Soothing Investors: The Impact of Manager Communication on Mutual Fund Flows

Ahmed Guecioueur*

August 30, 2023
Latest: ahmedgc.com/files/latest.pdf

Abstract

I uncover an economically meaningful channel through which fund managers foster investor risk-taking, build trust, and impact asset prices. I measure communication by fund managers about unrealized risks to their clients, and document its prevalence. Using an institutional setting that enables causal identification, I find that increased communication about future risks encourages investors to *increase* their holdings in the market portfolio. I rule out belief-based persuasion, uncertainty reduction, ambiguity alleviation and other potential explanations. Additional analyses show this communication soothes investors' anxiety and alleviates their effective risk aversion, consistent with the money doctors framework of Gennaioli, Shleifer, and Vishny (2015).

Keywords: Fund Flows, Anxiety, Trust, Persuasion, Text Data
JEL Codes: G11, G23, G41, G50, D01, D83

*INSEAD, Boulevard de Constance, 77300 Fontainebleau, France, email: ahmed.guecioueur@insead.edu
I would like to thank Alexandru Barbu, Pedro Bordalo, Aditya Chaudhry, Olivier Dessaint, Tilman Fries (discussant), Andreas Fuster, Sergei Glebkin, Naveen Gondhi, Richard Grice, Markus Ibert, Pascal Maenhout, Massimo Massa, Cristina Ortiz (discussant), Joël Peress (my advisor), Alessandro Previtero, Rodolfo Prieto, Dmitrii Pugachev, Vesa Pursiainen, Lucie Teplá, Scott Yonker (discussant), Bart Zhou Yueshen, and many others, as well as participants at the 2023 Financial Intermediation Research Society (FIRS) Conference PhD Sessions, the 7th SAFE Household Finance Workshop, the 16th Behavioural Finance Working Group Conference, the LSE-Oxford-Imperial Behavioural Finance Group, the 16th International Risk Management Conference, the 11th Wharton-INSEAD Doctoral Consortium and the 7th HEC Paris Finance PhD Workshop for their helpful comments. I am grateful to Yehuda Izhakian, Gerald S. Martin, Melissa Porras Prado and Rüdiger Weber for sharing data, and to Google for granting me access to its Trends API. I gratefully acknowledge the support of the Henry Grunfeld Foundation. A previous version of this paper was circulated under the title "Money Doctors and Their Prognoses."

1 Introduction

Investing is a risky endeavor. Low equity holding rates indicate that many individuals are reluctant to bear this risk.[1] To the extent that certain behavioral biases are at fault, these might be overcome by governments nudging savers to take more risks by regulating their choice architecture (Thaler and Sunstein, 2008), or by the asset management industry designing investment products that foster risk-taking (Calvet, Célérier, Sodini, and Vallée, 2023). However, the role of persuasion has received less attention, despite the prevalence of communication by asset managers. Can investors be encouraged to bear risk through information provision?

Gennaioli, Shleifer, and Vishny (2015, henceforth **GSV**) argue that trusted financial intermediaries can give their clients the confidence to overcome their anxieties and take risks, including through "communication" and "persuasive advertising." More generally, persuasion is a commonplace economic activity (McCloskey and Klamer, 1995) and can act through different mechanisms, including by inducing changes in the beliefs or preferences of recipients (DellaVigna and Gentzkow, 2010).[2]

If investors experience anxiety about risks that are yet to be realized, it is possible that communication about those future risks could soothe their concerns and encourage them to accept and bear risk — effectively reducing their risk aversion.[3] In fact, mere exposure to and familiarity with a stimulus reduces negative attitudes towards it while increasing preferences for it (Fechner, 1876; Zajonc, 1968; Harmon-Jones and Allen, 2001; Tom, Nelson, Srzentic, and King, 2007). Motivated both by this evidence and by GSV's model of anxiety-reduction as a reduction in risk aversion, I investigate how investors adjust their equity holdings in response to the provision of information about future risks, including the level of detail conveyed.

Mutual funds typically communicate with their investors at regular intervals (Hillert, Niessen-Ruenzi, and Ruenzi, 2021). In this paper, I examine responses to the dissemination of free-form letters written to investors in mutual funds that track the market portfolio (S&P 500

[1] The majority of the US population does not participate at all in the stock market (Mankiw and Zeldes, 1991; Haliassos and Bertaut, 1995; Melcangi and Sterk, 2020). In addition, many households that *do* hold stocks hold a low share of their financial wealth in them (Calvet, Campbell, and Sodini, 2007; Calvet, Célérier, Sodini, and Vallée, 2023). This paper focuses on investments in diversified large-cap equity portfolios such as the S&P 500, which are generally considered beneficial for households' welfare (in contrast to trading in individual stocks).

[2] Like McCloskey and Klamer (1995) & DellaVigna and Gentzkow (2010), I use the term "persuasion" in a very general sense. Two streams of the belief-based persuasion literature reserve this label either for games with verifiable information (Milgrom and Roberts, 1986), or for "Bayesian persuasion" models with commitment (Kamenica and Gentzkow, 2011). The present paper does not restrict itself to either.

[3] The evidence that risk aversion is not stable within individuals is extensive (Schildberg-Hörisch, 2018). In particular, heightened anxiety causes an increase in risk aversion (Raghunathan and Pham, 1999; Lerner and Keltner, 2001; Maner and Schmidt, 2006; Kuhnen and Knutson, 2011; Giorgetta et al., 2012; Guiso, Sapienza, and Zingales, 2018). The medical literature has long studied the anxiety-alleviating benefits of information: Hayward (1975) finds that providing information to surgical patients results in lower reported anxiety and pain.

index funds).[4] I find that fund managers who provide more detailed forecasts about future risks in their fund letters – in the sense of dedicating a larger proportion of the letter to discussing future risk – enjoy *higher* net flows. The flow effect is economically significant: a 1 standard deviation growth in the level of detail communicated predicts inflows of 0.23 standard deviations in magnitude. The micro-level relation is present both over time and in the cross-section of S&P 500 index funds. Decomposing net flows into gross inflows and outflows, I find that the effect is driven by increased inflows. Given that the communications I examine are targeted at existing investors, my results suggest that communication about future risks encourages existing investors to increase their holdings, and thus to bear additional risk. The findings are consistent with a mechanism similar to GSV's in which communication by trusted money doctors to their investor "patients" can decrease investors' effective risk aversion.

I consider, and ultimately rule out, a number of alternative explanations for my findings. *First*, investors might simply be updating their beliefs about the level of risk, which they may infer from either the amount of detail managers provide about it, or the sentiment they convey when discussing this topic. Indeed, managers seem to be honest: within-manager variation in both the level of communication and sentiment expressed about future risks appears to faithfully convey the information relayed in the VIX index, and so investors *should* be able to infer a measure of future risks, and therefore reduce or delay their investment accordingly (Hassan, Hollander, van Lent, and Tahoun, 2019; Bloom, 2009; Baker, Bloom, and Davis, 2016). Yet investors do not act this way: I find that they increase their allocations even in the face of openly pessimistic forecasts, and their actions do not vary based on the relative optimism of communicated beliefs compared to their priors. Furthermore, the strength of investors' priors about risk has no bearing on the strength of their actions, which argues against belief-updating as the underlying mechanism. *Second*, it is possible that more communication about risks might lead investors to infer a higher level of risk in general, and thus to rebalance from riskier assets into less risky ones (i.e. engage in a flight to quality). To the extent that investors perceive the market portfolio to be relatively safe, this could explain the observed flows to S&P 500 index funds. However, I find no evidence of this within the same fund family, which is precisely where any rebalancing should be the least costly (Massa, 2003). *Third*, receiving signals about an asset's payoff might simply mechanically decrease its perceived risk, leading to a higher portfolio allocation (e.g. Grossman and Stiglitz, 1980). Using a placebo analysis of the contents of fund letters, I show such a mechanism cannot explain my empirical findings on the effect of more detailed communication about risk itself. *Fourth*, it might be the case that communication about future risks is intended to alleviate investors'

[4]Focusing on this homogeneous class of passive funds aids in identification (Hortaçsu and Syverson, 2004). First, objective measures of risk and return are common across funds. Second, there is no role for active manager skill in this setting, which rules out a channel in which investors attempt to learn about it based on communication.

ambiguity aversion. However, I find that investors tend to ignore the letters in times of high perceived ambiguity – precisely the time when one would expect communication to matter the most. *Finally*, I conduct a battery of robustness checks to rule out other hypotheses based on financial literacy, pandering-based persuasion, preferences over the timing of uncertainty resolution, or signaling about funds' tracking errors.

Instead, I find support for an anxiety-alleviating mechanism. I geographically locate the readers of fund letters in order to infer their current anxiety attitudes: to do this, I use the frequency of local Google searches (comprising a multitude of phrases, such as "stressed", "anxiety" & "stop worrying"). I then aggregate up to a fund letter readership level by taking an asset-weighted average of fund letter readers' local anxiety attitudes. I find that low-anxiety readers barely respond to communication. By contrast, high-anxiety readers respond strongly and positively: the effect of communication on fund flows is concentrated among investor bases with above-median anxiety levels, as proxied by Google search volume. This finding indicates that readers' prior high anxiety levels had been holding them back from bearing risks – until being soothed by communication. The finding is thus consistent with GSV's hypothesis that financial intermediaries can alleviate investors' anxiety and give them the confidence to bear risk. By alleviating investors' effective risk aversion, this communications acts as a form of preference-based persuasion (DellaVigna and Gentzkow, 2010).

I next uncover the key role played by trust. The money doctors theory holds that *trusted* intermediaries, specifically, can alleviate investors' anxiety. While trust levels are hard to measure, prior work uses negative trust shocks to trace out the effect of trust on investors' behavior (Kostovetsky, 2016; Giannetti and Wang, 2016; Gurun, Stoffman, and Yonker, 2018). Likewise, I find that negative shocks to trust dampen the effectiveness of communication as an anxiety-reducing tool. Using the exogenous revelation of local corporate securities fraud (Giannetti and Wang, 2016; Karpoff, Koester, Lee, and Martin, 2017), I find that more detailed communications about risks to investors with a high exposure to such negative local trust shocks produce no discernible flows. By contrast, when a fund letter's readership has a low (or no) exposure to such negative trust shocks, the persuasive effect is strong and statistically significant. These results apply within-fund (i.e. holding the identify of a fund constant) and within the cross-section of funds (i.e. incorporating time effects). This set of findings confirms that trust is a necessary ingredient for money doctors to reassure anxious investors through communication.

Finally, I quantify the aggregate asset pricing implications, which are economically important. My focus on funds that track the market portfolio allows me to measure the extent to which communication-driven inflows into the stock market affect the total market value. To translate fund flow effects into return effects, I use Gabaix and Koijen's (2021) estimate of the macro multiplier of stock market inflows. A back-of-the-envelope estimate is that the

communication-driven net flows in my sample are responsible for an increase of almost 1% a year in the value of the S&P 500. That is, the S&P 500 would have been approximately 11% to 14% lower at the end of my sample (in August 2021) without the effect of anxiety-alleviating communication by my sample of fund managers over the period 2006-2021. This simple analysis suggests that the asset pricing implications of anxiety alleviation are meaningful in aggregate.

Related Literature

This study contributes to the literature on persuasion – a phenomenon that is prevalent in economic settings (McCloskey and Klamer, 1995). My findings uncover a persuasive mechanism in which financial intermediaries encourage risk-taking by anxious households, and with economically meaningful effects. I provide evidence that the persuasion occurs via what DellaVigna and Gentzkow (2010) term a preference-based channel, rather than a belief-based form of persuasion. Preference-based persuasion has hitherto been examined mainly in the context of advertising (for example, Becker and Murphy, 1993).

Relatedly, this paper contributes to a growing behavioral literature on the noninstrumental or hedonic nature of information, which studies agents who do not use information to form more accurate beliefs, but for other purposes instead. This prior work tends to focus on the potential drawbacks of information aversion (Andries and Haddad, 2020) and selective attention (Sicherman, Loewenstein, Seppi, and Utkus, 2016), or considers the case when information is used as entertainment (Ely, Frankel, and Kamenica, 2015). My study contributes evidence on the bright side of noninstrumental or hedonic information: by calming investors' effective risk aversion, communication by fund managers has the effect of encouraging more stock market participation. To the extent that a greater share of wealth held in the stock market is beneficial (Calvet, Campbell, and Sodini, 2007; Calvet, Célérier, Sodini, and Vallée, 2023), I therefore contribute evidence on the benefits of persuasion and social influence in an important financial setting.[5]

I contribute to the literature on affect, emotion and the psychological costs of financial markets. Loewenstein, Weber, Hsee, and Welch (2001) argue via their "risk-as-feelings" hypothesis that emotions play a crucial part in individuals' risk-related behavior, over and above cognitive-consequentialist processes. Engelberg and Parsons (2016) show that stock market-induced anxiety increases hospital admissions. Sergeyev, Lian, and Gorodnichenko (2023) show that individual psychic costs aggregate up to macroeconomic outcomes. To my knowl-

[5]Besides being a form of persuasion, such communications are a social interaction (see Kuchler and Stroebel (2021)), and may also be a means of maintaining relationships (see Crawford (2019)) or an appeal to investors' "animal spirits" (Akerlof and Shiller, 2010). Communication tools are of interest to policymakers (Blinder, 2009).

edge, my paper is the first to link the anxiety attitudes of mutual fund investors to their behavior, and furthermore uncover a mechanism that alleviates such psychological costs.

