0.020***	0.015***	0.015***	0.011***
	(33.08)	(32.97)	(21.93)	(19.43)	(19.68)	(13.92)
$\text{Return}_{Report,Filing}$	0.131***	0.144***	0.136***	0.008	0.011	0.013**
	(12.66)	(12.99)	(13.41)	(1.20)	(1.51)	(2.15)
1-year Return Family	0.007*	0.008*	0.004	-0.000	0.001	0.007*
	(1.76)	(1.91)	(0.85)	(-0.05)	(0.29)	(1.96)
1-year Flow Family	0.011***	0.010***	0.008***	-0.001	-0.002	-0.004***
	(5.58)	(5.68)	(4.16)	(-0.69)	(-1.36)	(-2.82)
ln(Fund size)	-0.002***	-0.002***	-0.009***	-0.004***	-0.003***	-0.007***
	(-8.76)	(-8.74)	(-16.48)	(-12.15)	(-15.50)	(-15.92)
ln(Family size)	0.001***	0.000***	0.001*	0.001***	0.001***	-0.001
	(4.43)	(3.67)	(1.95)	(6.89)	(6.27)	(-1.11)
ln(fund age)	-0.002***	-0.002***	-0.006***	-0.002***	-0.002***	-0.001
	(-4.54)	(-4.58)	(-4.57)	(-4.22)	(-5.72)	(-0.51)
Expense ratio	-0.279***	-0.330***	-0.794***	0.284***	0.041	-0.781***
	(-3.54)	(-4.31)	(-3.69)	(3.67)	(0.52)	(-4.55)
Fund Risk	-0.034	-0.068**	-0.089***	0.093***	0.088***	0.025
	(-1.44)	(-2.40)	(-3.07)	(3.91)	(3.52)	(1.10)
Time Difference	-0.000	-0.000	-0.000	0.000**	0.000	-0.000
	(-0.10)	(-0.30)	(-0.72)	(2.15)	(1.14)	(-0.59)
Segment Flow	0.428***	0.380***	0.436***			
	(7.31)	(2.69)	(3.11)			
Absolute Segment Flow				0.585***	0.616***	0.417***
				(8.56)	(5.13)	(3.43)
Fund FE	N	N	Y	N	N	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y
Filing Month FE	Y	N	N	Y	N	N
Combined Filing Month x Inv. Obj. FE	N	Y	Y	N	Y	Y
Adj. R^2	0.113	0.123	0.159	0.077	0.100	0.193
Observations	83,027	83,034	82,837	83,027	83,034	82,837

Table IA-1: Availability of shareholder letters and monthly fund flows (cont'd)

Panel B: annual N-CSR filings						
Dependent Variable	Flow Filing Month			Absolute Flow Filing Month		
	(1)	(2)	(3)	(4)	(5)	(6)
Letter Dummy	-0.004	-0.003	-0.003	0.000	-0.000	0.000
	(-1.53)	(-1.58)	(-0.91)	(0.08)	(-0.23)	(0.02)
6-month Return	0.047***	0.051***	0.045***	0.016***	0.015**	0.008
	(6.29)	(5.98)	(4.91)	(3.03)	(2.40)	(1.43)
1-year Return	0.023***	0.033***	0.044***	-0.009**	-0.006	0.005
	(4.32)	(5.69)	(6.97)	(-2.28)	(-1.25)	(1.27)
1-year Flow	0.028***	0.028***	0.021***	0.016***	0.015***	0.011***
	(24.15)	(23.82)	(15.98)	(15.29)	(14.99)	(11.00)
$\text{Return}_{Report,Filing}$	0.122***	0.134***	0.127***	0.005	0.011	0.016*
	(8.31)	(8.56)	(9.00)	(0.55)	(0.92)	(1.71)
1-year Return Family	0.014***	0.012**	0.004	0.003	0.002	0.003
	(2.78)	(2.32)	(0.80)	(0.42)	(0.42)	(0.60)
1-year Flow Family	0.009***	0.009***	0.006**	-0.002	-0.004	-0.004*
	(3.54)	(3.51)	(2.33)	(-0.79)	(-1.57)	(-1.90)
ln(Fund size)	-0.002***	-0.002***	-0.009***	-0.003***	-0.003***	-0.007***
	(-7.73)	(-7.63)	(-12.99)	(-10.33)	(-13.28)	(-13.19)
ln(Family size)	0.001***	0.001***	0.001	0.001***	0.001***	-0.001
	(3.56)	(3.16)	(1.28)	(5.02)	(4.68)	(-1.13)
ln(fund age)	-0.002***	-0.002***	-0.005***	-0.003***	-0.003***	-0.002
	(-4.83)	(-4.76)	(-3.21)	(-5.13)	(-7.14)	(-1.41)
Expense ratio	-0.334***	-0.352***	-0.315	0.347***	0.102	-0.511***
	(-3.14)	(-3.38)	(-1.13)	(3.26)	(1.17)	(-2.70)
Fund Risk	-0.023	-0.069*	-0.067*	0.087***	0.075**	0.025
	(-0.83)	(-1.92)	(-1.81)	(3.39)	(2.36)	(0.78)
Time Difference	0.000	0.000	0.000	0.000*	0.000	0.000
	(0.09)	(0.02)	(0.51)	(1.93)	(1.23)	(0.22)
Segment Flow	0.393***	0.335**	0.309**			
	(6.15)	(2.28)	(2.10)			
Absolute Segment Flow				0.476***	0.464**	0.322*
				(5.80)	(2.52)	(1.74)
Fund FE	N	N	Y	N	N	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y
Filing Month FE	Y	N	N	Y	N	N
Combined Filing Month x Inv. Obj. FE	N	Y	Y	N	Y	Y
Adj. R^2	0.117	0.125	0.170	0.077	0.106	0.194
Observations	42,052	42,003	41,612	42,052	42,003	41,612

Table IA-1: Availability of shareholder letters and monthly fund flows (cont'd)

Panel C: semi-annual N-CSRS filings						
Dependent Variable	Flow Filing Month			Absolute Flow Filing Month		
	(1)	(2)	(3)	(4)	(5)	(6)
Letter Dummy	0.000	0.000	-0.002	-0.000	-0.000	0.000
	(0.41)	(0.11)	(-1.56)	(-0.30)	(-0.44)	(0.43)
6-month Return	0.052***	0.059***	0.052***	0.006	0.009	0.011*
	(6.71)	(6.06)	(5.34)	(1.01)	(1.33)	(1.68)
1-year Return	0.018***	0.022***	0.031***	-0.007	-0.006	0.001
	(3.94)	(4.07)	(5.22)	(-1.55)	(-1.05)	(0.15)
1-year Flow	0.027***	0.026***	0.020***	0.015***	0.014***	0.010***
	(27.05)	(26.01)	(16.73)	(14.96)	(15.26)	(10.39)
$\text{Return}_{Report,Filing}$	0.141***	0.155***	0.145***	0.013	0.010	0.011
	(13.27)	(13.19)	(13.31)	(1.45)	(0.98)	(1.41)
1-year Return Family	0.001	0.003	-0.000	-0.003	-0.000	0.008
	(0.15)	(0.55)	(-0.00)	(-0.49)	(-0.05)	(1.49)
1-year Flow Family	0.012***	0.011***	0.007***	-0.001	-0.002	-0.003*
	(4.31)	(4.29)	(3.03)	(-0.48)	(-1.27)	(-1.85)
ln(Fund size)	-0.002***	-0.002***	-0.008***	-0.004***	-0.003***	-0.007***
	(-6.12)	(-6.21)	(-13.00)	(-10.61)	(-13.73)	(-12.18)
ln(Family size)	0.001***	0.000***	0.001**	0.001***	0.001***	-0.000
	(3.83)	(3.11)	(2.03)	(5.53)	(4.81)	(-0.00)
ln(fund age)	-0.001*	-0.001**	-0.004***	-0.002***	-0.002***	0.001
	(-1.83)	(-2.04)	(-2.83)	(-2.61)	(-3.75)	(0.59)
Expense ratio	-0.224**	-0.321***	-1.269***	0.268**	-0.031	-0.917***
	(-2.39)	(-3.43)	(-4.41)	(2.42)	(-0.29)	(-3.64)
Fund Risk	-0.037	-0.068*	-0.104***	0.085***	0.070***	0.030
	(-1.20)	(-1.96)	(-2.64)	(3.05)	(2.64)	(1.09)
Time Difference	-0.000	-0.000	-0.000*	0.000	-0.000	-0.000
	(-0.24)	(-0.38)	(-1.86)	(1.48)	(-0.24)	(-0.20)
Segment Flow	0.438***	0.444*	0.400			
	(4.88)	(1.72)	(1.46)			
Absolute Segment Flow				0.689***	0.611***	0.361**
				(6.88)	(4.50)	(2.39)
Fund FE	N	N	Y	N	N	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y
Filing Month FE	Y	N	N	Y	N	N
Combined Filing Month x Inv. Obj. FE	N	Y	Y	N	Y	Y
Adj. R^2	0.112	0.123	0.157	0.087	0.121	0.210
Observations	40,975	40,922	40,497	40,975	40,922	40,497