My empirical evidence lends support to the money doctors hypothesis of GSV, which links trust to a reduction in investor anxiety and risk aversion. The retail investors surveyed by Choi and Robertson (2020) state that trust-related concerns are important determinants of their investment decisions. Kostovetsky (2016) and Gurun, Stoffman, and Yonker (2018) both analyze the effects of a breakdown of trust on the behavior of the clients of different types of financial intermediary, without examining what actions these intermediaries take to foster risk-taking in the first place. To my knowledge, my study is the first to illuminate by what actions such trusted intermediaries can foster risk-taking beyond an initial matching or allocation – specifically, by communicating about risk.[6]

By identifying communication as an economically meaningful determinant of flows into S&P 500 index funds, this paper contributes direct evidence on the drivers of flows into the aggregate stock market, answering the call by Gabaix and Koijen (2021). I also add to a long literature documenting the drivers of mutual fund investors' capital allocation decisions by studying fund flows, including Sirri and Tufano (1998), Chevalier and Ellison (1999), Jain and Wu (2000), Berk and Green (2004), Huang, Wei, and Yan (2007), Del Guercio and Tkac (2008), Bailey, Kumar, and Ng (2011), and many others.

I contribute novel findings on persuasion to the literature on information transmission in the asset management industry. This literature has mostly concentrated on disclosure and manager skill; for example, Cassar et al. (2018) analyze the choice of hedge funds to disclose past information. Some studies examine the contents of mandatory prospectus disclosures: Abis, Buffa, Javadekar, and Lines (2022), Sheng, Xu, and Zheng (2022) and Mitali (2022) analyze strategy descriptions, risk factor exposures, and confidence in writing style (respectively). Other studies examine different forms of voluntary communication: Hillert, Niessen-Ruenzi, and Ruenzi (2021) and Du, Jiao, Ye, and Fan (2019) correlate the tone and confidence in writing style (respectively) of shareholder letters with investor outcomes, and Gil-Bazo and Imbet (2022) correlate the sentiment of fund manager tweets with investor outcomes. Cao, Yang, and Zhang (2021) relate the sentiment of risk-related discussions by fund managers to their performance. I differ in my approach of examining the forward-looking

[6]These results on trust have implications far beyond the money doctors framework. Trust has been linked to economic growth (Fukuyama, 1995; Knack and Keefer, 1997; Algan and Cahuc, 2010), financial development (Guiso, Sapienza, and Zingales, 2004), investor behavior (Guiso, Sapienza, and Zingales, 2008; Giannetti and Wang, 2016; Wei and Zhang, 2020; Massa, Wang, Zhang, and Zhang, 2022), and other economic outcomes (Arrow, 1972). That communication by a trusted sender can foster risk-taking complements prior findings: in cross-country comparisons, Fisman and Khanna (1999) found trust is increasing in the ease of communication. Earlier works by Deutsch (1958) and Loomis (1959) find that communication establishes trust between players in a non-cooperative game. Morgan and Hunt (1994) argue that communication fosters trust and thereby contributes to the continuation of ongoing customer-supplier relationships.

risk-related statements transmitted by S&P 500 tracker funds, which aids in identification. To my knowledge, this paper is also the first to measure the prior attitudes (and prior beliefs) of fund letter readers. My approach uncovers a persuasion channel in which fund managers influence investors' preferences through communication, and highlights the roles of trust & anxiety.

A smaller literature studies advertising by mutual fund families (Jain and Wu, 2000; Gallaher, Kaniel, and Starks, 2015; Aydogdu and Wellman, 2011). Advertising may attract new investors by raising awareness or attracting attention. Focusing on existing investors rules out such channels, and exposes a persuasion effect.

2 Money Doctors

This section briefly reviews the core of GSV's money doctors theory, which provides a concise empirical framework for how communication can encourage investor risk-taking.

Classic Portfolio Choice Consider the simple case of an investor who can access a riskless asset that returns R_0 and a risky asset with excess return R. I interpret the risky asset as the market portfolio; note that the present paper studies investors in S&P 500 index funds. The investor will allocate fraction x of her portfolio to the risky asset and the remainder $(1-x)$ to the riskless asset. Assume the investor has mean-variance preferences and γ is her coefficient of risk aversion.[7] She therefore seeks to maximize the utility function

$$U(x) = R_0 + x\mathbb{E}[R] - \frac{\gamma}{2}x^2\mathbb{V}\text{ar}(R), \tag{1}$$

and will therefore allocate the fraction

$$x = \frac{\mathbb{E}[R]}{\gamma\mathbb{V}\text{ar}(R)} \tag{2}$$

of her portfolio to the risky asset, based upon her risk aversion and return expectations.

[7]The assumptions of mean-variance preferences and a static horizon are not crucial. A dynamic continuous-time investor behaves identically under various sets of assumptions on the utility function, investment opportunity set, and so on; in any case, the mean-variance optimal portfolio is an important component of dynamic portfolio choice. In static settings, a log utility investor behaves similarly (Koijen and Yogo, 2019, pp. 1481). Under widely-used CARA-Normal assumptions, the behavior is identical.

Incorporating Anxiety GSV augment the investor's utility function with the parameter $a \geq 1$, which they interpret as the "anxiety suffered by [the investor] for bearing risk:"

$$U(x) = R_0 + x\mathbb{E}[R] - \frac{a\gamma}{2}x^2\mathbb{V}\text{ar}(R), \tag{3}$$

This now leads the investor to allocate a (weakly) smaller portfolio share to the risky asset:

$$x = \frac{1}{a\gamma} \cdot \frac{\mathbb{E}[R]}{\mathbb{V}\text{ar}(R)}. \tag{4}$$

The limit $a \to \infty$ captures the behavior of an investor who experiences so much anxiety that she does not participate in the stock market at all. Conversely, the lower the anxiety a that the investor experiences, the higher the fraction x of wealth that she invests in the risky asset (instead of the riskless one). Thus the term $a\gamma$ is the investor's effective risk aversion, which varies with her anxiety levels.

Money Doctors Can Reduce Investor Anxiety and Build Trust GSV argue that financial intermediaries play a key role in giving investors the confidence to take risks. In their model, a money manager (or "money doctor") can reduce the anxiety a of an investor. Money doctors thus "provide investors peace of mind." Furthermore, GSV define "trust" as "reducing investor anxiety about taking risk;" under this view, trust and anxiety are inextricably linked. Finally, GSV suggest some possible sources of trust: "personal relationships, familiarity, persuasive advertising, connections to friends and colleagues, communication, and schmoozing."

Consider an investor in a mutual fund. Assuming that $\mathbb{E}[R] > 0$, Eqn. (4) predicts that

$$\frac{\partial x}{\partial a} < 0. \tag{5}$$

That is, holding the investor's beliefs constant, any decreases in her anxiety $\Delta a < 0$ (or, equivalently, increases in her trust) encourage her to increase the fraction of her wealth x that she places in the fund. These increases in her portfolio allocation $\Delta x > 0$ empirically manifest as inflows into the fund.

Existing Evidence There are two streams of literature that support the arguments made by GSV. The first examines exogenous trust-breaking events and finds they lead to outflows, consistent with Eqn. (5). This includes observational studies by Kostovetsky (2016) examining mutual fund clients' flows, and by Gurun, Stoffman, and Yonker (2018) analyzing investment advisor clients' flows. The second stream of literature directly measures or isolates the effective risk aversion $a\gamma$ of investors, and finds it does indeed increase with their anxiety a. This

includes experimental studies by Raghunathan and Pham (1999), Lerner and Keltner (2001), Maner and Schmidt (2006), Kuhnen and Knutson (2011), Giorgetta et al. (2012), and Guiso, Sapienza, and Zingales (2018).

Roles for Communication Equation (4) makes clear that communication can influence investor risk-taking along two separate paths.

First, under a classic view, communication might alter the beliefs $\mathbb{E}[R]$ and $\mathbb{V}\text{ar}(R)$ that investors hold about their future payoff. In the case of active management, these would be dependent on active fund manager skill, while this paper focuses on the passive market portfolio. Nevertheless, information of all kinds could shift these beliefs, as might belief-based persuasion (DellaVigna and Gentzkow, 2010).

The second path for communication to foster risk-taking is by reducing the investor's anxiety a. As per Eqn. (5), a mutual fund investor whose anxiety is soothed by this communication would react with inflows into the fund. A number of experimental studies show that providing information decreases anxiety. In a medical context, Hayward (1975) finds that surgical patients report lower anxiety (and even less physical pain) when provided with information. In a more closely related setting, Eliaz and Schotter (2010) find that subjects who must bear a risky prospect exhibit a demand for information even when it cannot be used instrumentally. Furthermore, the subjects reported that possessing information gave them "confidence in [their] decision" and "a sense of security" (Eliaz and Schotter, 2010, pp. 316). Since effective risk aversion $a\gamma$ is interpretable as a preference parameter, communication of this type can be categorized as preference-based persuasion (DellaVigna and Gentzkow, 2010).

Empirical Approach The GSV empirical framework informs this paper's approach to studying the impact of communication on investor flows. I examine investors in the S&P 500 market portfolio, intermediated by mutual fund managers. I study the effect of communications about unrealized risks transmitted by mutual fund managers to these investors, including their interplay with investors' prior anxiety levels and trust shocks. Importantly, I control for and rule out various belief-updating channels.

3 Institutional Setting and Data

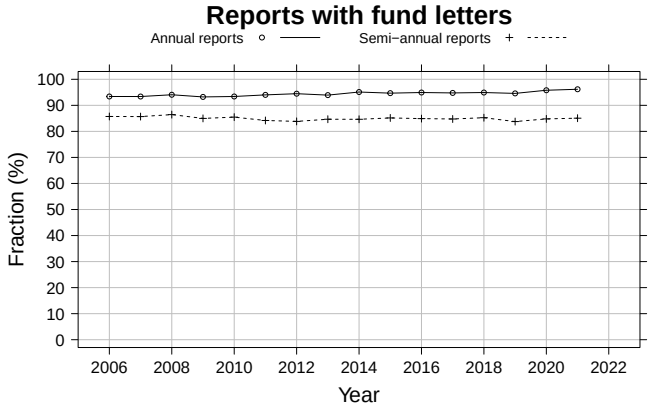
3.1 Letters to Investors Contained in Funds' Annual and Semi-Annual Reports

The SEC requires funds to disseminate annual reports and semi-annual reports to their investor clients, and these must also be filed within 10 days on the SEC's EDGAR system as Form N-CSR and N-CSRS, respectively. These reports contain information about historical performance,

fees and holdings. They also often contain free-form text communications to shareholders. Such communications do not follow a standardized format, are completely optional, and may contain unaudited statements. I call these statements "fund letters."[8]

Fund letters are often formatted as just that, beginning with phrases such as "Dear investors" and ending with a signoff. However, they may also appear in other guises; for example, under a section titled "Management discussion and analysis" or "Interview with portfolio manager." Extracting them requires casting a net wide enough to cover the many forms that these letters can take. My solution to this challenge is to define and then iteratively refine a set of regular expressions that identify the start and end locations of letters based on common phrases and section headings. I repeatedly sample from reports in which no letters were detected and manually read through the reports to select new delimiting phrases, then update the regular expressions and iterate once more. I also sample among extracted letters and cross-check them to the reports from which they were extracted to verify the accuracy of the results. Figure 1 indicates that my letter extraction methodology identifies letters in over nine out of ten annual reports and eight out of ten semi-annual reports in each year of my sample. Therefore, such free-form text is a very common form of communication from mutual fund managers to their investor clients.

Figure 1: Fraction of reports with letters. The vast majority of annual (N-CSR) and semi-annual (N-CSRS) reports published by fund families since 2006 contain at least one letter intended for investor clients. Fund letters within reports were detected using my text extraction methodology.



Both annual and semi-annual reports are published at a fund family level (for example, "Schwab Capital Trust") rather than at the level of an individual fund (for example, "Schwab S&P 500 Index Fund"). This presents a further challenge for matching letters to individual funds, but also the opportunity to use contemporaneous sibling funds' letter contents as part

[8]The literature also uses the more general phrase "shareholder letter" (Hillert, Niessen-Ruenzi, and Ruenzi, 2021) to connote that letter contents are not matched to individual funds. I use "fund letter" to emphasize that the letters I extract are targeted at the investors in an individual fund.

of an identification strategy.[9] To match letters in the fund family-level report to individual funds, I search for fund names or tickers in the contents of each letter. Some fund families only contain a single fund; in that case, matching is trivial but my identification strategy cannot be employed.

My sample begins in 2006, which is the first year individual fund letters can be identified; prior to that, the SEC EDGAR database did not record the mappings from funds to their families that are needed for my matching procedure. The sample ends in August 2021.

3.2 Measuring Letter Contents

I contribute a methodology to precisely measure the forward-looking contents of fund letters. My approach is capable of producing fine-grained measures of multiple types of forward-looking statements contained in fund letters; for example, the number of words, or fraction of a letter, devoted to forward-looking statements about risk or volatility, or other topics.

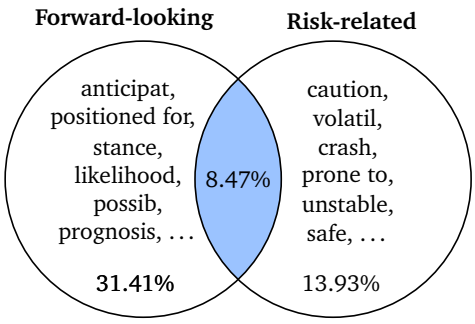
Procedure to Extract Statements from Fund Letters I provide a brief summary of the procedure: I extract grammatically-valid sentences from fund letters using Stanford CoreNLP tools (Manning et al., 2014) and discard boilerplate legal statements. I use a set of regular expressions (refined through an iterative procedure) to tag sentences containing forward-looking statements. I use an additional set of regular expressions to classify forward-looking statements into various categories. I can then count the number of words in each letter tagged as a particular category of forward-looking statement, and use this to characterize the forward-looking signals conveyed by funds to their investor clients. Using this procedure I find that statements about unrealized risks are prevalent in the fund letters corpus, with 8.47% of all sentences relating to future risks.

Figure 2 summarizes my procedure. All the regular expressions used are detailed in Internet Appendix E, with the twin goals of (i) encouraging replication & extension, while also (ii) highlighting the interpretability of the procedure in comparison to less transparent approaches.

Validation of Textual Analysis Procedure In Internet Appendix A.1, I compare my methodology to other approaches that are adapted to different contexts, or that use off-the-shelf software. I find that my customized approach detects a comprehensive set of forward-looking and risk-related statements, while also being transparent and adapted to the fund letters context. Quantitatively, my methodology meets or exceeds the fraction of forward-looking or risk-related statements detected by state-of-the-art software used in the Psychology literature

[9] Both identification strategies used in this paper are discussed in detail in Section 4.2.

Figure 2: Forward-looking risk-related statements. Forward-looking risk-related statements within fund letters are defined as sentences that are matched by both forward-looking and risk-related regular expressions. The percentage figures denote the fraction of all sentences in the fund letter corpus that are matched by each type of regular expression.



(which by contrasts operates as a proprietary "black-box" that gives no details of the criteria it uses).

Measuring Sentiment I also calculate the net sentiment of extracted text, based on word counts, as

$$\text{Net sentiment} = \frac{\text{Positive words} - \text{Negative words}}{\text{Total words}} \tag{6}$$

Individual words are classified as having positive or negative (or neutral) sentiment using Loughran and McDonald's (2011) financial sentiment dictionary, with an adjustment to correct for negations in the 3 words preceding positive terms.

Timeseries Variation & Determinants Even within a fund, there is considerable variation in both the level of detail conveyed, as well as the sentiment of these communications about future risks. The mean (median) within-fund autocorrelation of the fraction of the letter on this topic is only 0.12 (0.09). The mean (median) within-fund autocorrelation of the number of words devoted to this topic is similar, at 0.13 (0.19). The mean (median) within-fund autocorrelation of the net sentiment conveyed by these words is only 0.02 (0.08). As I show next, market conditions appear to be an important determinant of this timeseries variation.

Table 1 analyzes the timeseries means across funds of two measures of detail, as well as the sentiment measure. All measures are an equally-weighted cross-sectional year-month t mean of the within-fund-standardized (i.e. z-scores) equivalents; each measure is thus an average index of within-fund variation. The primary driver of all three timeseries is the contemporaneous level of the VIX. Within-fund variation thus appears to be honest, in the sense

Table 1: Timeseries indices of communicated detail and sentiment. The independent variables are measures of the level of detail in fund letters devoted to future risks (columns 1-2), and a measure of the sentiment of the same textual communication (column 3). Each independent variable y_t is the cross sectional mean over individual fund i variables, $y_t = \frac{1}{N_t}\sum_i z_{i,t}$. These individual fund variables $z_{i,t}$ are pre-standardized to focus purely within-fund variation, $z_{i,t} = (x_{i,t} - \overline{x_i})/\sigma(x_i)$. As well as the (log) VIX and prior month's return of the S&P 500 index, the independent variables comprise a measure of aggregate expected returns (proxied by the AAII survey employed by Greenwood and Shleifer (2014)), and a measure of aggregate perceived ambiguity of the S&P 500 market index (measured by Brenner and Izhakian (2018)). The sample of funds used to construct the dependent variables consists of S&P 500 index funds only.

	Detail		
Dependent Variables:	Number of Words$_t$	Fraction of Letter$_t$	Sentiment$_t$
	(1)	(2)	(3)
Log VIX$_t$	0.2719**	0.2332**	-0.3205***
	(0.1191)	(0.1151)	(0.1205)
Log Ambiguity$_t$	0.0387	0.0115	0.1327**
	(0.0556)	(0.0596)	(0.0603)
AAII Expected Return$_t$ (%)	-0.0002	0.0003	-0.0060
	(0.0022)	(0.0024)	(0.0040)
S&P 500 Return$_{t-1}$ (%)	0.0087	0.0040	-0.0029
	(0.0060)	(0.0063)	(0.0079)
Intercept	✓	✓	✓
N	183	183	183
R^2	0.05	0.04	0.14

Newey-West standard errors (in parentheses) use automatically-selected lags.
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

that managers provide more detail about future risks at precisely the moment when the VIX is highest. Managers also write with a more pessimistic tone (i.e. decreased sentiment) about future risks when the VIX is high – again showing that they are reflecting an objective measure of risk in their communications.

Cross-Sectional Dispersion Figures 3 and 4 summarize the cross-sectional and dispersion in the level of detail related to future risks (by the log word count) and the net sentiment of that content, respectively. Both variables exhibit considerable dispersion in the cross-section. Sentiment is on average negative (as might be expected in a discussion of risks) but both

net-positive and net-negative sentiment letters coexist at any given point in time.

Figure 3: Dispersion in detail. Quantiles and the mean of the (log) word count of the forward-looking risk-related portions of S&P 500 index fund letters.

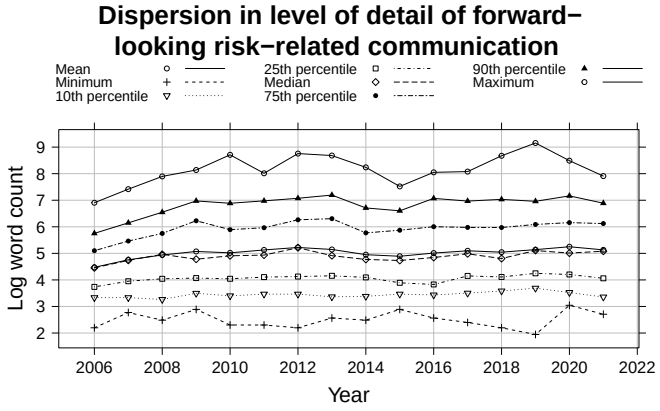
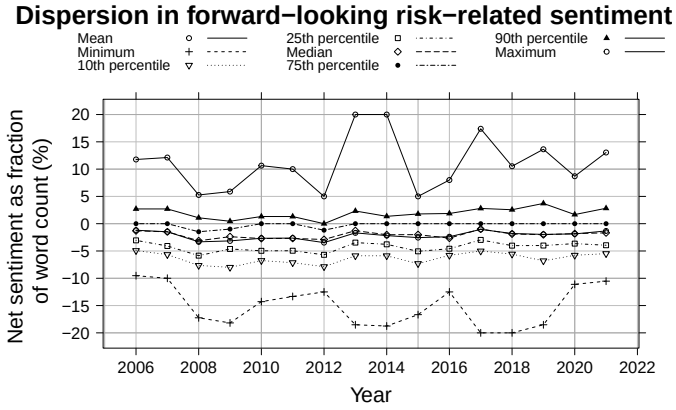


Figure 4: Dispersion in sentiment. Quantiles and the mean of the net sentiment contained in the forward-looking risk-related portions of S&P 500 index fund letters. The Loughran and McDonald (2011) sentiment dictionary is used as part of the net sentiment calculation.



3.3 Mutual Fund Classifications, Characteristics and Fees

I use the CRSP US Survivor-Bias-Free Mutual Fund Database (MFDB) to compute the mutual fund and fund family characteristics used in this study. These characteristics include classifications of funds as domestic equity funds, or otherwise.

I clean fee data and calculate total fund fees by closely following the procedure of Roussanov, Ruan, and Wei (2021, Internet Appendix). Total fees are therefore defined as the sum of expense ratios (which include management fees, and other expenses) and front loads (annualized by dividing by 7).

Funds are defined based on EDGAR identifiers; namely, CIKs and Series IDs. CRSP fund characteristics are matched to SEC identifiers using the "CRSP_cik_map" table from the CRSP Mutual Fund Database. Share class-level characteristics are aggregated to the fund level by total net assets (TNA).

I obtain fund-level monthly net returns from the CRSP MFDB using a similar procedure. These are then winsorized at the 5% and 95% levels over the entire sample to avoid the undue influence of outliers.

One of my identification strategies relies on measuring communication by active funds; to identify these, I first classify index funds (in general) based on their names, following the list of keywords proposed by Ben-David, Li, Rossi, and Song (2022). Active funds are then defined as non-index funds.

3.4 Fund Flows

Monthly net flows for mutual funds are calculated following the extant literature, based on the fund returns and Total Net Assets (TNA) recorded in the CRSP US Survivor-Bias-Free Mutual Fund Database, as

$$\text{Net flow}_{i,t\to t+1}(\$) \;=\; \text{TNA}_{i,t+1} \;-\; (1+R_{i,t\to t+1}) \times \text{TNA}_{i,t}. \tag{7}$$

In addition to estimated net flows, I also make use of data on gross monthly inflows and outflows reported to the SEC at a fund level. These gross dollar flows are extracted from mandatory SEC filing documents (N-SAR, N-PORT, and any amendments), with the latest available observation retained per month.

All dollar flows are normalized to a percentage of lagged assets:

$$\text{Net flow}_{i,t\to t+1}(\%) \;=\; \frac{\text{Net flow}_{i,t\to t+1}(\$)}{\text{TNA}_{i,t}} \times 100\%. \tag{8}$$

A similar normalization to Eqn. (8) is applied to gross inflows and gross outflows.

To avoid the undue influence of outliers, including those due to corporate actions, the net flow percentage variable is winsorized at the 5% and 95% levels over the full sample. Similarly, each gross flow percentage variable is winsorized at the 95% level.

3.5 S&P 500 Index Funds

S&P 500 index funds may be either mutual funds or exchange traded funds (ETFs). This study focusses on index mutual funds, which are more prevalent. To select S&P 500 index mutual funds, I source monthly fund return information from the CRSP US Survivor-Bias-Free Mutual

Fund Database: I aggregate returns from the share class level to the fund level, then correlate these returns with the returns on the S&P 500 index and discard funds with a correlation less than 98%. Finally, I filter out funds whose names suggest they are not following a pure passive indexing strategy (such as quant funds, funds tilted to a particular sector, or funds that track a different index). This procedure selects a total of 211 S&P 500 index mutual funds.[10]

3.6 SEC EDGAR Usage

I measure the geographic locations of fund letter readers using SEC EDGAR usage logs, beginning in 2006 and ending in June 2017. I follow Lee, Ma, and Wang (2015) in identifying and removing robots from the usage logs, and de-anonymize IP addresses using the lookup table provided by Chen et al. (2020). Similarly to Grice and Gueciouer (2023), I then geolocate IP addresses using historical mapping tables provided by the MaxMind geolocation service. Internet Appendix A.2 validates my sample of fund readership data by comparing the geographic distribution of fund letter readers to the overall population.

4 Fund Manager Communication and Investor Risk-taking

This section shows that more detailed fund manager communication encourages investor risk-taking. I first quantify the effects of detail and sentiment on individual S&P 500 tracker fund flows. I show that sentiment does not, in fact, predict fund flows, which is initial evidence against a belief-updating channel; in Section 6 I rule out such a channel more strongly. Section 4.2 formulates two (very different) identification strategies that both confirm the presence of a causal effect.

4.1 Communication and Fund Flows

I begin by analyzing the relationship between net flows into S&P 500 index funds and the amount of forward-looking risk-related information transmitted by managers to investors, and find that they are positively associated, irrespective of whether the conveyed sentiment is positive or negative.

Level of Detail My main covariate of interest is the (log) word count of forward-looking risk-related information contained in an S&P 500 index tracker fund letter. The panel that

[10] By way of comparison, I repeated the procedure on the ETFs recorded in the CRSP monthly stock database, and found only 12 S&P 500 index ETFs. Focussing on mutual funds thus captures the vast majority of S&P 500 tracker funds.

Table 2: Detail communicated about future risks vs. net flows into S&P 500 index funds. The main covariate measures the level of detail contained in the letter written by fund i at time t about forward-looking risks, using the log of the number of words devoted to that topic. Net flows are calculated using the CRSP Mutual Fund Database, as in Eqn. (8). Columns 1-4 present results with no variable transformations. Columns 5-8 present results where both the dependent and the main independent variables have been standardized (z-scores). *Fund controls* (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. *Market controls* (not shown) are for the (log) VIX level and for aggregate expected returns (proxied by the AAII survey employed by Greenwood and Shleifer (2014)); note these are not compatible with time effects. The sample consists of S&P 500 index funds only.

Dependent Variables:	Net Flow$_{i,t\to t+1}$ (%)				z(Net Flow$_{i,t\to t+1}$)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Detail$_{i,t}$	0.3846***	0.4746***	0.4893***	0.5078***				
	(0.1242)	(0.1528)	(0.1528)	(0.1801)				
z(Detail$_{i,t}$)					0.1741***	0.2148***	0.2215***	0.2299***
					(0.0562)	(0.0692)	(0.0692)	(0.0815)
Fund Controls$_{i,t}$	✓	✓	✓	✓	✓	✓	✓	✓
Market Controls$_t$			✓				✓	
Year-month FEs	✓			✓	✓			✓
Fund FEs		✓	✓	✓		✓	✓	✓
N	1,155	1,155	1,155	1,155	1,155	1,155	1,155	1,155
R^2	0.20	0.31	0.31	0.44	0.20	0.31	0.31	0.44

*Clustered (Fund & Year-month) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 3: Detail communicated about future risks vs. gross flows to/from S&P 500 index funds. The main covariate measures the level of detail contained in the letter written by fund i at time t about forward-looking risks, using the log of the number of words devoted to that topic. Gross dollar flows are extracted from mandatory SEC filing documents (N-SAR, N-PORT), and then normalized as a percentage of lagged total net assets. Columns 1-4 focus on the association with gross inflows. Columns 5-8 focus on the association with gross outflows. *Fund controls* (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. *Market controls* (not shown) are for the (log) VIX level and for aggregate expected returns (proxied by the AAII survey employed by Greenwood and Shleifer (2014)); note these are not compatible with time effects. The sample consists of S&P 500 index funds only.