Table IA-2: Determinants of shareholder letter availability

This table explores the determinants of the availability of shareholder letters in the N-CSR/N-CSRS filings. In Panel A, we regress a shareholder letter dummy on various fund and filing characteristics. The shareholder letter dummy is equal to one if there is at least one letter in the filing and zero otherwise. In Panel B, the dependent variable is a matched shareholder letter dummy that is equal to one if we can uniquely match a fund with a shareholder letter and zero otherwise. The sample in Panel B includes all funds from filings that contain at least one shareholder letter. All specifications include fund performance over various horizons, fund flows, fund family performance and flows, fund size and age, family size, the fund's expense ratio, fund risk, the time between the report and filing date, and a dummy for the annual N-CSR filings. In columns (1) and (3) in Panel A, we additionally include a dummy indicating whether the previous year's filing includes a shareholder letter. In columns (1), (3), and (5) in Panel B, we additionally include a dummy indicating whether the fund is uniquely matched to a letter in the previous year's filing. Reporting month fixed effects and combined investment objective x filing month fixed effects are always included in specifications (1) to (4). In columns (3) and (4), we add fund fixed effects. Columns (5) and (6) of Panel B additionally include filing fixed effects. In these specification, we remove all fund family and filing controls as well as the reporting month fixed effects, as there is no within-filing variation. All control variables are defined in detail in Section 3. The unit of observation is the fund-filing level. Standard errors are double-clustered by fund family (CIK) and time (filing month). t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table IA-2: Determinants of shareholder letter availability (cont'd)

Panel A: Determinants of Shareholder Letter Availability				
Dependent Variable		Filing with Letter		
	(1)	(2)	(3)	(4)
$\text{Return}_{Report,Filing}$	-0.014	-0.030	-0.031	-0.034
	(-0.33)	(-0.83)	(-0.80)	(-0.93)
6-month Return	-0.017	0.001	-0.015	-0.005
	(-0.61)	(0.03)	(-0.58)	(-0.23)
1-year Return	0.025	0.020	0.013	0.014
	(1.28)	(1.34)	(0.80)	(0.96)
1-year Flow	0.014***	0.006**	0.006	0.006**
	(3.45)	(2.57)	(1.51)	(2.08)
1-year Return Family	-0.006	-0.025	-0.052	-0.027
	(-0.12)	(-0.69)	(-1.37)	(-0.78)
1-year Flow Family	-0.021	0.005	0.014	0.010
	(-0.94)	(0.38)	(1.05)	(0.75)
ln(Fund size)	-0.002	-0.002	-0.003	-0.002
	(-0.60)	(-1.36)	(-1.01)	(-1.23)
ln(Family size)	0.010***	0.002*	0.001	-0.000
	(3.22)	(1.90)	(0.13)	(-0.15)
ln(fund age)	0.011	0.005*	0.007	0.004
	(1.57)	(1.75)	(0.63)	(0.58)
Expense ratio	3.874***	0.689	2.489	0.558
	(3.05)	(1.30)	(1.57)	(0.51)
Fund Risk	0.186	0.032	0.267	0.102
	(1.06)	(0.34)	(1.56)	(0.76)
Time Difference	-0.001	-0.000	0.000	-0.000
	(-0.62)	(-0.50)	(0.19)	(-0.33)
ln(Number Funds in Filing)	-0.004	-0.002	0.049***	0.026**
	(-0.25)	(-0.31)	(2.78)	(2.24)
N-CSR Dummy	0.115***	0.041***	0.118***	0.060***
	(8.93)	(6.02)	(9.36)	(8.27)
Letter Previous Year		0.649***		0.495***
		(31.47)		(20.68)
Fund FE	N	N	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.118	0.494	0.362	0.527
Observations	83,584	74,135	83,385	73,910

Table IA-2: Determinants of shareholder letter availability (cont'd)

Panel B: Determinants of Shareholder Letter Matching						
Dependent Variable	Matched Shareholder Letter					
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Return}_{Report,Filing}$	0.056	0.044	-0.022	0.024	-0.000	0.006
	(0.77)	(1.08)	(-0.40)	(0.59)	(-0.01)	(0.15)
6-month Return	0.038	-0.008	0.048	0.000	-0.000	0.010
	(1.14)	(-0.33)	(1.37)	(0.00)	(-0.00)	(0.36)
1-year Return	0.034	0.023	-0.020	0.009	-0.004	-0.007
	(1.01)	(1.22)	(-0.76)	(0.42)	(-0.22)	(-0.34)
1-year Flow	-0.010	-0.003	-0.009	-0.006	-0.002	0.000
	(-1.09)	(-0.66)	(-1.52)	(-1.19)	(-0.51)	(0.06)
1-year Return Family	-0.057	-0.024	-0.042	-0.024		
	(-0.57)	(-0.75)	(-0.84)	(-0.61)		
1-year Flow Family	-0.009	-0.003	-0.010	0.001		
	(-0.22)	(-0.21)	(-0.50)	(0.04)		
ln(Fund size)	0.014***	0.003**	0.010**	0.004	0.007*	0.004
	(3.43)	(2.30)	(2.30)	(1.31)	(1.82)	(1.24)
ln(Family size)	0.030***	0.006***	-0.000	0.001		
	(5.28)	(4.30)	(-0.06)	(0.24)		
ln(fund age)	0.032***	0.007**	-0.027	-0.016*	0.009	0.001
	(3.12)	(2.28)	(-1.65)	(-1.75)	(0.98)	(0.11)
Expense ratio	7.160***	1.611**	0.939	-0.440	2.637	1.206
	(2.64)	(2.07)	(0.39)	(-0.27)	(0.84)	(0.50)
Fund Risk	0.163	0.058	0.099	0.131	-0.036	-0.083
	(0.34)	(0.40)	(0.37)	(0.67)	(-0.19)	(-0.50)
Time Difference	0.003	0.000	0.003**	-0.000		
	(1.37)	(0.24)	(2.16)	(-0.10)		
ln(Number Funds in Filing)	-0.189***	-0.037***	-0.136***	-0.057***		
	(-9.79)	(-6.56)	(-7.81)	(-4.85)		
N-CSR Dummy	0.177***	0.037***	0.184***	0.084***		
	(9.26)	(7.10)	(9.45)	(9.69)		
Match Letter Previous Year		0.792***		0.554***		0.339***
		(81.46)		(32.48)		(14.03)
Fund FE	N	N	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y	N	N
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y	Y	Y
Filing FE	N	N	N	N	Y	Y
Adj. R^2	0.200	0.702	0.598	0.738	0.826	0.852
Observations	74,775	64,045	74,531	63,772	60,993	52,012

Table IA-3: Publication Dates of N-CSR and N-CSRS filings

This table shows the distribution of reporting months (columns (1) and (2)) and filing months (columns (3) and (4)) during the year. Reporting month is the month of the fiscal (half-) year end. Filing month is the month of the actual SEC filing. Columns (1) and (3) ((2) and (4)) show absolute (relative) frequencies.

	Report Date		Filing Date	
	Number	Percentage	Number	Percentage
	(1)	(2)	(3)	(4)
January	954	2.69%	3,696	10.43%
February	2,076	5.86%	2,316	6.54%
March	2,865	8.08%	4,764	13.44%
April	3,679	10.38%	1,564	4.41%
May	1,387	3.91%	1,925	5.43%
June	4,873	13.75%	3,332	9.4%
July	1,300	3.67%	2,788	7.87%
August	2,073	5.85%	3,312	9.35%
September	3,400	9.59%	3,151	8.89%
October	5,671	16%	1,335	3.77%
November	1,242	3.5%	2,560	7.22%
December	5,917	16.7%	4,694	13.25%
Total	35,437	100%	35,437	100%

Table IA-4: Top 25 negative words

This table shows the 25 most frequent negative words according to the Loughran and McDonald (2011) dictionary. Absolute frequency is the total number of occurrences of the negative word across all shareholder letters in the sample. Inverse document frequency is the log of the number of shareholder letters in the sample divided by the number of shareholder letters containing the negative word.

Rank	Negative word	Abs. frequency	Inverse document frequency
1	volatility	32,236	0.82
2	concerns	22,889	1.00
3	negative	20,614	1.09
4	detracted	19,951	1.19
5	declined	15,939	1.33
6	underperformed	13,756	1.42
7	crisis	13,651	1.56
8	decline	13,471	1.43
9	weak	11,243	1.55
10	hurt	11,020	1.59
11	against	10,163	1.56
12	losses	9,943	1.72
13	recession	9,706	1.75
14	weakness	9,690	1.70
15	poor	9,662	1.67
16	loss	9,180	1.72
17	volatile	9,176	1.67
18	lagged	8,049	1.82
19	slowing	7,976	1.87
20	underperformance	7,789	1.88
21	unemployment	7,718	1.90
22	worst	7,390	1.95
23	defensive	7,203	2.04
24	sharply	7,175	1.91
25	difficult	7,004	1.94

Table IA-5: Correlations of main variables

This table shows correlations of shareholder letter tone measures and fund characteristics. Panel A shows the correlations of the different tone measures. Panel B displays correlations of negative tone and fund characteristics. LMD^- (LMD^+) is the fraction of negative (positive) words according to the Loughran and McDonald (2011) dictionary of negative (positive) words. Fwd. LMD^- and Non-fwd. LMD^- distinguish between the negativity in forward-looking and non-forward-looking sentences. Forward-looking sentences are defined as in Li (2010). Fwd. sentences is the percentage of forward-looking sentences in the letter. Number of words is the number of words of the shareholder letter. Flow adj. filing month is the net fund flow in the filing month (in the month after the filing month) for letters that are filed by (after) the 15th calendar day of the month. 1-year return (1-year flow) is the fund return (fund flow) during the twelve months before the report date. 1-year family return is the return of the fund family during the twelve months before the report date. All fund characteristics are defined in detail in Section 3. p-values are provided in parentheses.

Variables	LMD^-	LMD^+	Fwd. LMD^-	Non-fwd. LMD^-	Number of words	Fwd. sentences
LMD^-	1.000					
LMD^+	0.125	1.000				
	(0.000)					
Fwd. LMD^-	0.506	0.059	1.000			
	(0.000)	(0.000)				
Non-fwd. LMD^-	0.950	0.125	0.306	1.000		
	(0.000)	(0.000)	(0.000)			
Number of words	0.047	-0.016	0.140	0.031	1.000	
	(0.000)	(0.004)	(0.000)	(0.000)		
Fwd. sentences	-0.219	-0.166	0.029	-0.176	0.069	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

Table IA-5: Correlations of main variables (cont'd)

Variables	LMD^-	Flow adj. Filing Month	ln(Fund size)	ln(fund age)	Expense ratio	Fund Risk	1-year Return	1-year Return Family	1-year Flow
LMD^-	1.000								
Flow adj. Filing Month	-0.042	1.000							
	(0.000)								
ln(Fund size)	0.014	-0.047	1.000						
	(0.009)	(0.000)							
ln(Fund age)	0.012	-0.137	0.485	1.000					
	(0.019)	(0.000)	(0.000)						
Expense Ratio	-0.025	-0.030	-0.399	-0.126	1.000				
	(0.000)	(0.000)	(0.000)	(0.000)					
Fund Risk	0.122	-0.016	-0.029	0.080	0.089	1.000			
	(0.000)	(0.003)	(0.000)	(0.000)	(0.000)				
1-year Return	-0.236	0.079	0.099	0.051	-0.056	-0.163	1.000		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
1-year Return Family	-0.222	0.050	0.076	0.033	-0.066	-0.197	0.789	1.000	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
1-year Flow	-0.041	0.302	-0.055	-0.364	-0.039	-0.076	0.105	0.058	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

Table IA-6: Determinants of shareholder letter tone

This table shows regressions of shareholder letter tone on various fund characteristics. In column (1) and (2), tone is measured as the percentage of negative words according to the Loughran and McDonald (2011) dictionary. In columns (3) and (4) ((5) and (6)), we analyze the percentage of negative words in forward-looking (non-forward-looking) sentences. Forward-looking sentences are defined as in Li (2010). Columns (1), (3), and (5) ((2), (4), and (6)) include filing month fixed effects (combined investment objective x filing month fixed effects). Reporting month fixed effects and fund fixed effects are always included. In Panel B, we rerun columns (1) and (2) from Panel A separately for N-CSR (columns (1) and (2)) and N-CSRS (columns (3) and (4)) filings. In Panel C, we repeat the regressions from columns (1) and (2) of Panel A but use alternative measures for past fund performance. In columns (1) and (2), we replace the 6-month and 12-month return by the return rank and squared rank. In columns (3) and (4) ((5) and (6)), we add the CAPM 1-factor (Carhart (1997) 4-factor) alpha. All control variables are defined in detail in Section 3. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table IA-6: Determinants of shareholder letter tone (cont'd)

Panel A: Determinants of Tone and Content						
Dependent Variable	LMD$^-$		Forward-looking LMD$^-$		Non-forward looking LMD$^-$	
	(1)	(2)	(3)	(4)	(5)	(6)
Return$_{Report,Filing}$	-0.246	-0.245	0.114	0.398	-0.294	-0.388*
	(-1.55)	(-1.39)	(0.49)	(1.41)	(-1.61)	(-1.95)
6-month Return	-0.365***	-0.372***	-0.496***	-0.393**	-0.387***	-0.403***
	(-3.10)	(-3.02)	(-2.85)	(-2.17)	(-2.79)	(-2.81)
1-year Return	-0.884***	-0.865***	-0.226*	-0.182	-0.999***	-0.994***
	(-11.01)	(-10.76)	(-1.85)	(-1.54)	(-11.05)	(-10.94)
6-month Flow	-0.126***	-0.131***	-0.181**	-0.188**	-0.141***	-0.140***
	(-2.90)	(-3.08)	(-2.25)	(-2.30)	(-2.81)	(-2.74)
1-year Flow	0.012	0.009	0.056	0.045	0.019	0.015
	(0.55)	(0.40)	(1.37)	(1.11)	(0.76)	(0.58)
6-month Return Family	-0.335	-0.247	-0.320	-0.101	-0.425*	-0.356
	(-1.63)	(-1.19)	(-0.93)	(-0.31)	(-1.79)	(-1.51)
1-year Return Family	0.258	0.204	0.422*	0.264	0.142	0.105
	(1.47)	(1.19)	(1.67)	(1.11)	(0.72)	(0.55)
6-month Flow Family	-0.101	-0.093	0.008	-0.020	-0.124	-0.118
	(-0.73)	(-0.69)	(0.04)	(-0.11)	(-0.77)	(-0.75)
1-year Flow Family	0.134*	0.146**	0.073	0.097	0.152*	0.165**
	(1.89)	(2.19)	(0.78)	(1.02)	(1.90)	(2.19)
Fund Risk	0.600	1.011	1.052	1.093	0.572	1.017
	(0.93)	(1.53)	(1.02)	(0.97)	(0.76)	(1.34)
ln(Fund size)	-0.009	-0.000	0.001	0.003	-0.012	-0.001
	(-0.76)	(-0.03)	(0.04)	(0.20)	(-0.86)	(-0.10)
ln(Family size)	0.002	-0.003	-0.007	-0.010	-0.000	-0.005
	(0.12)	(-0.17)	(-0.32)	(-0.45)	(-0.02)	(-0.27)
ln(fund age)	0.058	0.057	0.031	0.056	0.071*	0.067
	(1.54)	(1.46)	(0.53)	(0.94)	(1.72)	(1.54)
Expense ratio	-9.650	-9.191	-8.829	-12.787	-8.347	-7.660
	(-1.38)	(-1.32)	(-0.94)	(-1.40)	(-1.00)	(-0.92)
Fund FE	Y	Y	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y
Filing Month FE	Y	N	Y	N	Y	N
Combined Filing Month x Inv. Obj. FE	N	Y	N	Y	N	Y
$Adj.R^2$	0.514	0.524	0.269	0.290	0.469	0.481
Observations	35,182	35,136	35,182	35,136	35,182	35,136

Table IA-6: Determinants of shareholder letter tone (cont'd)

Panel B: Separate Analysis for N-CRS and N-CSRS Filings				
Dependent Variable	LMD$^-$			
Filing	N-CSR		N-CSRS	
	(1)	(2)	(3)	(4)
Return$_{Report,Filing}$	-0.147	-0.210	-0.412	-0.429
	(-0.78)	(-0.98)	(-1.56)	(-1.39)
6-month Return	-0.299*	-0.283*	-0.314	-0.394
	(-1.95)	(-1.83)	(-1.32)	(-1.62)
1-year Return	-1.049***	-1.081***	-0.657***	-0.516***
	(-10.24)	(-9.94)	(-4.51)	(-3.95)
6-month Flow	-0.104*	-0.099*	-0.090	-0.116
	(-1.94)	(-1.82)	(-0.93)	(-1.27)
1-year Flow	-0.004	-0.012	0.015	0.005
	(-0.16)	(-0.43)	(0.34)	(0.12)
6-month Return Family	-0.353	-0.179	-0.711**	-0.689*
	(-1.18)	(-0.59)	(-2.09)	(-1.95)
1-year Return Family	0.248	0.143	0.457*	0.421
	(1.05)	(0.63)	(1.89)	(1.64)
6-month Flow Family	-0.133	-0.095	0.072	0.103
	(-0.73)	(-0.52)	(0.33)	(0.48)
1-year Flow Family	0.118	0.109	0.069	0.058
	(1.30)	(1.29)	(0.71)	(0.54)
Fund Risk	0.168	0.241	1.008	1.692
	(0.24)	(0.33)	(1.03)	(1.60)
ln(Fund size)	-0.000	0.012	-0.008	-0.010
	(-0.02)	(0.89)	(-0.42)	(-0.49)
ln(Family size)	0.012	0.006	-0.020	-0.028
	(0.69)	(0.31)	(-0.86)	(-1.16)
ln(fund age)	0.029	0.032	0.101*	0.125**
	(0.73)	(0.77)	(1.90)	(2.30)
Expense ratio	-2.211	-1.276	-12.622	-18.273**
	(-0.31)	(-0.18)	(-1.41)	(-2.01)
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Filing Month FE	Y	N	Y	N
Combined Filing Month x Inv. Obj. FE	N	Y	N	Y
$Adj.R^2$	0.544	0.557	0.538	0.550
Observations	22,258	22,138	12,337	12,126