Dependent Variables:	$\text{Inflow}_{i,t\to t+1}$ (%)				$\text{Outflow}_{i,t\to t+1}$ (%)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(9)
$\text{Detail}_{i,t}$	0.2124	0.4950***	0.4823***	0.4777**	-0.0866	0.0236	0.0037	0.0524
	(0.2361)	(0.1756)	(0.1734)	(0.2008)	(0.1768)	(0.0825)	(0.0797)	(0.0983)
$\text{Fund Controls}_{i,t}$	✓	✓	✓	✓	✓	✓	✓	✓
$\text{Market Controls}_{t}$			✓				✓	
Year-month FEs	✓			✓	✓			✓
Fund FEs		✓	✓	✓		✓	✓	✓
N	1,155	1,155	1,155	1,155	1,155	1,155	1,155	1,155
R^2	0.25	0.74	0.74	0.79	0.24	0.83	0.83	0.85

*Clustered (Fund & Year-month) standard-errors in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

I analyze consists of fund-month pairs for each published letter; letters are published at a semi-annual frequency (approximately).

Table 2 columns 1-4 present the results of panel regressions of the net flows of S&P 500 funds regressed against the level of detail relating to forward-looking risks contained in fund letters disseminated the previous month. Importantly, controlling for year-month time effects in columns (1) and (4) implicitly controls for any time-varying potential confounders; this includes past or future market returns and measures of risk or risk premia, market-wide sentiment, any (common) prior held by investors, the growth of passive investing – and, indeed, any other market-wide and economy-wide events or shocks. All OLS coefficients in columns 1-4 are positive and significantly different to zero, with the coefficient magnitudes increasing as more controls are incorporated. Furthermore, the positive and significant association survives a number of controls for variables that have been previously found to drive fund flows (Sirri and Tufano, 1998; Chevalier and Ellison, 1999) including the fund's past returns, fees and size, the family's age and size, as well as fund fixed effects and year-month effects. I also control for the (log of the) total number of words in a letter, which effectively normalizes by letter length.

To illustrate that the effect is not only cross-sectional (i.e. present when time effects are included) but also within-fund, the specifications in Columns 2-3 do not include time effects at all. In column (3) I explicitly control for investors' beliefs about the distribution of the S&P 500's future return: I use survey expectations (Greenwood and Shleifer, 2014) and the (log) level of the VIX for the first and second moment, respectively. In column (4), I add both fund fixed effects and time effects and find the magnitude strengthens.

Economic Significance The effect is economically significant. This is most clear when regressing standardized (i.e. z-scored) net flows against standardized detail measures in Table 2 columns 5-8. In column (8), the richest specification, a growth of 1 standard deviation in the level of detail above its average predicts a next-month increase in the fund's assets of 0.23 standard deviations in magnitude.

Un-standardized estimates also convey the magnitude of the effect. In column (4), the richest specification, the estimated coefficient value of +0.51 on the log word count suggests that a fund that doubles the detail it communicates about future risks attracts an additional $0.51 \times \log(1+1) = 0.35\%$ worth of assets under management in the month following its dissemination of a fund letter. The economic magnitude can alternatively be assessed based on the alternative specification in Internet Appendix D, or by assessing an aggregate effect in Section 7, and remains notable when taking those approaches.

Unobservables The stability of the estimated coefficients in Table 2 is informative about what effect any unobservables might have. These results show that including additional controls, especially fund-level fixed effects, does not change the main coefficient estimates in any qualitative or economically-meaningful way, but does increase the R^2: comparing columns (1) to (4) and columns (5) to (8), we see that including fund fixed effects more than doubles the R^2. Unobservables are hence unlikely to explain the main effect, reducing potential concerns about selection bias and omitted variable bias (Altonji, Elder, and Taber, 2005; Oster, 2019). This finding further justifies the present paper's focus on S&P 500 index funds to aid in identification. Section 4.2.1 elaborates further by formally applying the Oster (2019) test.

Decomposition into Inflows & Outflows I now decompose net flows into gross inflows and gross outflows. Table 3 repeats the baseline regression results on each gross flow variable, including the same fixed/time effects and controls as previously. These estimates show that the effect acts through increasing inflows, rather than decreasing outflows: the estimates for the effect on inflows in columns 1-4 are positive and significant; by contrast, the estimates for the effect on outflows in columns 5-8 are not significantly different to zero. Since fund letters are targeted at existing investors rather than potential clients (who are attracted by fund prospectuses or advertising), this positive coefficient can be interpreted as existing investors increasing their share of the risky asset.

Role of Sentiment (or Lack Thereof) Given that the sentiment conveyed in fund letters about future risks is honest and accurate, in the sense that it negatively covaries with the current level of the VIX, I now analyze how this sentiment relates to the baseline detail effect. Table 4 augments the baseline regressions of Table 2 and now includes the net sentiment of the forward-looking risk-related communications within each letter (columns 2-3), or of the entire letter content (columns 4-5), with the addition of an interaction term.

The sentiment coefficients in the second row of Table 4 are not significantly different to zero. Similarly, the interaction terms in the third row are also insignificant. These results obtain no matter whether the sentiment measure is computed solely on forward-looking risk-related statement (columns 2-3) or on the overall letter (columns 4-5). By contrast, the coefficients on the level of detail communicated in the first row are positive, significant, and of the same order of magnitude as before. Taken together, these results suggest that the primary drivers of fund flows are the level of detail communicated *irrespective* of the sentiment conveyed. Further analyses in Section 6.2 show that investors do not appear to update their beliefs about the level of future risks based on the sentiment conveyed in these fund letters.

Table 4: Interaction between communication detail & sentiment, and the relation with S&P 500 index fund flows. This table shows the relationship between net flows, communication detail, the sentiment of the communication, and their interaction. The level of detail is measured as the log of the number of words devoted to forward-looking risk-related statements. The sentiment measure in columns 2-3 is calculated on forward-looking risk-related statements. The sentiment measure in columns 4-5 is calculated on the overall fund letter. Controls (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. The sample consists of S&P 500 index funds only. Columns 1-2 show the effect for net flows.

Dependent Variable:	Net Flow$_{i,t\to t+1}$ (%)				
	(1)	(2)	(3)	(4)	(5)
Detail$_{i,t}$	0.5109***	0.5078***	0.4837***	0.5103***	0.4505**
	(0.1795)	(0.1771)	(0.1844)	(0.1796)	(0.1820)
Sentiment$_{i,t}$		-0.0243	0.0194	0.0046	0.1051
		(0.0281)	(0.0754)	(0.0200)	(0.0888)
Detail$_{i,t}$ × Sentiment$_{i,t}$			-0.0098		-0.0239
			(0.0154)		(0.0193)
Sentiment Calculation		Future Risk-Related		Overall Letter	
Fund Controls$_{i,t}$	✓	✓	✓	✓	✓
Year-month FEs	✓	✓	✓	✓	✓
Fund FEs	✓	✓	✓	✓	✓
N	1,185	1,185	1,185	1,167	1,167
R^2	0.44	0.44	0.44	0.44	0.44

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

4.2 Identification

The fundamental pillar of this paper's ability to identify the effect of communication is my focus on quasi-identical S&P 500 index funds — a setting where many channels such as learning about or signaling/disclosing skill do not apply. This section verifies that the previous estimates are not due to omitted variables or other forms of selection bias. To do so, I use three radically different (hence complementary) methodologies to test for endogeneity.

The first relies on partial identification results (Altonji, Elder, and Taber, 2005; Oster, 2019). The second relies on an exclusion restriction, by employing an instrumental variables approach. The third relies on a shape restriction, by exploiting a discontinuity induced by corner bunching (Caetano, Caetano, and Nielsen, forthcoming; Caetano, Caetano, Nielsen, and Sanfelice, 2023). Despite the very different assumptions underlying these methods, the

three analyses consistently confirm the presence of a causal effect.

4.2.1 Ruling Out Unobservable Drivers

Building on the seminal work of Altonji, Elder, and Taber (2005), Oster (2019) shows how to test for the presence of unobservable selection using empirical measures of coefficient stability. I now apply her test to confirm my baseline results are not driven by further unobservable omitted variables. Reassuringly, the test implies unobservables do not drive the findings.

The test is based on calculating a coefficient of proportionality $\tilde{\delta}$ that measures the proportional degree of selection on unobservables relative to observables that would be needed to nullify the estimated treatment effect. The calculation of $\tilde{\delta}$ itself involves a sensitivity analysis of the baseline regression, and is described in detail by Oster (2019). I adopt her suggested assumption for the hypothetical share of variance that could be explained as $\tilde{R}^2 = \min\{1.3\hat{R}^2, 1\}$. I set $\hat{R}^2$ to that of Table 2 column (4), which is my richest specification, and which has the most explanatory power. Given the above parameterization, Oster (2019) argues a coefficient of proportionality with value $\tilde{\delta} > 1$ signifies the effect is robust to the influence of unobservables.

The resulting degree of proportionality is calculated to be $\tilde{\delta} = 1.11$. Since this is strictly greater than 1, it provides reassuring evidence that unobservable drivers are not responsible for the baseline effect.

4.2.2 Instrumental Variables Identification Strategy

As a further check, I now formulate an instrumental variables (IV)-based identification strategy that shifts the explanatory variable (the level of detail communicated). The strategy is facilitated by this paper's focus on passive funds.

The key quantity of interest in my analyses is the (log of the) number of words within an S&P 500 fund letter devoted to forward-looking risk-related statements. To construct an instrument for this quantity, I calculate similar measures for the contemporaneous letters of other funds within the same fund family, then take the family-level mean over all active domestic equity funds. This excludes any passive index funds, including the focal S&P 500 fund of interest.

To understand the information spillover mechanism captured by this instrument, recall first that the market portfolio return, in which S&P 500 index fund investors are invested in, is the aggregated return of all constituent firms or industries i:

$$R_{t+1} = \frac{1}{\sum_i \text{Size}_{i,t}} \sum_i \text{Size}_{i,t} R_{i,t+1} \tag{9}$$

In the ordinary course of business, an actively-managed fund uncovers a series of signals ε_i

Table 5: Instrumented effect of the level of detail communicated about future risks on the net flows of S&P 500 index funds. The main covariate measures the level of detail contained in the letter written by fund i at time t about forward-looking risks, using the log of the number of words devoted to that topic. The instrument for the main covariate used in the 2SLS specifications is constructed from the log word counts (about the same topic of forward-looking risk) of the letters written by active domestic equity sibling funds within the focal index fund's overall family. *Fund controls* (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. *Market controls* (not shown) are for the (log) VIX level and for aggregate expected returns (proxied by the AAII survey employed by Greenwood and Shleifer (2014)); note these are not compatible with time effects. The sample consists of S&P 500 index funds only. Columns 1-3 show un-instrumented panel specifications. Columns 4-6 show 2SLS specifications, with F-stats corresponding to the first stage; the first-stage regressions are shown in Table IA.4.

Dependent Variable:	Net Flow$_{i,t\to t+1}$ (%)							
Regression:	OLS				2SLS IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Detail$_{i,t}$	0.3846***	0.4746***	0.4893***	0.5078***	0.4806***	0.6088***	0.6252***	0.6527***
	(0.1242)	(0.1528)	(0.1528)	(0.1801)	(0.1317)	(0.1776)	(0.1795)	(0.2038)
Fund Controls$_{i,t}$	✓	✓	✓	✓	✓	✓	✓	✓
Market Controls$_t$			✓				✓	
Year-month FEs	✓			✓	✓			✓
Fund FEs		✓	✓	✓		✓	✓	✓
N	1,155	1,155	1,155	1,155	1,155	1,155	1,155	1,155
R^2	0.20	0.31	0.31	0.44	0.20	0.31	0.31	0.44
F-stat (IV 1st-stage)					289.5	335.2	261.9	412.7

Clustered (Fund & Year-month) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

about some subset of individual returns $R_{i,t+1}$ as part of its portfolio choice process. These signals ε_i about individual firms or industries i can be aggregated up to a signal ε about the market portfolio return R_{t+1}. I do not observe the internal research of active funds within the family (i.e. the ε_i realizations), but can observe the number of the level of detail about future risks that these active funds communicate to their own clients, via their own fund letters. $\bar{w}_A = \frac{1}{N_k} \sum_{k \in \text{Active}} w_k$, the mean over all these variables w_k, captures the level of detail about future risks that emerges from active funds' internal research. This instrument captures information diffusion among funds in the same family, which drives variation in the content of the focal S&P 500 index fund's letter.[11]

Two assumptions must hold for this instrument to be valid. First, there must be no reverse information spillovers from passive funds to active funds. Since passive funds do not conduct fundamental research of their own, this is unlikely.[12] Second, any information about future risks contained in active funds' letters should not affect flows to passive funds (other than via information spillovers to passive funds' letters). The analysis that follows in Section 6.1 shows that more detailed communication about future risks in sibling active funds' letters does *not* trigger outflows from these active funds themselves, which is the most likely source of rebalancing inflows to the passive funds. Hence, such a potential violation is also unlikely.

My IV-based identification strategy produces similar results to the paper's baseline regressions. Table 5 columns 1-4 repeats these baseline results (from Table 2), and then columns 5-8 show two-stage least squares (2SLS) estimates that have been instrumented using the sibling domestic active equity fund communication detail instrument. The positive and significant effect remains, and can now be interpreted as the causal effect of increased communication about future risks related to the risky portfolio, and which leads to increased flows into the risky portfolio managed by the fund.

Further details of the IV identification strategy – including evidence of information spillovers within fund families – are provided in Internet Appendix B.1.