Table IA-6: Determinants of shareholder letter tone (cont'd)

Panel C: Alternative Performance Controls						
Dependent Variable			LMD$^-$			
	(1)	(2)	(3)	(4)	(5)	(6)
Return$_{Report,Filing}$	-0.210	-0.213	-0.257	-0.251	-0.250	-0.248
	(-1.29)	(-1.21)	(-1.62)	(-1.43)	(-1.58)	(-1.41)
6-month Return			-0.346***	-0.347***	-0.357***	-0.363***
			(-2.90)	(-2.76)	(-3.05)	(-2.96)
1-year Return			-0.679***	-0.611***	-0.763***	-0.737***
			(-4.99)	(-4.50)	(-7.07)	(-6.67)
6-month Return Rank	-0.129*	-0.153**				
	(-1.76)	(-2.17)				
6-month Return Rank^2	0.029	0.054				
	(0.41)	(0.80)				
1-year Return Rank	-0.072	-0.072				
	(-0.99)	(-1.00)				
1-year Return Rank^2	-0.192***	-0.191***				
	(-2.83)	(-2.79)				
1-year 1F-Alpha			-3.546*	-4.374**		
			(-1.83)	(-2.25)		
1-year 4F-Alpha					-2.934*	-3.090*
					(-1.74)	(-1.83)
6-month Flow	-0.115***	-0.108**	-0.122***	-0.128***	-0.121***	-0.128***
	(-2.71)	(-2.57)	(-2.78)	(-2.99)	(-2.77)	(-2.99)
1-year Flow	0.007	0.004	0.013	0.010	0.013	0.010
	(0.32)	(0.17)	(0.60)	(0.44)	(0.59)	(0.43)
6-month Return Family	-0.301	-0.186	-0.329	-0.240	-0.334	-0.246
	(-1.47)	(-0.89)	(-1.60)	(-1.15)	(-1.63)	(-1.18)
1-year Return Family	0.129	0.153	0.260	0.208	0.261	0.208
	(0.74)	(0.89)	(1.47)	(1.21)	(1.48)	(1.21)
6-month Flow Family	-0.072	-0.081	-0.097	-0.088	-0.099	-0.091
	(-0.51)	(-0.60)	(-0.70)	(-0.65)	(-0.72)	(-0.68)
1-year Flow Family	0.128*	0.145**	0.134*	0.146**	0.135*	0.146**
	(1.81)	(2.18)	(1.89)	(2.18)	(1.91)	(2.19)
Fund Risk	0.574	1.081	0.497	0.892	0.514	0.913
	(0.90)	(1.64)	(0.79)	(1.37)	(0.80)	(1.39)
ln(Fund size)	-0.022*	-0.007	-0.010	-0.001	-0.011	-0.001
	(-1.78)	(-0.60)	(-0.84)	(-0.09)	(-0.87)	(-0.12)
ln(Family size)	0.004	-0.002	0.002	-0.003	0.002	-0.003
	(0.21)	(-0.13)	(0.14)	(-0.16)	(0.14)	(-0.17)
ln(fund age)	0.065*	0.060	0.060	0.059	0.060	0.059
	(1.72)	(1.53)	(1.60)	(1.50)	(1.61)	(1.51)
Expense ratio	-10.121	-8.715	-9.660	-9.098	-9.654	-9.164
	(-1.45)	(-1.26)	(-1.38)	(-1.30)	(-1.38)	(-1.31)
Fund FE	Y	Y	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y
Filing Month FE	Y	N	Y	N	Y	N
Combined Filing Month x Inv. Obj. FE	N	Y	N	Y	N	Y
$Adj.R^2$	0.514	0.525	0.514	0.524	0.514	0.524
Observations	35,182	35,136	35,182	35,136	35,182	35,136

Table IA-7: Shareholder letter tone and flows in the unadjusted filing month

This table shows regressions of monthly fund flows on shareholder letter tone and various fund characteristics as in Panel A of Table 2 of our paper. The dependent variable is the net fund flow (in percent) in the month of the SEC filing independent of the filing day of the shareholder letter. Tone is the fraction of negative words in a shareholder letter based on the LMD^- dictionary. We include all control variables and fixed effects from Panel A of Table 2 of our paper. All control variables are defined in detail in Section 3. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	Flow Unadjusted Filing Month			
	(1)	(2)	(3)	(4)
LMD^-	-0.107***	-0.103***	-0.084**	-0.084**
	(-2.79)	(-2.72)	(-2.21)	(-2.22)
Controls Table 2 Panel A	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.198	0.198	0.201	0.201
Observations	33,423	33,423	33,423	33,423

Table IA-8: Shareholder letter tone and monthly fund flows: alternative clustering of standard errors

This table shows regressions of monthly fund flows on shareholder letter tone and various fund characteristics as in Panel A of Table 2 of our main paper but we change the clustering of standard errors. The dependent variable is the net fund flow (in percent) in the month of the SEC filing. We replace the flows of the filing month by the flows of the subsequent month whenever the filing date is after the 15th calendar day. Tone is the fraction of negative words in a shareholder letter based on the Loughran and McDonald (2011) (LMD^{-}) dictionary and is standardized to unit variance. In Panel A, standard errors are double-clustered by fund and reporting month. In Panel B (C), we double-cluster standard errors at the fund family and filing (reporting) month dimension. All control variables and fixed effects from Panel A of Table 2 of our paper are included. All control variables are defined in detail in Section 3. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table IA-8: Shareholder letter tone and monthly fund flows: alternative clustering of standard errors (cont'd)

Panel A: standard errors double-clustered by fund and reporting month				
Dependent Variable	Flow Filing Month			
	(1)	(2)	(3)	(4)
LMD^{-}	-0.134***	-0.131***	-0.111***	-0.111***
	(-3.29)	(-3.22)	(-2.76)	(-2.77)
$\text{Return}_{Report,Filing}$	12.936***	13.004***	12.903***	12.905***
	(11.78)	(11.89)	(11.79)	(11.79)
6-month Return	4.730***	4.470***	3.793***	3.714***
	(5.57)	(5.26)	(4.79)	(4.66)
1-year Return	3.288***	0.843		
	(5.57)	(0.77)		
1-year 1F-Alpha		42.258***		
		(3.14)		
1-year Return Rank			0.875*	
			(1.96)	
1-year Return Rank^2			0.630	
			(1.45)	
1-year Bottom Quintile				1.258*
				(1.81)
1-year Mid Quintiles				1.404***
				(7.94)
1-year Top Quintile				2.601***
				(3.13)
1-year Return Family	0.334	0.251	0.026	-0.006
	(0.55)	(0.41)	(0.04)	(-0.01)
1-year Flow	2.161***	2.139***	2.100***	2.096***
	(15.06)	(14.90)	(14.51)	(14.48)
1-year Flow Family	0.699**	0.681**	0.691**	0.694**
	(2.44)	(2.37)	(2.42)	(2.44)
ln(Fund size)	-0.931***	-0.922***	-0.906***	-0.906***
	(-12.12)	(-12.05)	(-11.90)	(-11.89)
ln(Family size)	0.024	0.021	0.023	0.023
	(0.34)	(0.30)	(0.33)	(0.32)
ln(fund age)	-0.593***	-0.611***	-0.623***	-0.624***
	(-3.35)	(-3.43)	(-3.53)	(-3.53)
Expense ratio	-92.739***	-93.072***	-95.016***	-95.104***
	(-3.06)	(-3.07)	(-3.16)	(-3.17)
Fund Risk	-7.119	-5.952	-6.361	-6.276
	(-1.63)	(-1.32)	(-1.40)	(-1.38)
ln(Words)	0.088	0.088	0.084	0.085
	(1.59)	(1.58)	(1.53)	(1.53)
Time Difference	-0.003	-0.003	-0.003	-0.003
	(-0.20)	(-0.27)	(-0.27)	(-0.26)
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.191	0.192	0.195	0.195
Observations	33,398	33,398	33,398	33,398

Table IA-8: Shareholder letter tone and monthly fund flows: alternative clustering of standard errors (cont'd)