[11] I do *not* assume that funds share fundamental research or trade ideas with one another, although there is evidence of this in prior work (Cici, Jaspersen, and Kempf, 2017; Auh and Bai, 2020). I assume only that the market-wide views expressed in fund letters by passive funds incorporate information produced by active funds within the same fund family.

[12] It might be argued that passive funds still conduct fundamental research as part of their shareholder duties (voting and monitoring), but the literature finds this is not the case. Iliev, Kalodimos, and Lowry (2021) measure research activity at the fund family level using SEC EDGAR usage, and find that fund families with a larger share of passive assets conduct less research. Heath, Macciocchi, Michaely, and Ringgenberg (2022) find that index funds are poor monitors as a result. Moreover, while some index funds are closet active funds, these tend to use custom indices (Akey, Robertson, and Simutin, 2021); the present paper focuses on S&P 500 index funds.

4.2.3 Corner Bunching Identification Strategy

My third and final analysis exploits corner bunching in the detail treatment variable for identification. Recall that the baseline specifications in the paper so far normalize for the total (log) word count of the entire fund letter by including that variable as a control. I now construct a related measure of the level of detail: the fraction of a fund letter devoted to forward-looking risk-related statements, without taking logs of either the numerator or denominator.

Figure 5: Corner bunching used in identification. These figures visualize the distribution of the fractions of S&P 500 index fund letters that are devoted to forward-looking risk-related statements in this paper's sample. The left chart shows the empirical CDF of this variable, and the right chart shows its histogram. The histogram highlights corner bunching at the value 0%: this is shown by the single bar colored red, while all other bars are colored blue. A corresponding sharp discontinuity is visible in the empirical CDF at that value. The presence of this feature of the empirical distribution is used in the corner bunching identification strategy described in Section 4.2.3.

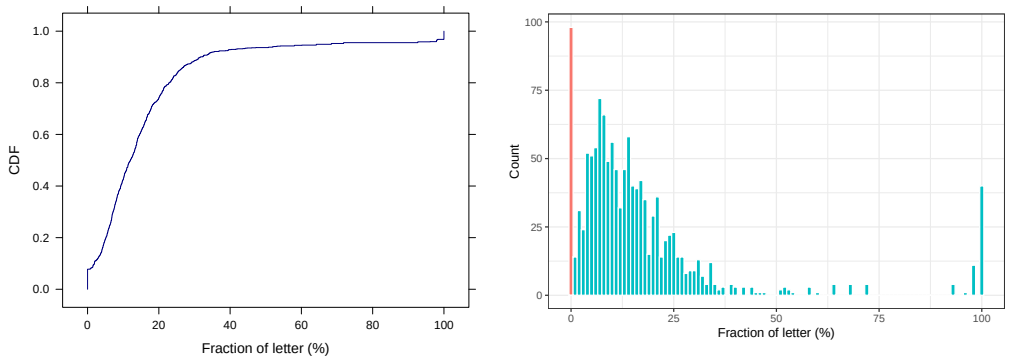


Figure 5 highlights the presence of corner bunching at the value 0%. Recent methodological contributions by Caetano, Caetano, and Nielsen (forthcoming) have been developed to exploit precisely such a feature of the distribution for identification purposes. The key intuition is to use shape restrictions on the CDF to detect the presence of any potential endogeneity and then correct for it. Further details of the methodology are described in Internet Appendix B.2. I now leverage their approach.

Table 6 repeats this paper's main baseline regression of net flows on the recent level of detail about future risks in a fund letter, with these coefficients labeled β. Specification (1) conducts an OLS/panel regression: here, β is positive and significantly different to zero. The usual additive year-month and fund fixed effects are included, together with fund controls.

The remaining specifications in Table 6 apply Caetano, Caetano, and Nielsen's (forthcoming) correction. β remains the main coefficient of interest, while an additional δ coefficient is estimated. δ is the coefficient on the endogeneity correction term of Caetano, Caetano, and Nielsen's (forthcoming) model; the model is parameterised using a Tobit in specification

Table 6: Corner bunching-based identification of the effect of the level of detail communicated about future risks on the net flows of S&P 500 funds. β is the coefficient capturing the corrected effect of the level of detail contained in the letter written by fund i at time t about forward-looking risks, using the fraction of the letter devoted to that topic. δ is the coefficient on the endogeneity correction control. Fund controls (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. Standard errors are based on 1,000 bootstrap replicates. The sample consists of S&P 500 index funds only.

Dependent Variable:	Net Flow$_{i,t\to t+1}$ (%)		
Specification:	OLS	Tobit-corrected	Semiparam. Tobit
	(1)	(2)	(3)
β	0.028***	0.069**	0.099**
	(0.008)	(0.035)	(0.045)
δ		-0.037	-0.067
		(0.031)	(0.042)
Fund Controls$_{i,t}$	✓	✓	✓
Year-month FEs	✓	✓	✓
Fund FEs	✓	✓	✓
N	1,259	1,259	1,259

Bootstrapped standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

(2) and a semiparametric tobit in specification (3) to apply the correction. In each case β remains positive and significant. Importantly, δ is not significantly different to zero in either specification.

The absence of detectable endogeneity provides further support for the causality of the communication effect. Furthermore, the conclusion using the present approach is consistent with the findings from the previous IV-based identification strategy.

5 Inspecting the Mechanism

This section provides evidence that my findings on communication-induced inflows are consistent with the money doctors hypothesis.

5.1 Detailed Communication Alleviates Anxiety

The money doctors hypothesis by GSV posits that financial intermediaries can alleviate investors' anxiety. This section tests whether the anxiety attitudes of fund letter readers impact their response to the provision of information about future risks by fund managers.

To test for investors' anxiety playing a role, I examine the differential effect of communication on high-anxiety versus low-anxiety readerships. My measure for the *ex ante* level of anxiety experienced by a letters' readers is constructed from Google Trends data. Google's search engine groups together related search terms under umbrella "topics," and defines two anxiety-related topics:[13]

1. The first anxiety topic, which Google labels "Psychological Stress", includes the search terms "stressed", "what is stress", "anxiety" & "relieve stress."

2. The second anxiety topic, which Google labels "Worry", includes the search terms "worry about", "anxiety", "stop worrying" & "bible worry."

For each topic's geographic state s-month t SVI panel, I aggregate across fund letter readers to construct a fund i-month t panel:

$$\text{Readership SVI}_{i,t} = \frac{\sum_j \text{Investment}_{j,t} \times \text{Geographic SVI}_{j,t}}{\sum_j \text{Investment}_{j,t}}. \tag{10}$$

I then repeat the detail-sentiment-flow panel regressions (of Table 4), this time cross-sectionally splitting the sample based on letter readers' anxiety attitudes. That is, at each year-month t, I split fund letter readerships into those with below-median and above-median anxiety attitudes.

Table 7 presents the results of such panel regressions, and shows evidence that readers' anxiety attitudes are crucial determinants of their response to information provision about future risks. Low-anxiety readers do not respond to more detailed communication in a statistically-significant manner (columns 1-4). By contrast, high-anxiety readers respond strongly to information provision (columns 5-8).

Altogether, these results support a mechanism in which detailed prognoses by fund managers alleviate investor anxiety.

[13] Access to Google's restricted Trends API is not necessary to examine these topics or replicate this paper's results, since the publicly-accessible Google Trends website also allows SVI data to be downloaded. At the time of writing (August 2023), the "Psychological Stress" topic SVI was accessible at `https://trends.google.com/trends/explore?q=/m/0glpkn8&date=all&geo=US`. Similarly, the "Worry" topic SVI was accessible at `https://trends.google.com/trends/explore?q=/m/0bg_x1&date=all&geo=US`. However, these identifiers may change in the future.

Table 7: Readership anxiety levels, and the differential effects of the level of detail & sentiment communicated about future risks on S&P 500 index fund flows. The main covariate measures the level of detail contained in the letter written by fund i at time t about forward-looking risks, using the log of the number of words devoted to that topic. The sentiment measure is calculated on forward-looking risk-related statements. The sample consists of S&P 500 index funds only, and is split cross-sectionally (i.e. at every year-month t) into fund letter readerships with below-median anxiety levels (columns 1-4), and above median anxiety levels (columns 5-8). In measuring readership anxiety levels, two Google Trends topics are used: Psychological Stress (columns 1-2 & 5-6), and Worry (columns 3-4 & 7-8). Additional controls (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. Each flow variable is the next month's flow, in percentage points of TNA.

Dependent Variable:	Net Flow$_{i,t\to t+1}$ (%)							
Sub-sample:	Low Anxiety				High Anxiety			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Detail$_{i,t}$	0.1512	0.1504	0.0212	0.0505	0.6223**	0.4976**	0.9940***	0.8697***
	(0.1729)	(0.1692)	(0.1430)	(0.1663)	(0.2610)	(0.2499)	(0.2992)	(0.2840)
Sentiment$_{i,t}$		0.0284		-0.1199		0.0873		0.2167**
		(0.1024)		(0.1182)		(0.0960)		(0.0999)
Detail$_{i,t}$ × Sentiment$_{i,t}$		0.0012		0.0169		-0.0339		-0.0508**
		(0.0237)		(0.0294)		(0.0227)		(0.0230)
Google Trends Topic	Psychological Stress		Worry		Psychological Stress		Worry	
Fund Controls$_{i,t}$	✓	✓	✓	✓	✓	✓	✓	✓
Year-month FEs	✓	✓	✓	✓	✓	✓	✓	✓
Fund FEs	✓	✓	✓	✓	✓	✓	✓	✓
N	586	586	600	600	653	653	639	639
R^2	0.59	0.59	0.57	0.57	0.51	0.51	0.53	0.54

Clustered (Fund & Year-month) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

5.2 Detailed Communication Fosters Trust

In GSV's view, trust and anxiety are intertwined. Market participants also attribute a link between increased communication and trust-building.[14] In this section, I analyze whether communication builds up the trust between fund managers and investors, hand-in-hand with alleviating their anxiety.

Stocks of trust and perceived trustworthiness are challenging to measure at the level of a particular fund's investor base. In order to formulate a test of my trust-building hypothesis, I follow the literature in exploiting the occurrence of negative external shocks to trust (for example, Kostovetsky, 2016; Gurun, Stoffman, and Yonker, 2018). Specifically, I use a geographic measure of negative shocks to investors' trust in financial markets that has been validated by the prior literature: Giannetti and Wang (2016) show that the revelation of local corporate fraud acts as an exogenous shock to the trust held by nearby investors in the stock market in general.

The logic of my test is as follows. If more detailed communication builds trust, one should observe a negative interaction between communication detail and exposure to negative trust shocks. That is, if one variable increases trust and the other decreases it, the variables should have opposing effects, and this interaction should be significantly different to zero. Otherwise, the effects of communication and changes to trustworthiness are unrelated.

Like Giannetti and Wang (2016), I measure negative trust shocks using the public revelation of securities fraud incidents by listed companies, as compiled by Karpoff, Koester, Lee, and Martin (2017).Using the state in which each company was headquartered in, I construct a state-level measure of the local investor population's exposures to securities fraud incidents. Specifically, for an individual j, her exposure to securities fraud during a period t is measured based on the number of fraud incidents by companies headquartered in the state in which she is located, divided by the total number of public companies headquartered in the same state:

$$\text{Geographic Fraud Exposure}_{j,t} = \frac{\#\text{ Fraud incidents by in-state public firms}}{\#\text{ Total in-state public firms}} \times 100\%. \tag{11}$$

Given the geographic locations of fund letter readers that I measured from the SEC EDGAR logs, I aggregate from individuals' geographic fraud exposures to fund-level fraud exposures:

$$\text{Readership Fraud Exposure}_{i,t} = \frac{1}{\sum_j \text{Investment}_{j,t}} \sum_j \text{Investment}_{j,t} \times \text{Geographic Fraud Exposure}_{j,t}. \tag{12}$$

This aggregation is performed similarly to that in the previous section; i.e. at an annual frequency, and proxying for investment weights using SIPP survey values. The fund level readership fraud exposure captured by Equation (12) proxies the extent to which the readers

[14] A Financial Times survey of investment advisors in September 2012 highlighted that communication is one of the three top drivers of fund manager trustworthiness in the eyes of investment advisors.

of the letter issued by fund i at time t have been exposed to the revelation of securities fraud by nearby listed firms.

Table 8: Communication-driven flows and readership exposure to securities fraud. The level of detail contained in the letter written by fund i at time t about forward-looking risks is measured using the log of the number of words devoted to that topic. The indicator variable $\mathbb{1}\{\text{High Fraud Exposure}\}_{i,t}$ is 1 when the readership of the fund's letter currently has an above-median Fraud Exposure Index, and zero otherwise; a baseline effect is estimated in addition to the interaction shown in the table. The variable $\text{High Fraud Exposure}_{i,t}$ is rank-standardized, and thus runs from 0 to 1; a baseline effect is estimated in addition to the interaction shown in the table. The Fraud Exposure Index is a weighted average (by unique report downloads) of revelations of securities fraud (recorded by Karpoff, Koester, Lee, and Martin (2017)) committed by publicly listed companies headquartered in the local areas of each fund i's annual/semi-annual report readers, varying annually. The overall sample period begins in 2006 (due to EDGAR data availability) and ends in 2012 (with the securities fraud database). *Fund controls* (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. *Market controls* (not shown) are for the (log) VIX level and for aggregate expected returns (proxied by the AAII survey employed by Greenwood and Shleifer (2014)); note these are not compatible with time effects. The sample consists of S&P 500 index funds only.