Panel B: standard errors double-clustered by fund family and filing month				
Dependent Variable		Flow Filing Month		
	(1)	(2)	(3)	(4)
LMD$^-$	-0.134***	-0.131***	-0.111***	-0.111***
	(-3.35)	(-3.30)	(-2.89)	(-2.90)
Return$_{Report,Filing}$	12.936***	13.004***	12.903***	12.905***
	(9.44)	(9.50)	(9.55)	(9.56)
6-month Return	4.730***	4.470***	3.793***	3.714***
	(5.40)	(5.00)	(4.56)	(4.47)
1-year Return	3.288***	0.843		
	(4.66)	(0.78)		
1-year 1F-Alpha		42.258***		
		(3.25)		
1-year Return Rank			0.875**	
			(2.03)	
1-year Return Rank^2			0.630	
			(1.52)	
1-year Bottom Quintile				1.258*
				(1.83)
1-year Mid Quintiles				1.404***
				(7.68)
1-year Top Quintile				2.601***
				(3.45)
1-year Return Family	0.334	0.251	0.026	-0.006
	(0.50)	(0.37)	(0.04)	(-0.01)
1-year Flow	2.161***	2.139***	2.100***	2.096***
	(15.58)	(15.22)	(14.77)	(14.84)
1-year Flow Family	0.699**	0.681**	0.691**	0.694**
	(2.54)	(2.47)	(2.51)	(2.52)
ln(Fund size)	-0.931***	-0.922***	-0.906***	-0.906***
	(-11.41)	(-11.29)	(-11.17)	(-11.17)
ln(Family size)	0.024	0.021	0.023	0.023
	(0.38)	(0.34)	(0.37)	(0.36)
ln(fund age)	-0.593***	-0.611***	-0.623***	-0.624***
	(-3.75)	(-3.84)	(-3.97)	(-3.99)
Expense ratio	-92.739***	-93.072***	-95.016***	-95.104***
	(-2.89)	(-2.89)	(-2.96)	(-2.97)
Fund Risk	-7.119	-5.952	-6.361	-6.276
	(-1.57)	(-1.29)	(-1.39)	(-1.36)
ln(Words)	0.088	0.088	0.084	0.085
	(1.29)	(1.29)	(1.23)	(1.23)
Time Difference	-0.003	-0.003	-0.003	-0.003
	(-0.19)	(-0.25)	(-0.25)	(-0.24)
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.191	0.192	0.195	0.195
Observations	33,398	33,398	33,398	33,398

Table IA-8: Shareholder letter tone and monthly fund flows: alternative clustering of standard errors (cont'd)

Panel C: standard errors double-clustered by fund family and reporting month				
Dependent Variable		Flow Filing Month		
	(1)	(2)	(3)	(4)
LMD^{-}	-0.134***	-0.131***	-0.111***	-0.111***
	(-3.06)	(-3.00)	(-2.59)	(-2.60)
Return$_{Report,Filing}$	12.936***	13.004***	12.903***	12.905***
	(9.27)	(9.33)	(9.36)	(9.37)
6-month Return	4.730***	4.470***	3.793***	3.714***
	(5.05)	(4.71)	(4.27)	(4.13)
1-year Return	3.288***	0.843		
	(4.73)	(0.76)		
1-year 1F-Alpha		42.258***		
		(3.09)		
1-year Return Rank			0.875*	
			(1.87)	
1-year Return Rank^2			0.630	
			(1.40)	
1-year Bottom Quintile				1.258*
				(1.79)
1-year Mid Quintiles				1.404***
				(7.45)
1-year Top Quintile				2.601***
				(3.11)
1-year Return Family	0.334	0.251	0.026	-0.006
	(0.50)	(0.38)	(0.04)	(-0.01)
1-year Flow	2.161***	2.139***	2.100***	2.096***
	(14.78)	(14.45)	(14.01)	(14.06)
1-year Flow Family	0.699**	0.681**	0.691**	0.694**
	(2.28)	(2.21)	(2.25)	(2.27)
ln(Fund size)	-0.931***	-0.922***	-0.906***	-0.906***
	(-10.79)	(-10.69)	(-10.59)	(-10.59)
ln(Family size)	0.024	0.021	0.023	0.023
	(0.39)	(0.36)	(0.38)	(0.38)
ln(fund age)	-0.593***	-0.611***	-0.623***	-0.624***
	(-3.51)	(-3.59)	(-3.70)	(-3.71)
Expense ratio	-92.739***	-93.072***	-95.016***	-95.104***
	(-2.79)	(-2.80)	(-2.87)	(-2.88)
Fund Risk	-7.119	-5.952	-6.361	-6.276
	(-1.55)	(-1.27)	(-1.36)	(-1.33)
ln(Words)	0.088	0.088	0.084	0.085
	(1.35)	(1.35)	(1.28)	(1.28)
Time Difference	-0.003	-0.003	-0.003	-0.003
	(-0.20)	(-0.26)	(-0.26)	(-0.25)
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.191	0.192	0.195	0.195
Observations	33,398	33,398	33,398	33,398

Table IA-9: Shareholder letter tone and monthly fund flows: Excluding letters with reporting periods covering the financial and Covid-19 crisis

This table shows regressions of monthly fund flows on shareholder letter tone and various fund characteristics as in specifications (3) and (4) of Panel A of Table 3 of our paper. Here we exclude letters not based on their filing date but based on their reporting period. More specifically, we exclude all letters where the reporting period contains one of the crises months (September 2008 until March 2009 (column (1)) and February 2020 until April 2020 (column (2)), respectively). The dependent variable is the net fund flow (in percent) in the month of the SEC filing. We replace the flows of the filing month by the flows of the subsequent month whenever the filing date is after the 15th calendar day. Tone is the fraction of negative words in a shareholder letter based on the LMD^{-} dictionary. We include all control variables and fixed effects from column (1) of Panel A of Table 2 of our paper. All control variables are defined in detail in Section 3. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	Flow Filing Month	
Specification	excluding Fin. Crisis	excluding Covid
	(1)	(2)
LMD^{-}	-0.127***	-0.138***
	(-3.26)	(-3.74)
Controls Table 2 Panel A Column (1)	Y	Y
Fund FE	Y	Y
Reporting Month FE	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y
Adj. R^2	0.196	0.194
Observations	31,354	32,086

Table IA-10: Letter tone and monthly fund flows: Interaction effects using NSAR/NPORT flows

This table repeats the regressions of monthly fund flows on shareholder letter tone, various fund characteristics and interactions between tone and fund characteristics from Panel B of Table 2 in our paper but uses flows from NSAR and NPORT filings instead of the CRSP database. In Panel A, the dependent variable is the net fund flow (in percent) in the month of the SEC filing. Panel B (C) repeats the analysis for inflows (outflows). We replace flows of the filing month by flows of the subsequent month whenever the filing date is after the 15th calendar day. We analyze the interaction effects between the negativity of the shareholder letter (LMD^{-}) and 6-month returns (column (1)), 12-month returns (column (2)), a dummy for negative 6-month returns (column (3)), a dummy for negative 12-month returns (column (4)), a dummy that is one if the shareholder letter is part of the annual N-CSR filing (but not part of the semiannual N-CSRS filing; column (5)), a dummy for no-load funds (column (6)), retail funds (column (7)), and young funds (i.e., below median fund age; column (8)). Control variables and fixed effects from column (1) of Panel A in Table 2 of our paper are always included in the regressions and defined in detail in Section 3. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table IA-10: Letter tone and monthly fund flows: Interaction effects using NSAR/NPORT flows (cont'd)

Panel A: NSAR/NPORT Net Flows Dependent Variable	Flow Filing Month							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LMD$^-$	-0.107***	-0.109***	-0.118***	-0.104***	-0.106**	-0.091**	-0.151***	-0.049
	(-3.82)	(-3.89)	(-3.44)	(-3.16)	(-2.59)	(-2.34)	(-3.20)	(-1.52)
LMD$^-$ x 6-month Return	-0.126							
	(-0.89)							
LMD$^-$ x 1-year Return		-0.031						
		(-0.26)						
LMD$^-$ x Negative Return Dummy$_{6month}$			0.023					
			(0.54)					
LMD$^-$ x Negative Return Dummy$_{12month}$				-0.013				
				(-0.33)				
LMD$^-$ x N-CSR					-0.007			
					(-0.16)			
LMD$^-$ x No-Load Fund						0.042		
						(0.62)		
LMD$^-$ x Retail Fund							0.057	
							(1.10)	
LMD$^-$ x Young fund								-0.157***
								(-2.83)
Negative Return Dummy$_{6month}$			-0.422***					
			(-3.41)					
Negative Return Dummy$_{12month}$				-0.203				
				(-1.42)				
N-CSR					0.004			
					(0.05)			
No-Load Fund						-0.190		
						(-0.52)		
Retail Fund							0.058	
							(0.37)	
Young Fund								0.159
								(1.03)
Controls Table 2 Panel A Column (1)	Y	Y	Y	Y	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.267	0.267	0.268	0.267	0.267	0.246	0.266	0.267
Observations	22,881	22,881	22,881	22,881	22,881	14,046	22,751	22,881

Table IA-10: Letter tone and monthly fund flows: Interaction effects using NSAR/NPORT flows (cont'd)

Panel B: NSAR/NPORT Inflows Dependent Variable	Flow Filing Month							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LMD$^-$	-0.054**	-0.055**	-0.070**	-0.052*	-0.091**	-0.027	-0.072	-0.007
	(-2.12)	(-2.11)	(-2.46)	(-1.94)	(-2.37)	(-0.76)	(-1.63)	(-0.25)
LMD$^-$ x 6-month Return	-0.137							
	(-1.13)							
LMD$^-$ x 1-year Return		-0.050						
		(-0.47)						
LMD$^-$ x Negative Return Dummy$_{6month}$			0.034					
			(0.95)					
LMD$^-$ x Negative Return Dummy$_{12month}$				-0.013				
				(-0.36)				
LMD$^-$ x N-CSR					0.052			
					(1.31)			
LMD$^-$ x No-Load Fund						0.031		
						(0.54)		
LMD$^-$ x Retail Fund							0.026	
							(0.51)	
LMD$^-$ x Young fund								-0.127**
								(-2.55)
Negative Return Dummy$_{6month}$			-0.107					
			(-1.00)					
Negative Return Dummy$_{12month}$				0.030				
				(0.26)				
N-CSR					-0.080			
					(-0.95)			
No-Load Fund						-0.104		
						(-0.41)		
Retail Fund							-0.016	
							(-0.11)	
Young Fund								0.323**
								(2.43)
Controls Table 2 Panel A Column (1)	Y	Y	Y	Y	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.386	0.386	0.386	0.386	0.386	0.377	0.387	0.386
Observations	23,098	23,098	23,098	23,098	23,098	14,176	22,968	23,098

Table IA-10: Letter tone and monthly fund flows: Interaction effects using NSAR/NPORT flows (cont'd)

Panel C: NSAR/NPORT Outflows Dependent Variable	Flow Filing Month							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LMD$^-$	0.062***	0.062***	0.064***	0.067***	0.033	0.059**	0.086**	0.053**
	(3.12)	(3.12)	(2.81)	(2.91)	(1.13)	(2.05)	(2.37)	(2.53)
LMD$^-$ x 6-month Return	0.031							
	(0.30)							
LMD$^-$ x 1-year Return		0.029						
		(0.37)						
LMD$^-$ x Negative Return Dummy$_{6month}$			-0.006					
			(-0.21)					
LMD$^-$ x Negative Return Dummy$_{12month}$				-0.014				
				(-0.51)				
LMD$^-$ x N-CSR					0.048			
					(1.59)			
LMD$^-$ x No-Load Fund						0.026		
						(0.60)		
LMD$^-$ x Retail Fund							-0.029	
							(-0.72)	
LMD$^-$ x Young fund								0.026
								(0.85)
Negative Return Dummy$_{6month}$			0.337***					
			(3.84)					
Negative Return Dummy$_{12month}$				0.211**				
				(2.43)				
N-CSR					-0.083			
					(-1.31)			
No-Load Fund						0.049		
						(0.20)		
Retail Fund							-0.095	
							(-0.82)	
Young Fund								0.164*
								(1.71)
Controls Table 2 Panel A Column (1)	Y	Y	Y	Y	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.330	0.330	0.331	0.330	0.330	0.326	0.330	0.330
Observations	22,885	22,885	22,885	22,885	22,885	14,049	22,755	22,885

Table IA-11: Shareholder letter tone and monthly fund flows: controlling for confounding events

This table shows regressions of monthly fund flows on shareholder letter tone and various fund characteristics as in Panel A of Table 2 of our paper but includes additional controls for confounding events. The dependent variable is the net fund flow (in percent) in the month of the SEC filing. We replace the flows of the filing month by the flows of the subsequent month whenever the filing date is after the 15th calendar day. Tone is the fraction of negative words in the shareholder letter based on the Loughran and McDonald (2011) (LMD^-) dictionary and is standardized to unit variance. We include all control variables and fixed effects from Panel A of Table 2 of our paper and additionally control for fund manager changes, fund name changes, changes in funds' investment objective, whether a share class of the fund has been liquidated, whether a share class has been merged into the fund, and whether there was an expense ratio change and the size of the change. These variables are based on changes from seven months before to one month after the filing month. All control variables are defined in detail in Section 3. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	Flow Filing Month			
	(1)	(2)	(3)	(4)
LMD^-	-0.138***	-0.135***	-0.115***	-0.116***
	(-3.76)	(-3.69)	(-3.22)	(-3.23)
Manager Change Dummy	-0.227***	-0.224***	-0.215***	-0.214***
	(-3.36)	(-3.30)	(-3.24)	(-3.22)
Fund Name Change Dummy	0.099	0.102	0.086	0.087
	(0.93)	(0.95)	(0.81)	(0.82)
Investment Obj. Change Dummy	-0.583**	-0.589**	-0.583**	-0.584**
	(-2.12)	(-2.15)	(-2.15)	(-2.15)
Share Class Liquidated Dummy	-0.495***	-0.490***	-0.498***	-0.499***
	(-2.97)	(-2.95)	(-3.00)	(-3.01)
Share Class Merged Dummy	-0.183	-0.186	-0.180	-0.180
	(-1.13)	(-1.14)	(-1.12)	(-1.12)
Expense Ratio Change Dummy	0.213***	0.213***	0.209***	0.210***
	(2.69)	(2.69)	(2.65)	(2.66)
Expense Ratio	-309.570***	-307.181***	-306.174***	-305.466***
	(-5.91)	(-5.85)	(-5.85)	(-5.83)
Controls Table 2 Panel A	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.194	0.195	0.197	0.197
Observations	33,398	33,398	33,398	33,398

Table IA-12: Shareholder letter tone and monthly fund flows: cubic performance term

This table shows regressions of monthly fund flows on shareholder letter tone and various fund characteristics as in specification (3) of Panel A of Table 2 of our paper but including a cubic performance term. The dependent variable is the net fund flow (in percent) in the month of the SEC filing. We replace the flows of the filing month by the flows of the subsequent month whenever the filing date is after the 15th calendar day. Tone is the fraction of negative words in a shareholder letter based on the LMD^{-} dictionary. Column (2) controls for 6-month return ranks instead of 1-year ranks. We include all control variables and fixed effects from column (3) of Panel A of Table 2 of our paper. All control variables are defined in detail in Section 3. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	Flow Filing Month	
	(1)	(2)
LMD^{-}	-0.111***	-0.121***
	(-3.07)	(-3.29)
$\text{Return}_{Report,Filing}$	12.908***	12.941***
	(12.01)	(12.06)
6-month Return	3.693***	
	(5.06)	
1-year Return		2.795***
		(4.63)
6-month Return Rank		3.268***
		(3.30)
6-month Return Rank^2		-6.534***
		(-2.98)
6-month Return Rank^3		4.999***
		(3.45)
1-year Return Rank	1.732*	
	(1.82)	
1-year Return Rank^2	-1.467	
	(-0.67)	
1-year Return Rank^3	1.388	
	(0.95)	
Controls Table 2 Panel A Column (3)	Y	Y
Fund FE	Y	Y
Reporting Month FE	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y
Adj. R^2	0.195	0.194
Observations	33,398	33,398

Table IA-13: Who signs the shareholder letter?

This table shows descriptive statistics on the job positions of individuals signing shareholder letters. Column (1) ((2)) displays the absolute (relative) frequency of each job position. As individuals can have multiple positions at the same point in time, we scale the frequencies by the number of job positions an individual signing a letter holds. For example, if the letter is signed by an individual being CEO and CFO at the same time, the absolute frequency of each job position, CEO and CFO, increases by 0.5 each.

Letter signed by	Abs. frequency (1)	Rel. frequency (2)
Fund Manager	5,481	15.47%
President	4,980	14.05%
Chairman	1,684	4.75%
Chief Executive Officer	1,306	3.69%
Managing Director	1,063	3.00%
Chief Investment Officer	560	1.58%
Chief Financial Officer	13	0.04%
Chief Operating Officer	11	0.03%
Unknown	20,339	57.39%
Total	35,437	100.00%

Table IA-14: Letter tone and monthly fund flows: controlling for the role of the person signing the letter

This table shows regressions of monthly fund flows on shareholder letter tone and various fund characteristics as in specifications (1) and (4) of Panel A in Table 2 of our paper but we control for the person signing the letter. The dependent variable is the net fund flow (in percent) in the month of the SEC filing. We replace flows of the filing month by flows of the subsequent month whenever the filing date is after the 15th calendar day. Tone is the fraction of negative words in a shareholder letter based on the LMD^{-} dictionary. We include all control variables and fixed effects from column (1) and (4) of Panel A in Table 2 of our paper, respectively. We additionally add dummy variables for the job positions of the individuals signing the letters and include interactions between the job positions and tone in columns (2) and (4). Descriptive statistics on the job positions of individuals signing shareholder letters are displayed in Table IA-13. All control variables are defined in detail in Section 3. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table IA-14: Shareholder letter tone and monthly fund flows: Controlling for the person signing the letter (cont'd)

Dependent Variable	Flow Filing Month			
	(1)	(2)	(3)	(4)
LMD$^-$	-0.135***	-0.132***	-0.113***	-0.108**
	(-3.62)	(-2.87)	(-3.09)	(-2.39)
Signed by CEO	0.033	0.177	0.062	0.206
	(0.20)	(0.65)	(0.39)	(0.75)
Signed by CFO/CIO/COO	0.230	0.123	0.229	0.085
	(0.72)	(0.35)	(0.72)	(0.24)
Signed by Chairman	-0.035	-0.072	-0.045	-0.067
	(-0.23)	(-0.32)	(-0.30)	(-0.30)
Signed by President	-0.001	-0.106	-0.011	-0.095
	(-0.01)	(-0.67)	(-0.09)	(-0.60)
Signed by Managing Director	-0.019	0.148	-0.022	0.169
	(-0.11)	(0.63)	(-0.13)	(0.72)
Signed by Fund Manager	-0.265**	-0.164	-0.269***	-0.181
	(-2.59)	(-0.88)	(-2.62)	(-0.97)
CEO x LMD$^-$		-0.080		-0.081
		(-0.59)		(-0.60)
CFO/CIO/COO x LMD$^-$		0.063		0.085
		(0.33)		(0.45)
Chairman x LMD$^-$		0.026		0.018
		(0.24)		(0.17)
President x LMD$^-$		0.056		0.044
		(0.76)		(0.61)
Managing Director x LMD$^-$		-0.112		-0.126
		(-0.94)		(-1.06)
Fund Manager x LMD$^-$		-0.055		-0.048
		(-0.60)		(-0.52)
Controls Table 2 Panel A	Column (1)	Column (1)	Column (4)	Column (4)
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.191	0.191	0.195	0.195
Observations	33,398	33,398	33,398	33,398