Dependent Variable:	$\text{Net Flow}_{i,t\to t+1}$ (%)	
	(1)	(2)
$\text{Detail}_{i,t}$	0.6975*	0.9254*
	(0.4146)	(0.4839)
$\text{Detail}_{i,t} \times \mathbb{1}\{\text{High Fraud Exposure}\}_{i,t}$	-0.4727*	
	(0.2662)	
$\text{Detail}_{i,t} \times \text{Fraud Exposure}_{i,t}$		-0.8701*
		(0.4448)
$\text{Fund Controls}_{i,t}$	✓	✓
Year-month FEs	✓	✓
Fund FEs	✓	✓
N	440	440
R^2	0.54	0.54

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 8 shows the results of repeating this paper's baseline regression specification, with an additional interaction for local securities fraud exposure by the readership of each fund letter. The specification in column (1) employs an indicator variable for above-median exposure, and the specification in column (2) instead uses the rank-scaled exposure. Baseline effects are

also calculated, though not shown. In both specifications, there is a negative and statistically significant interaction effect between the detail communicated about future risks in a fund letter, and the contemporaneous exposure of the letter's readers to local securities fraud. The smaller sample size (and lower statistical power of the test) are due to the shorter sample, as the current version of the Karpoff, Koester, Lee, and Martin (2017) database ends in 2012.

I conclude from these findings that communication builds trust; that is, trust-destroying fraud and trust-building communication interact, and act in opposing directions. The results are consistent with the GSV framework.

6 Evaluating Potential Alternative Mechanisms

I now consider a number of potential alternative explanations for my main findings. Going through the proposed mechanisms in turn, I conduct additional empirical tests that rule them out as potential explanations. Notably, the evidence is not consistent with belief-based persuasion models in general.

6.1 Can Inflows Be Explained by Rebalancings Into Passive Trackers?

One potential explanation for the positive effect of communication about future risks related to the S&P 500 on tracker fund inflows is that consuming communication about future risks shifts investors' perceptions about risk higher in general; similarly, investors' risk aversion may in fact be increased. If there is indeed such a risk perception or risk aversion spillover, investors should withdraw their holdings from funds with a riskier mandate and increase their allocations to lower-risk S&P 500 index funds.

An ideal test of this candidate explanation would exploit individual investor-level data, instead of relying on aggregate flow data; yet this is not an important limitation due to the structure of investors' switching costs. Mutual fund investors are likely to keep their investments within the same fund family due to high across-family switching costs (Massa, 2003), as compared to within-family switching costs. In other words, studying rebalancing flows within fund families focuses precisely on the most frictionless setting, in which such flows are most likely to appear – if the rebalancing hypothesis is indeed true.

I test the hypothesis according to the following logic: a necessary condition for this explanation to hold is that communication about future risks induces investors to withdraw their holdings from riskier funds. Therefore, communication about future risks in S&P 500 index fund letters should predict an increase in outflows from riskier active domestic equity funds.

The regressions in Table 9 columns 1-2 test this hypothesis: instead of being positive, the predictive coefficient is negative in magnitude and not statistically significantly different to

Table 9: Testing for communication-driven rebalancing flows out of active funds. This table shows the relationship between outflows from active domestic equity funds and the level of detail about forward-looking risk-related statements contained in the letters of their sibling S&P index tracker funds (columns 1-2) or in their own letters (columns 3-4). The level of detail is measured as the log of the number of words devoted to forward-looking risk-related statements. Controls (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. The sample consists of active domestic equity funds that belong to fund families that also contain an S&P 500 index fund.

Dependent Variable:	Sibling Outflow$_{i,t\to t+1}$ (%)			
	(1)	(2)	(3)	(4)
S&P 500 tracker Detail$_{i,t}$	-0.1667	-0.1220		
	(0.2125)	(0.0950)		
Sibling Detail$_{i,t}$			-0.1220	-0.1530
			(0.2204)	(0.0995)
Fund Controls$_{i,t}$	✓	✓	✓	✓
Year-month FEs	✓		✓	
Fund FEs		✓		✓
N	9,123	9,123	9,123	9,123
R^2	0.36	0.79	0.36	0.79

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

zero. If investors are indeed rebalancing from riskier holdings after reading S&P 500 fund letters, these outflows do not originate from their holdings in active domestic equity funds within the same fund family.

I next conduct a further test of the rebalancing hypothesis: this time, I test whether communication about future risks contained in the letters of the higher-risk active domestic equity funds drives flows to lower-risk S&P 500 index funds. Columns 3-4 in Table 9 conduct this analysis, and the results are similar since there is no statistically significant effect.

6.2 Are Investors Updating Their Beliefs Based on Signals in the Letters?

Instead of communication decreasing investors' effective risk aversion, an obvious alternative is for investors to simply update their beliefs about risk given the communication about future risks. I now conduct a number of tests for this channel.

Bayesian Investor Who Learns About Risk To fix ideas in this section, consider a Bayesian investor who learns about the variance parameter σ^2 of the return distribution, or indeed some other distributional parameter that captures the notion of "risk." The Bayesian investor starts with a Gaussian prior on this parameter, $\sigma_p^2 \sim \mathcal{N}(m_p, {}^1/_{\rho_p})$, in which ρ_p denotes the prior precision. The investor receives an independent signal $s = m_s + \varepsilon$ about the variance, with $\varepsilon \sim \mathcal{N}(0, {}^1/_{\rho_s})$. Upon receipt of the signal, the investor updates her beliefs and so now perceives the posterior distribution of the variance to be the following:

$$\sigma^2|s \sim \mathcal{N}\left(\alpha s + (1-\alpha)m_p \ , \ \frac{1}{\rho_s + \rho_p}\right), \tag{13}$$

where $\alpha = \frac{\rho_s}{\rho_s+\rho_p}$. Upon receipt of the signal, therefore, the shift in the investor's expected variance from her prior mean m_p to the posterior is:

$$\Delta(\sigma^2) \ = \ [\alpha s + (1-\alpha)m_p] - m_p \ = \ (s - m_p)\frac{\rho_s}{\rho_s + \rho_p}. \tag{14}$$

As Equation 14 makes clear, the difference between the signal value and the investor's prior, $(s-m_p)$ is a first-order determinant of the change in her perceived risk. If the investor chooses her portfolio as in Section 2, the difference $(s-m_p)$ thus determines how much she will update her position in the risky asset. This paper takes the risky asset to be an S&P 500 index mutual fund, and therefore $(s-m_p)$ should be directly proportional to the investor's flow into/out of the fund.

Extension to Lognormal-Lognormal Updating The assumption that a return variance is Normally distributed is not necessarily ideal, since the variance should have positive support. However, the conclusions from the previous example still apply. In fact, the previous argument using Normal-Normal updating carries through exactly if the *log* of the variance is assumed to be Normally distributed, which guarantees a positive variance.

6.2.1 Detail or Sentiment as a Signal of Risk

Detail Table 1 shows that the level of detail

TODO: continue the below

I first consider whether investors respond to the level of detail conveyed for instrumental reasons – that is, whether they respond to it *because* it signals forward-looking risk. If the level of detail is high when an objective measure of future risk is low, then investors could be inferring a signal about future risks from this communication and be behaving appropriately. This hypothesis can be tested by comparing the measured variable against a contemporaneous

objective (and publicly available) benchmark that also relates to the S&P 500 market portfolio: the VIX index. The VIX captures the market's expected volatility over the ensuing month, which is precisely the period over which I measure fund flows. The results displayed earlier in Table IA.12 are inconsistent with the proposed hypothesis. Instead, they show that the level of the VIX is *positively* correlated with the level of detail provided by fund managers — managers write more when an objective measure of future risk indicates that risk is elevated. When investors increase their risk-taking in response, they do so precisely at moments when risk is high, instead of decreasing their risk-taking as predicted by classic portfolio choice models

Sentiment A similar argument could be made for the level of sentiment: whether the prognoses communicated are optimistic or pessimistic should matter if the channel involves investors updating their beliefs about the payoff or return of the S&P 500 market portfolio. If, on the other hand, investors treat the prognoses as a source of noninstrumental information that builds up their confidence and trust, the direction of the sentiment expressed within them should not matter. The results shown earlier in Table 4, indicate that the direction of the sentiment does not, in fact, matter to investors.

Furthermore, it cannot be the case that investors are ignoring the sentiment of the letters because these signals are misleading. Under this alternative view, it could be that fund letters convey an overly-rosy view of future risks in comparison to an objective benchmark, and if so, investors might internalize the deception by rationally ignoring the conveyed sentiment, as in a cheap talk "babbling equilibrium." Once again, this hypothesis can be rejected by comparing sentiment to the objective VIX index. The regressions in columns (3) and (4) of Table IA.12 investigate the determinants of the sentiment expressed in fund letters about future risks. I find that managers appear to be collectively honest in how they adjust the sentiment of their communications in response to changing market conditions. The negative coefficients in the first row indicate that fund letters are optimistic (i.e. high sentiment) when the risk is low (i.e. low VIX), and vice versa. Investors should therefore be able to infer the objective level of future risks by listening to the sentiment conveyed. That they do not implies they are not treating these letters as a source of information with which to update their beliefs.

6.2.2 Updates From the Prior Sentiment

My discussion of sentiment so far implicitly assumes fund letter readers share a common prior — in such a scenario, the common prior is controlled for by the inclusion of time effects in Table 4. However, it is possible that investors could differ in their priors about future returns. If so, testing whether investors who read communications about risks use the conveyed sen-

timent to update their priors must also include a measure of prior sentiment. I now produce a measure of prior sentiment, and show that belief-updating does not explain the empirical findings.

The measure of prior sentiment that I use is the fund i-month t-varying asset-weighted average optimism of the readers of these letters, as revealed through local Google search activity. Prior work by Da, Engelberg, and Gao (2015) uses Google search activity to measure aggregate sentiment; I show how to measure the prior sentiment of each fund's letter-reading investor base.

I begin by creating a measure of who reads each letter. To meet this challenge, I focus on the readers of fund letters who download them in the form of shareholder reports from the SEC's EDGAR website. The SEC makes EDGAR website usage data up to June 2017 public (Lee, Ma, and Wang, 2015), and a number of studies have used these logs to measure information consumption. Further details on the SEC EDGAR usage logs are provided in Section 3.6 and Internet Appendix A.2.

Employing a methodology similar to Grice and Gueciouer (2023), I measure which states the readers of fund shareholder reports are located in; these fund shareholder reports contain the fund letters that the present paper focusses on. I then use the geo-locations of these readers to construct an aggregate Readership Prior Optimism index for each fund i-month t pair, based on the characteristics of the fund's unique report readers j:

$$\text{Readership Prior Optimism}_{i,t} = \frac{1}{\sum_j \text{Investment}_{j,t}} \sum_j \text{Investment}_{j,t} \times \text{Geographic Prior Optimism}_{j,t}. \tag{15}$$

Investment weights are proxied using SIPP survey values. The Geographic Prior Optimism index varies at a state-month level. The state-to-fund mappings for each fund's investor base are re-computed annually.

To define the Geographic Prior Optimism index, I obtain state s-month t-varying Search Volume Indices (SVI) from the Google Trends API, for search terms relating to the general topic of stock market crashes, as defined by Google's search engine itself. This comprises searches for "stock market crash" and related terms such as "market crash" and "stock crash". I obtain a monthly timeseries of the aggregated SVI across the entire United States, as well as within-month SVIs across the cross-section of states. For each term, I separately merge and rescale individual SVIs to produce a term-specific month-state panel of comparable SVI values. The Geographic Prior Optimism index is then calculated as negative of the stock market crash search topic index:

$$\text{Geographic Prior Optimism}_{s,t} = -\text{SVI}_{s,t}^{\text{stock market crash topic}}. \tag{16}$$

The United States-wide aggregate version of the index, $\text{SVI}_t^{\text{stock market crash topic}}$, covaries strongly

with the VIX: in Internet Appendix A.3, I compare the two using an array of regression specifications including various time trends (to capture growth in search index usage) and varying functional forms, and show this relationship is robust. In the current section, I make use of state s-month t cross-sectional variation to proxy for the prior optimism held by different investors about market risk.[15]

Table 10: Interaction between communication detail & changes to sentiment about future risks, and the relation with S&P 500 index fund flows. This table shows the relationship between net flows, communication detail, the update to sentiment due to the communication, and their interaction. The main covariate measures the level of detail contained in the letter written by fund i at time t about forward-looking risks, using the log of the number of words devoted to that topic. The sentiment measure is calculated on forward-looking risk-related statements. As shown in Eqn. (17), the change in sentiment, Δ Sentiment$_{i,t}$, is calculated as the difference between the (standardized) net sentiment of the forward-looking risk-related communication contained in the letter, and the (standardized) prior sentiment described in Section 6.2.2. The standardization is performed across the whole sample for columns 2-3, and within-fund for columns 4-5. Controls (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. The sample consists of S&P 500 index funds only.

Dependent Variable:	Net Flow$_{i,t\to t+1}$ (%)				
	(1)	(2)	(3)	(4)	(5)
Detail$_{i,t}$	0.5228***	0.5235***	0.5138***	0.5192***	0.5160***
	(0.1846)	(0.1838)	(0.1848)	(0.1816)	(0.1829)
Δ Sentiment$_{i,t}$		-0.1136	0.0493	-0.1452	-0.0659
		(0.0951)	(0.2610)	(0.1094)	(0.3042)
Detail$_{i,t}$ $\times$ Δ Sentiment$_{i,t}$			-0.0369		-0.0157
			(0.0497)		(0.0474)
Z-score Calculation		Unconditional		Within-fund	
Fund Controls$_{i,t}$	✓	✓	✓	✓	✓
Year-month FEs	✓	✓	✓	✓	✓
Fund FEs	✓	✓	✓	✓	✓
N	901	901	901	898	898
R^2	0.45	0.45	0.45	0.45	0.45

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

[15]This part of the analysis would be unnecessary if it were possible to compute an index such as the VIX that is localized for specific geographic subgroups of investors — however no such index exists. I therefore compute my own index based on Google search traffic. The validation checks in Internet Appendix A.3 confirm that the aggregate version of the index and the VIX both capture investors' perceptions of future risk. At the time of writing (August 2023), the publicly-accessible location to download the "Stock Market Crash" topic SVI was at `https://trends.google.com/trends/explore?q=/m/0gz_4&date=all&geo=US`.