Table IA-15: Alternative tone measures and monthly fund flows

This table shows regressions of monthly fund flows on different measures of shareholder letter tone and various fund characteristics similar to the regression in column (1) of Panel A in Table 2 of our paper. The dependent variable is net fund flow (in percent) in the month of the SEC filing. We replace flows of the filing month by flows of the subsequent month whenever the filing of the shareholder letter takes place after the 15th calendar day. In column (1), we use an orthogonalized tone measure ($\text{LMD}^{-}_{adj.}$), which is the residuals from regressing LMD^{-} on the controls and fixed effects from column (1) of Table IA-6 of this Internet Appendix. In column (2), we measure tone by the change in negativity from last year's letter to the current letter (Δ LMD^{-}) and additionally control for the lagged fund flows, i.e., the flows in the filing month of last year's letter. Column (2) does not include fund fixed effects. In column (3), we weight negative words by their term frequency and inverse document frequency as in Loughran and McDonald (2011). In column (4), we include both positive and negative tone according to the Loughran and McDonald (2011) dictionaries (LMD^{+} and LMD^{-}). In columns (5) and (6), tone is defined according to the Harvard IV-4 psychosocial dictionary of negative words (HVD^{-}). In columns (7) and (8), we measure negativity based on our mutual fund dictionary of negative words MF^{-}. In columns (3), (6) and (8), we use the term frequency and inverse document frequency weighting from Loughran and McDonald (2011) ($\text{LMD}^{-}_{tf.idf}$, $\text{HVD}^{-}_{tf.idf}$, and $\text{MF}^{-}_{tf.idf}$). Control variables and fixed effects from column (1) of Panel A in Table 2 of our paper are always included in the regressions (except for the fund fixed effects in column (2)) and defined in detail in Section 3. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table IA-15: Alternative tone measures and monthly fund flows (cont'd)

Dependent Variable	Flow Filing Month							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\text{LMD}^{-}_{adj.}$	-0.076***							
	(-3.01)							
ΔLMD^{-}		-0.076**						
		(-2.22)						
$\text{LMD}^{-}_{tf.idf}$			-0.116***					
			(-3.19)					
LMD^{-}				-0.134***				
				(-3.62)				
LMD^{+}				0.011				
				(0.30)				
HVD^{-}					-0.034			
					(-0.93)			
$\text{HVD}^{-}_{tf.idf}$						-0.042		
						(-1.08)		
MF^{-}							-0.156***	
							(-4.31)	
$\text{MF}^{-}_{tf.idf}$								-0.141***
								(-3.77)
Controls Table 2 Panel A Column (1)	Y	Y	Y	Y	Y	Y	Y	Y
Fund FE	Y	N	Y	Y	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.191	0.132	0.191	0.191	0.191	0.191	0.191	0.191
Observations	33,398	27,766	33,398	33,398	33,398	33,398	33,398	33,398

Table IA-16: Other dimensions of tone: uncertainty and modal weak

This table shows regressions of monthly fund flows on shareholder letter tone and various fund characteristics as in column (1) of Panel A in Table 2 of our paper. The dependent variable is the net fund flow (in percent) in the month of the SEC filing. We replace flows of the filing month by flows of the subsequent month whenever the filing date is after the 15th calendar day. Negativity (uncertainty, modal weak) is the fraction of negative (uncertainty, modal weak) words in a shareholder letter based on the Loughran and McDonald (2011) dictionaries. Columns (2) and (4) include the interaction effects between uncertainty and negativity and modal weak and negativity, respectively. Control variables and fixed effects from column (1) of Panel A in Table 2 of our paper are always included in the regressions and defined in detail in Section 3. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	Flow Filing Month			
	(1)	(2)	(3)	(4)
LMD^{-}	-0.137***	-0.157**	-0.135***	-0.193***
	(-3.64)	(-2.30)	(-3.66)	(-3.66)
$\text{LMD}^{uncertainty}$	0.029	0.007		
	(0.71)	(0.09)		
LMD^{-} x $\text{LMD}^{uncertainty}$		0.012		
		(0.35)		
$\text{LMD}^{modal\ weak}$			-0.017	-0.100
			(-0.47)	(-1.62)
LMD^{-} x $\text{LMD}^{modal\ weak}$				0.050
				(1.63)
Controls Table 2 Panel A Column (1)	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Combined Filing Month x Inv. Obj. FE	Y	Y	Y	Y
Adj. R^2	0.191	0.191	0.191	0.191
Observations	33,398	33,398	33,398	33,398

Table IA-17: Shareholder letter tone and future fund performance: alternative clustering of standard errors

This table shows regressions of fund performance (in %) on shareholder letter tone and various fund characteristics as in Table 7 of our paper. In this table, standard errors are double-clustered at the fund family and time (not fund and time as in the paper) dimension. The dependent variable is the fund's return (column (1)), CAPM 1-factor alpha (column (2)), and Carhart (1997) 4-factor alpha (column (3)). Performance is measured from one (t+1) to six months (t+6) after the filing date. Alphas are computed using beta coefficients obtained from a regression using daily fund returns over the previous twelve months (t-12 to t-1). In Panel A, tone is the fraction of negative words based on the Loughran and McDonald (2011) dictionary. In Panel B, we split negativity into negativity in forward-looking and non-forward-looking sentences. Forward-looking sentences are defined as in Li (2010). All control variables are defined in detail in Section 3. LMD^- is standardized to unit variance. All regressions use the Pástor, Stambaugh, and Taylor (2015) recursive demeaning estimator which recursively forward-demeans all variables. We instrument for forward-demeaned negativity, letter length, fund size, lagged 1-year fund returns and flows using the corresponding backward-demeaned variables. Regressions include time fixed effects for the reporting and filing month. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Negativity and future fund performance			
Dependent variable	Raw Return	1-Factor Alpha	4-Factor Alpha
	(1)	(2)	(3)
LMD^-	0.007	-0.137**	-0.032
	(0.08)	(-2.22)	(-0.58)
ln(Fund size)	-1.925	-0.529	-0.387
	(-1.51)	(-0.54)	(-0.53)
ln(Family size)	0.241	-0.058	-0.041
	(0.70)	(-0.21)	(-0.21)
ln(fund age)	1.399*	-0.465	-0.347
	(1.83)	(-1.09)	(-0.91)
Expense ratio	-255.183	31.894	48.176
	(-1.18)	(0.20)	(0.42)
1-year Return	0.003	-0.015	-0.027
	(0.09)	(-0.52)	(-1.13)
1-year Flow	-1.649*	-0.527	-0.351
	(-1.66)	(-0.74)	(-0.72)
1-year Return Family	-0.070*	-0.035	-0.009
	(-1.92)	(-1.36)	(-0.45)
1-year Flow Family	0.941	-0.058	-0.237
	(0.87)	(-0.08)	(-0.44)
ln(Words)	0.094	-0.007	0.018
	(0.38)	(-0.05)	(0.14)
Time Difference	-0.008	0.006	0.016
	(-0.27)	(0.36)	(1.10)
Recursive demeaning	Y	Y	Y
Fund FE	N	N	N
Reporting Month FE	Y	Y	Y
Filing Month FE	Y	Y	Y
Adj. R^2	0.009	0.014	0.010
Observations	34,539	34,539	34,539

Table IA-17: Shareholder letter tone and future fund performance: alternative clustering of standard errors (cont'd)

Panel B: Forward-looking negativity and future fund performance			
Dependent variable	Raw Return	1-Factor Alpha	4-Factor Alpha
	(1)	(2)	(3)
Forward-looking LMD$^-$	0.018	0.025	0.018
	(0.25)	(0.51)	(0.39)
Non-forward-looking LMD$^-$	-0.029	-0.159***	-0.036
	(-0.34)	(-2.62)	(-0.64)
Forward-looking sentences	-0.933	-1.218	-0.651
	(-0.69)	(-1.61)	(-0.98)
Controls from Panel A	Y	Y	Y
Recursive demeaning	Y	Y	Y
Fund FE	N	N	N
Reporting Month FE	Y	Y	Y
Filing Month FE	Y	Y	Y
Adj. R^2	0.009	0.014	0.010
Observations	34,539	34,539	34,539

Table IA-18: Shareholder letter tone and managerial risk-taking: alternative clustering of standard errors