It remains to measure the strength of belief updating induced by the receipt of a fund letter on its readership. This is simply the difference between the standardized (i.e. z-scored) net sentiment of the forward-looking risk-related communication in a letter, and the standardized prior optimism of the letter's readership:

$$\Delta\text{Sentiment}_{i,t} = z(\text{Net Sentiment}_{i,t}) - z(\text{Readership Prior Optimism}_{i,t}). \tag{17}$$

I conduct a set of panel regressions that are partly similar to the previous specifications (of Table 4), with a crucial modification: instead of including the sentiment of the letter's risk-related communication, I instead include the proxy for the *change in sentiment* among the letter's readership (Eqn. 17) induced by letter's message. The results are presented in Table 10.

The coefficients on $\Delta\text{Sentiment}_{i,t}$ in columns (2) and (4) of Table 10 are negative in sign, and not significantly different to zero. If the readers of a letter update their sentiment about risks in an optimistic direction, these coefficients should be positive and significant: i.e. an increase in sentiment from the prior should induce an inflow to the fund. When an interaction term is included in column (5), the sign becomes positive but remains insignificant. By contrast, the main coefficient on the level of detail in the first row always remains positive and significant. These results are therefore not consistent with a classic belief-updating mechanism, such as Bayesian learning about some risk parameter (such as the variance parameter of a conditional distribution of market returns).

6.2.3 Do Fund Letters Invoke Belief-Updating at All?

Having ruled out a number of belief-related hypotheses, I now consider more broadly whether communication about future risks acts to change investor beliefs about risk at all. A wide class of what DellaVigna and Gentzkow (2010) term belief-based persuasion models involve the sender inducing a shift in the receiver's posterior belief through communication. DellaVigna and Gentzkow (2010) suggest a single test for this class of models: "persuasion will tend to be more effective when receivers are less certain about the truth. In Bayesian models, the weaker receivers' priors are, the more their beliefs are affected by a given piece of new information (holding everything else constant)."[16] Classic strategic models of communication and persuasion include a core Bayesian updating element, including those of Crawford and Sobel (1982) and Kamenica and Gentzkow (2011).

[16] A key advantage of this test is that it also applies to belief-based models in which agents deviate from rational Bayesian updating, as long as their behavior varies less as they are (*ex ante*) more certain about the state of the world before receiving the communication. Furthermore, by testing another relationship between flows and belief-based parameters, the current test is complementary to the previous ones, including that in Section 6.2.2.

Effect of the Strength of the Investor's Prior As a concise illustration of how the strength of the prior affects the strength of belief updating, consider again the Bayesian investor who learns about the variance parameter σ^2 of the return distribution, as described at the beginning of Section 6.2. Define the strength of her update from the prior mean m_p to the posterior similarly as the magnitude of the signed update (Eqn. 14):

$$\left|\Delta(\sigma^2)\right| \;=\; \left|s - m_p\right| \frac{\rho_s}{\rho_s + \rho_p}. \tag{18}$$

It is then clear that the strength of the update is decreasing in the strength of the prior:

$$\frac{\partial}{\partial \rho_p}\left|\Delta(\sigma^2)\right| \;=\; -\left|s - m_p\right| \frac{\rho_s}{(\rho_s + \rho_p)^2} \;<\; 0. \tag{19}$$

As before, this conclusion is not overly dependent on the assumption of Normal-Normal updating; for instance, Lognormal-Lognormal updating about the variance leads to a similar implication.

Empirical Test I now test whether more detailed communication about future risks interacts with the strength of the representative investor's prior about the level of future risks. I measure the prior by the VVIX, the so-called "VIX of the VIX." This public index measures the expected volatility of the 30-day forward price of the VIX. The VVIX is therefore informative of "the degree of confidence the market has in its forecast of future values of the VIX," according to its publisher, the CBOE.[17]

I interpret a higher level of the VVIX as corresponding to a weaker, more diffuse prior about future risks. That is, a high value of the VVIX suggests the representative investor is less confident about her assessment of future risks. Under a belief-based persuasion model, two identical signals (with the same level of precision) should invoke different responses according to whether the investor has high or low prior confidence: with high prior confidence, the investor should respond only weakly and take little action; with low prior confidence, the investor should update her beliefs more strongly, thus manifesting as higher inflows.

Table 11 presents the results of just such an analysis. To be precise, I repeat this paper's main flow-versus-detail panel regressions, but this time include an interaction term for low prior confidence in the level of future risks. This corresponds to market states with above-median levels of the VVIX. The coefficient on this interaction between having low prior confidence about future risks and the level of detail about future risks conveyed in fund letters is insignificant. Moreover, the direction is opposite to that expected: instead of being positive, the estimate is in fact negative, suggesting the representative investor updates *less* strongly

[17] Source: `https://www.cboe.com/us/indices/dashboard/vvix/`.

Table 11: Communication-driven flows and *ex ante* strength of the prior about future risk. The level of detail contained in the letter written by fund i at time t about forward-looking risks is measured using the log of the number of words devoted to that topic. The indicator variable $\mathbb{1}\{\text{Low Prior Confidence}\}_t$ is 1 when the $\underline{\text{V}}$VIX is above the median, and 0 otherwise; a baseline effect is estimated in addition to the interaction shown in the table. *Fund controls* (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. *Market controls* (not shown) are for the (log) VIX level and for aggregate expected returns (proxied by the AAII survey employed by Greenwood and Shleifer (2014)); note these are not compatible with time effects. The sample consists of S&P 500 index funds only. Each flow variable is the next month's flow, in percentage points of TNA.

Dependent Variables:	Net Flow$_{i,t\to t+1}$ (%)		Inflow$_{i,t\to t+1}$ (%)		Outflow$_{i,t\to t+1}$ (%)	
	(1)	(2)	(3)	(4)	(5)	(6)
Detail$_{i,t}$	0.3826***	0.4205***	0.4332***	0.4912***	0.0336	0.0862
	(0.1298)	(0.1398)	(0.1592)	(0.1627)	(0.0591)	(0.0678)
Detail$_{i,t}$	-0.0309	-0.0137	-0.1078	-0.0827	-0.0844	-0.1041
$\times\ \mathbb{1}\{\text{Low Prior Confidence}\}_t$	(0.0939)	(0.1037)	(0.1072)	(0.1182)	(0.0676)	(0.0709)
Fund Controls$_{i,t}$	✓	✓	✓	✓	✓	✓
Fund FEs	✓	✓	✓	✓	✓	✓
Market Controls$_t$	✓		✓		✓	
Year-month FEs		✓		✓		✓
N	1,588	1,588	1,588	1,588	1,588	1,588
R^2	0.32	0.41	0.73	0.77	0.84	0.86

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

when she is less confident *ex ante*.

These results suggest the underlying persuasive mechanism is unlikely to be belief-based.

6.2.4 Are Investors Learning About Returns and Converging in Posterior Precision?

Instead of updating their beliefs about risk directly, could the results be explained by investors learning about future *returns* from the content of the fund letters, and *then* forming beliefs about the second moment of the return/payoff? Put differently: is an uncertainty-reduction channel at play?

Example with Normal-Normal Bayesian Updating About the Return To illustrate such a mechanism, consider an investor who learns about the return r (or payoff, or fundamental value) of an asset in general — instead of learning about the variance or some risk-related parameter, specifically. The investor starts with a Gaussian prior on the return, $r_p \sim \mathcal{N}(m_p, {}^1/_{\rho_p})$,

in which ρ_p denotes the prior precision. The investor receives an independent signal $s = m_s + \varepsilon$ about the overall return, with $\varepsilon \sim \mathcal{N}(0, {}^1/_{\rho_s})$. Upon receipt of the signal, the investor updates her beliefs and so now perceives the posterior distribution of the return to be the following:

$$r|s \sim \mathcal{N}\left(\alpha s + (1-\alpha)m_p \, , \, \frac{1}{\rho_s + \rho_p}\right), \tag{20}$$

where $\alpha = \frac{\rho_s}{\rho_s + \rho_p}$. Importantly, the precision has increased from ρ_p to $(\rho_s + \rho_p)$; equivalently the posterior variance $\mathbb{V}\text{ar}[r|s] = {}^1/_{(\rho_s + \rho_p)}$ has decreased in comparison to the prior variance ${}^1/_{\rho_p}$. Uncertainty is thus reduced, irrespective of the actual value (realization) of the signal s that the investor receives. Consider now that the investor chooses between a risky asset (with return r) and a riskless asset, and the quantity $\mathbb{V}\text{ar}[r|s]$ is defined to be her perceived risk — as in Grossman and Stiglitz (1980), for example. Then the receipt of *any* signal about the return will mechanically reduce the investors' perceived uncertainty, and hence lead her to increase her allocation to the risky asset.

Learning From the Full Letter Text In the case of the full text of fund letters, this channel is already controlled for since a control for the (log) total number of words is always included in this paper's flow-detail regressions. In other words, even in the case that investors are forming a more precise posterior about returns from *all* the text contained in fund letters, my empirical analyses detect an incremental effect that is specifically due to communication about risks. This effect of communication about future risks is the focus of the present paper.

Learning From Statements Specifically About Performance or Returns I now conduct an explicit test for the presence of such an uncertainty-reduction channel, precisely where it is most likely to be apparent: I detect statements in fund letters that are forward-looking and devoted to performance or returns specifically (rather than risk). I do so by modifying the set of regular expressions used, as illustrated by Figure 6. While statements devoted to performance and returns occur frequently in the fund letter corpus, only 23% of these are also forward-looking; by contrast, 61% of risk-related statements were also forward-looking.

I next examine how fund flows respond to the level of detail in fund letters concerning future performance or returns, explicitly. The logic of this test is the following: if investors are learning about the payoff/return from statements about future risk, they should do the same from statements that are explicitly about the payoff/return itself. If anything, the effect should be even stronger when using the measure. Conversely, if a different channel specifically related to risk-related information is at play, the analysis will be a placebo with no effect detected.

Table 12 presents the results of repeating the paper's main detail-flow regressions, but

Figure 6: Forward-looking performance- and return-related statements. Forward-looking statements within fund letters relating to performance & returns are defined as sentences that are matched by both forward-looking and performance or return-related regular expressions. The percentage figures denote the fraction of all sentences in the fund letter corpus that are matched by each type.

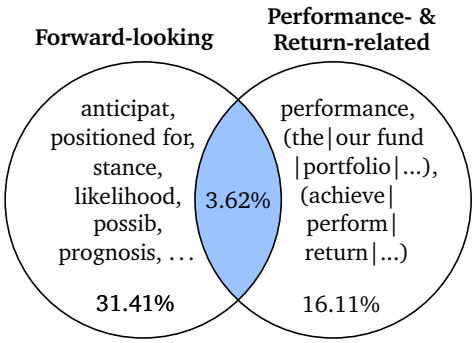


Table 12: Placebo test for the level of detail communicated about future performance & returns on the flows of S&P 500 index funds. The main covariate measures the level of detail contained in the letter written by fund i at time t about forward-looking performance & returns, using the log of the number of words devoted to that topic. Controls (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. The sample consists of S&P 500 index funds only. Each flow variable is the next month's flow, in percentage points of TNA.

Dependent Variables:	Net flow	Inflow	Outflow
	(1)	(2)	(3)
Detail about Performance & Returns$_{i,t}$	-0.0488	0.0803	0.1067
	(0.0927)	(0.1305)	(0.1054)
Fund Controls$_{i,t}$	✓	✓	✓
Year-month FEs	✓	✓	✓
Fund FEs	✓	✓	✓
N	882	882	882
R^2	0.50	0.81	0.87

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

where the level of detail in fund letters is measured from forward-looking performance- and return-related statements. The estimated coefficients in Table 12 exhibit no significant relation between the level of detail conveyed and fund flows. The absence of such an effect suggests that detailed textual communication about performance & returns (explicitly) also does not act

to reduce investors' perceived risks, as in the stylized setup of Grossman and Stiglitz (1980).

6.3 Does Detailed Communication Alleviate Investors' Ambiguity Aversion?

An extensive literature on ambiguity aversion in portfolio choice suggests that investors who are increasingly averse to ambiguity allocate decreasing fractions of their portfolio to the risky asset.[18] In principle, it is therefore possible that communication about future risks acts to decrease investors' ambiguity about the distribution of future returns. If this does not shift investors' risk aversion (or their subjective perception of risk), or even increases it to some small extent, it is possible that the net effect of this communication will be to increase the risk-bearing appetite of investors, leading them to increase their holdings in the risky asset and thus driving inflows into the S&P 500 index fund.

Testing this potential explanation without measures of investors' ambiguity aversions, risk aversions and beliefs is challenging. My empirical test is based on the following logic: if communication about future risks soothes the ambiguity aversion of investors who fear ambiguity, it will act both at times of low perceived ambiguity (when it is least needed) and at times of high perceived ambiguity (when it is most needed). Instead, I find that investors do not seem to respond to communication about future risks during high-ambiguity periods.

I use Brenner and Izhakian's (2018) measure of the degree of perceived ambiguity, inferred from S&P 500 returns, to split the sample of fund letters into those that are published during below-median perceived ambiguity periods and above-median perceived ambiguity periods. I conduct 2SLS regressions of net flows, gross inflows and gross outflows on the level of detail communicated about future risks, adding fund fixed effects and measures of expected risk (the log of the VIX) and expected return (based on the AAII survey) in addition to the standard fund-level and fund-family level controls. Table 13 presents the results of this analysis. Much like the main findings, increased communication about future risks induces investors to increase their holdings in S&P 500 index funds during low ambiguity periods (columns 1-3). However, this effect mostly disappears during high-ambiguity periods (columns 4-6): coefficient magnitudes shrink greatly and are no longer statistically distinguishable from zero.