This table shows regressions of fund risk and investment style on shareholder letter tone and various fund characteristics as in Table 8 of our paper. In this table, standard errors are double-clustered at the fund family and time (not fund and time as in the paper) dimension. Tone is the fraction of negative words based on the Loughran and McDonald (2011) (LMD^-) dictionary. LMD^- is standardized to unit variance. The dependent variable is the fund's total risk (column (1)), systematic risk (column (2)), idiosyncratic risk (column (3)), and style extremity (column (4)). Fund risk is the standard deviation of daily returns from one to six months after the filing month. Systematic (idiosyncratic) risk is the market beta (standard deviation of residuals) obtained from estimating a Carhart 4-Factor-Model using daily data from one to six months after the filing month. Total risk and idiosyncratic risk are expressed in %. Style extremity is the absolute difference of a fund's SMB-, HML-, and UMD-beta from the average beta in the fund's investment objective. Betas are calculated using daily returns over the twelve months after the filing month. For a detailed description of style extremity, see Baer, Kempf, and Ruenzi (2011). All control variables are defined in detail in Section 3. All regressions include fund fixed effects and time fixed effects for the month of the fiscal (half-) year end (report month) and for the month when the filing takes place (filing month). t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent variable	Total Risk	Systematic Risk	Idiosyncratic Risk	Style Extremity
	(1)	(2)	(3)	(4)
LMD^-	0.002	0.001	-0.002**	-0.010**
	(0.87)	(0.55)	(-2.08)	(-2.05)
ln(Fund size)	0.024***	0.016***	0.009***	0.022**
	(6.12)	(5.85)	(4.48)	(2.34)
ln(Family size)	0.003	0.003	-0.003	-0.001
	(0.82)	(1.38)	(-1.53)	(-0.16)
ln(fund age)	-0.046***	-0.029***	-0.014**	-0.088***
	(-4.35)	(-3.98)	(-2.46)	(-3.32)
Expense ratio	5.033**	3.188**	1.928**	-0.895
	(2.45)	(2.25)	(2.12)	(-0.22)
ln(Words)	0.005	0.003	0.004**	0.007
	(1.09)	(1.18)	(2.18)	(0.84)
Time Difference	0.001	0.001	-0.000	-0.001
	(1.44)	(1.63)	(-0.26)	(-0.34)
1-year Return	0.050	0.060**	0.047***	0.097*
	(1.32)	(2.48)	(2.84)	(1.77)
1-year Flow	-0.000	-0.006**	-0.003	-0.015
	(-0.01)	(-2.55)	(-1.14)	(-1.55)
1-year Return Family	-0.005	0.024	0.006	-0.176**
	(-0.14)	(1.14)	(0.28)	(-2.55)
1-year Flow Family	-0.019	-0.006	0.006	0.014
	(-1.37)	(-0.93)	(0.84)	(0.46)
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Filing Month FE	Y	Y	Y	Y
Adj. R^2	0.901	0.703	0.862	0.577
Observations	33,699	33,678	33,678	32,627

Table IA-19: Top 50 negative mutual fund dictionary words

This table shows the 50 most frequent negative words according to our mutual fund dictionary of negative shareholder letter words (MF^-). Absolute frequency is the total number of occurrences of the negative word across all shareholder letters in the sample. The last column shows whether the word is also part of the Loughran and McDonald (2011) dictionary of negative words.

Rank	Negative word	Abs. frequency	Included in LM dictionary
1	volatility	32,236	Yes
2	concerns	22,889	Yes
3	negative	20,614	Yes
4	detracted	19,951	Yes
5	inflation	18,119	No
6	declined	15,939	Yes
7	uncertainty	13,793	No
8	underperformed	13,756	Yes
9	crisis	13,651	Yes
10	decline	13,471	Yes
11	fluctuate	12,823	No
12	detractors	12,368	No
13	weak	11,243	Yes
14	hurt	11,020	Yes
15	losses	9,943	Yes
16	recession	9,706	Yes
17	weakness	9,690	Yes
18	poor	9,662	Yes
19	volatile	9,176	Yes
20	detractor	8,636	No
21	lagged	8,049	Yes
22	slowing	7,976	Yes
23	underperformance	7,789	Yes
24	unemployment	7,718	Yes
25	worst	7,390	Yes
26	defensive	7,203	Yes
27	sharply	7,175	Yes
28	difficult	7,004	Yes
29	slowdown	6,901	Yes
30	challenges	6,867	Yes

Table IA-19: Top 50 negative mutual fund dictionary words (cont'd)

Rank	Negative word	Abs. frequency	Included in LM dictionary
31	pressure	6,649	No
32	fears	6,560	Yes
33	declines	6,029	Yes
34	slow	5,843	Yes
35	sharp	5,610	No
36	weaker	5,604	Yes
37	disappointing	5,574	Yes
38	suffered	5,507	Yes
39	declining	5,412	Yes
40	challenging	5,076	Yes
41	lost	4,957	Yes
42	negatively	4,917	Yes
43	downturn	4,658	Yes
44	sell-off	4,621	No
45	headwinds	4,386	No
46	concern	4,329	Yes
47	mixed	4,110	No
48	hard	3,927	No
49	pressures	3,785	No
50	struggled	3,707	No

Table IA-20: Investors' reaction to market/economy-related and to non-market/related letter content

This table is similar to table 4 in our paper and shows regressions of monthly fund flows on negativity in market/economy and in non-market/economy-related sentences of shareholder letters and various fund characteristics. In column (1), the dependent variable is the net fund flow (in percent) in the month of the SEC filing from the CRSP database. In columns (2) to (4), we use the net fund flows, inflows, and outflows from NSAR and NPORT filings, respectively. We replace flows of the filing month by flows of the subsequent month whenever the filing date is after the 15th calendar day. We distinguish between negativity according to the Loughran and McDonald (2011) dictionary in market/economy-related sentences and in non-market/economy-related sentences. Market and economy related sentences are identified by the key words "economy" and or "market." Control variables and fixed effects from column (1) of Panel A in Table 2 of our paper are always included in the regressions and defined in detail in Section 3. Standard errors are double-clustered at the fund and time (filing month) dimension. t-statistics are provided in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Dependent Variable	CRSP	NSAR/NPORT		
	net flows	net flows	inflows	outflows
	(1)	(2)	(3)	(4)
Market/economy LMD$^-$	-0.059	-0.060**	-0.011	0.045***
	(-1.61)	(-2.18)	(-0.49)	(2.81)
Non-market/economy LMD$^-$	-0.067*	-0.065**	-0.037	0.042**
	(-1.80)	(-2.28)	(-1.53)	(2.31)
Market/economy sentences	-0.293	-0.036	-0.185	-0.159
	(-1.02)	(-0.17)	(-1.04)	(-0.99)
Controls Table 2 Panel A Column (1)	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y
Reporting Month FE	Y	Y	Y	Y
Filing Month FE	Y	Y	Y	Y
Adj. R^2	0.191	0.267	0.386	0.330
Observations	33,398	22,881	23,098	22,885

LEIBNIZ INSTITUTE SAFE Sustainable Architecture for Finance in Europe

Recent Issues

No. 379	Irene Mecatti, Tobias Tröger	Who Should Hold Bail-Inable Debt and How can Regulators Police Holding Restrictions Effectively?
No. 378	Sabine Bernard, Benjamin Loos, Martin Weber	How Speculative Asset Characteristics Shape Retail Investors' Selling Behavior
No. 377	Konstantin Egorov	Optimal Policy under Dollar Pricing
No. 376	Raphael Abiry, Marien Ferdinandusse, Alexander Ludwig, Carolin Nerlich	Climate Change Mitigation: How Effective is Green Quantitative Easing?
No. 375	Olivia S. Mitchell, Orly Sade, Abigail Hurwitz	Testing Methods to Enhance Longevity Awareness
No. 374	Franco Fiordelisi, Giulia Fusi, Angela Maddaloni, David Marqués-Ibáñez	Pandemic Lending: Micro and Macro Effects of Model-Based Regulation
No. 373	Emily Kormanyos, Tobin Hanspal, Andreas Hackethal	Do Gamblers Invest in Lottery Stocks?
No. 372	Alessandro Pollastri, Paulo Rodrigues, Christian Schlag, Norman Seeger	A Jumping Index of Jumping Stocks? An MCMC Analysis of Continuous-Time Models for Individual Stocks
No. 371	Mariano Massimiliano Croce, Tatyana Marchuk, Christian Schlag	The Leading Premium
No. 370	Benjamin M. Abdel-Karim, Kevin Bauer, Oliver Hinz, Michael Nofer	The Effects of Discontinuing Machine Learning Decision Support
No. 369	Katja Langenbucher	Consumer Credit in The Age of AI – Beyond Anti-Discrimination Law
No. 368	Vanya Horneff, Raimond Maurer, Olivia S. Mitchell	How Would 401(k) 'Rothification' Alter Saving, Retirement Security, and Inequality?
No. 367	Jens Lausen, Benjamin Clapham, Peter Gomber, Micha Bender	Drivers and Effects of Stock Market Fragmentation - Insights on SME Stocks
No. 366	Satchit Sagade, Stefan Scharnowski, Christian Westheide	Broker Colocation and the Execution Costs of Customer and Proprietary Orders