This analysis suggests that investors either ignore or increasingly discount the signals transmitted by fund letters during periods of high ambiguity, which is inconsistent with a mechanism in which fund letters alleviate investors' ambiguity aversion.

[18] For example, the model of Maenhout (2004, Proposition 1) makes this point clearly: he shows that an ambiguity-averse investor with CRRA utility who solves the Merton portfolio choice problem holds a position in the risky asset that is equivalent to what she would have held if she had increased her risk aversion parameter γ to the level $\gamma + \theta$, where θ captures the strength of the investor's preference for robustness. This result highlights that ambiguity aversion and risk aversion both act in the direction of dampening an investor's risk-bearing ability.

Table 13: Aggregate ambiguity levels, and the differential effects of the level of detail communicated about future risks on S&P 500 index fund flows. The main covariate measures the level of detail contained in the letter written by fund i at time t about forward-looking risks, using the log of the number of words devoted to that topic. The sample consists of S&P 500 index funds only, and is split into periods of below-median market ambiguity (columns 1-3) and above-median market ambiguity (columns 4-6), according to the ambiguity measure of Brenner and Izhakian (2018). Two market-level controls with timeseries variation are included: (log) VIX for the level of S&P 500 forward-looking risk, and aggregate expected returns (proxied by the AAII survey employed by Greenwood and Shleifer (2014)). Additional controls (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. Each flow variable is the next month's flow, in percentage points of TNA.

Sub-sample:	Low Ambiguity			High Ambiguity		
Dependent Variables:	Net flow	Inflow	Outflow	Net flow	Inflow	Outflow
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Detail}_{i,t}$	0.6187***	0.6618**	0.0093	0.0908	0.0916	-0.0793
	(0.2276)	(0.2543)	(0.1355)	(0.1132)	(0.1643)	(0.1062)
Log VIX_t	-0.4810	0.1039	0.7005**	0.5741**	0.7977**	0.2576**
	(0.5294)	(0.4298)	(0.2816)	(0.2655)	(0.3211)	(0.1224)
$\text{AAII Expected Return}_t$ (%)	0.0095	0.0167	0.0052	-0.0041	-0.0056	0.0016
	(0.0097)	(0.0122)	(0.0074)	(0.0082)	(0.0162)	(0.0064)
$\text{Fund Controls}_{i,t}$	✓	✓	✓	✓	✓	✓
Fund FEs	✓	✓	✓	✓	✓	✓
N	808	808	808	780	780	780
R^2	0.33	0.73	0.84	0.45	0.81	0.88

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

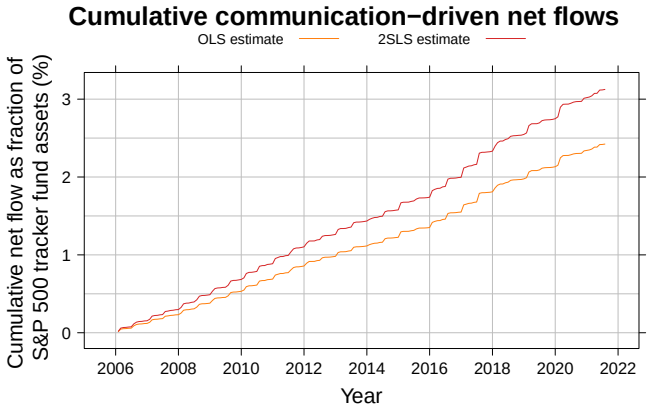
7 Effect of Aggregate Flows on Prices

This section quantifies the aggregate asset pricing implications of this paper's fund-level results. Given my focus on funds that track the market portfolio, it is simple to calculate of the effect of such flows on the market value. While reliant on macro multiplier estimates from other studies, this analysis nevertheless is a useful confirmation that the effects identified in this paper are economically meaningful and have important asset pricing implications.

I begin by aggregating the fund-level flow effects of Section 4.1; specifically, the estimates in Table 5 columns (4) and (8). Figure 7 charts these cumulative net flows into the S&P 500 (via index funds) as a percentage of index fund AuM. The flows are distributed evenly over

this paper's sample period of January 2006 to August 2021.

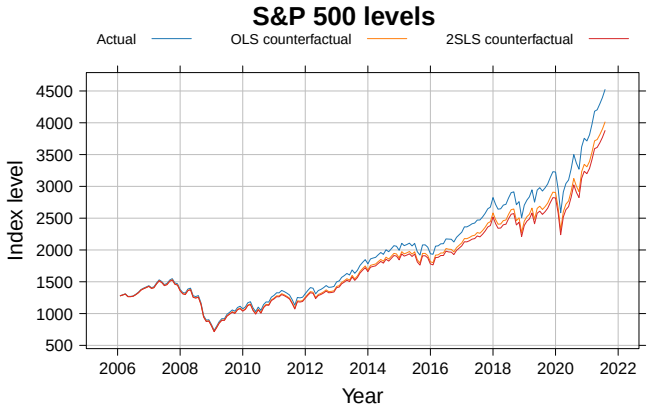
Figure 7: Aggregate communication-driven flows into S&P 500 index funds. Cumulative total of the net flows since January 2006 that were driven by increased communication about future risks, as a percentage of the assets under management of the S&P 500 index funds disseminating fund letters. The effect is aggregated from fund-level effects estimated using the specifications of Table 5 columns (2) and (6), and plotted as the orange and red series, respectively.



Aggregate Effect Based on Gabaix and Koijen's (2021) Multiplier Estimate Gabaix and Koijen (2021) show that flows into the stock market have a large price impact: their estimate of the macro multiplier is that a $1 inflow translates into an increase in value of $5. For the purposes of this paper, a 1% communication-driven flow would therefore produce a 5% return in the S&P 500 index. I use this measure of the stock market's (in)elasticity to translate communication-driven market flows into price impacts, according to a simple back-of-the-envelope calculation: I recalculate the communication-driven flows as a percentage of the total market capitalization of the S&P 500, and then multiply by 5 to estimate the causal effect of communication on market returns. To put the numbers in context, I adjust the historical levels of the S&P 500 by subtracting communication-driven returns induced from January 2006 onwards. Figure 8 plots the counterfactual levels of the S&P 500 that would have obtained in the absence of communication about future risks to the market portfolio.

The level of S&P 500 stood at 4,523 at the end of this paper's sample in August 2021. In the absence of communication-driven net flows, the counterfactual level of the index would have stood at only 4,012 (using OLS estimates) or a slightly lower 3,877 (using 2SLS estimates). The counterfactual gap is therefore on the order of 11% to 14% of the capitalization of the aggregate stock market. Although dependent on Gabaix and Koijen's (2021) estimate of the macro flow multiplier, this aggregation exercise puts the magnitude of the anxiety-alleviation mechanism into context, and confirms that it is economically meaningful. Furthermore, it suggests that communication-driven flows lead to substantial changes in asset prices.

Figure 8: Illustration of price effect of communication-driven flows on the S&P 500 index. The blue series plots the actual historical level of the S&P 500 over this paper's sample period of January 2006 to August 2021. The orange and red series plot counterfactual historical levels of the index in the absence of forward-looking risk-related communication disseminated the letters of S&P 500 index funds: First, the fund-level effect on net flows are aggregated from the fund-level specifications estimated in specifications of Table 5 columns (4) and (8). Next, the aggregate net flows are multiplied by 5 to compute the causal effect of market flows on the market index, based on the measured price (in)elasticity estimated by Gabaix and Koijen (2021). Finally, S&P 500 returns are adjusted (downward) to exclude these (positive) monthly return effects to compute counterfactual levels of the index.



Aggregate Effect Based on Hartzmark and Solomon's (2023) Multiplier Estimate As Gabaix and Koijen (2021) discuss, the market multiplier may be lower if the flows are considered to be anticipated. This would decrease the magnitude of the above estimates of the marketwide impact of anxiety-alleviating communication. It is simple to assess the robustness of the previous result to a different flow multiplier estimate.

Using anticipated divided payments, Hartzmark and Solomon (2023) find the market remains inelastic (albeit less so), with an estimated flow multiplier of 1.9. Combining this lower multiplier with my own estimates, the counterfactual level of the S&P 500 would have stood at 4,322 (using OLS estimates) or a slightly lower 4,266 (using 2SLS estimates) at the end of my sample period in the absence of anxiety-alleviating communication. The counterfactual gap thus shrinks to the order of 4% to 6% of the stock market's capitalization, which nonetheless remains economically meaningful.

8 Conclusion

In this paper, I find that money doctors who communicate in more detail about the riskiness of the fund's portfolio induce their investors to *increase* their share allocated to the risky asset. I contribute to the literature by illuminating a behavioral channel through which the theory of fund managers as money doctors (GSV) operates, with implications for the origins of trust

in other economic settings. My findings also highlight the potential benefits of persuasion and noninstrumental communication to investors, as well as the importance of preference-based persuasion (DellaVigna and Gentzkow, 2010) in the asset management industry.

References

Abis, S., Buffa, A. M., Javadekar, A., & Lines, A. (2022). Learning from prospectuses. *Available at SSRN*, *3865753*.

Akerlof, G. A., & Shiller, R. J. (2010). *Animal spirits: How human psychology drives the economy, and why it matters for global capitalism*. Princeton University Press.

Akey, P., Robertson, A., & Simutin, M. (2021). Closet active management of passive funds. *Rotman School of Management Working Paper*, (3874582).

Algan, Y., & Cahuc, P. (2010). Inherited trust and growth. *American Economic Review*, *100*(5), 2060–92.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, *113*(1), 151–184.

Andries, M., & Haddad, V. (2020). Information aversion. *Journal of Political Economy*, *128*(5), 1901–1939.

Arrow, K. J. (1972). Gifts and exchanges. *Philosophy & Public Affairs*, 343–362.

Auh, J. K., & Bai, J. (2020). Cross-asset information synergy in mutual fund families. *National Bureau of Economic Research Working Paper*, *26626*.

Aydogdu, M., & Wellman, J. W. (2011). The effects of advertising on mutual fund flows: Results from a new database. *Financial Management*, *40*(3), 785–809.

Bailey, W., Kumar, A., & Ng, D. (2011). Behavioral biases of mutual fund investors. *Journal of Financial Economics*, *102*(1), 1–27.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, *131*(4), 1593–1636.

Becker, G. S., & Murphy, K. M. (1993). A simple theory of advertising as a good or bad. *Quarterly Journal of Economics*, *108*(4), 941–964.

Ben-David, I., Franzoni, F., Moussawi, R., & Sedunov, J. (2021). The granular nature of large institutional investors. *Management Science*, *67*(11), 6629–6659.

Ben-David, I., Li, J., Rossi, A., & Song, Y. (2022). What do mutual fund investors really care about? *Review of Financial Studies*, *35*(4), 1723–1774.

Berk, J. B., & Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, *112*(6), 1269–1295.

Bhattacharya, U., Lee, J. H., & Pool, V. K. (2013). Conflicting family values in mutual fund families. *Journal of Finance*, *68*(1), 173–200.

Blinder, A. S. (2009). Talking about monetary policy: The virtues (and vices?) of central bank communication. *BIS Working Paper No. 274*.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, *77*(3), 623–685.

Brenner, M., & Izhakian, Y. (2018). Asset pricing and ambiguity: Empirical evidence. *Journal of Financial Economics*, *130*(3), 503–531.

Caetano, C., Caetano, G., Nielsen, E., & Sanfelice, V. (2023). The effect of maternal labor supply on children: Evidence from bunching.

Caetano, C., Caetano, G., & Nielsen, E. R. (forthcoming). Correcting for endogeneity in models with bunching.

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, *115*(5), 707–747.

Calvet, L. E., Célérier, C., Sodini, P., & Vallée, B. (2023). Can security design foster household risk-taking? *Journal of Finance*, *78*(4), 1917–1966.

Cao, S., Yang, B., & Zhang, A. L. (2021). Deep learning mutual fund disclosure: Risk sentiment, risk taking, and performance. *Available at SSRN*, *4060307*.

Cassar, G. J., Gerakos, J. J., Green, J. R., Hand, J. R., & Neal, M. (2018). Hedge fund voluntary disclosure. *The Accounting Review*, *93*(2), 117–135.

Chen, H., Cohen, L., Gurun, U., Lou, D., & Malloy, C. (2020). IQ from IP: Simplifying search in portfolio choice. *Journal of Financial Economics*, *138*(1), 118–137.

Chevalier, J., & Ellison, G. (1999). Are some mutual fund managers better than others? Cross-sectional patterns in behavior and performance. *Journal of Finance*, *54*(3), 875–899.

Choi, J. J., & Robertson, A. Z. (2020). What matters to individual investors? Evidence from the horse's mouth. *Journal of Finance*, *75*(4), 1965–2020.

Cici, G., Jaspersen, S., & Kempf, A. (2017). Speed of information diffusion within fund families. *Review of Asset Pricing Studies*, *7*(1), 144–170.

Crawford, V. P. (2019). Experiments on cognition, communication, coordination, and cooperation in relationships. *Annual Review of Economics*, *11*, 167–191.

Crawford, V. P., & Sobel, J. (1982). Strategic information transmission. *Econometrica*, 1431–1451.

Da, Z., Engelberg, J., & Gao, P. (2015). The sum of all FEARS: Investor sentiment and asset prices. *Review of Financial Studies*, *28*(1), 1–32.

Del Guercio, D., & Tkac, P. A. (2008). Star power: The effect of Morningstar ratings on mutual fund flow. *Journal of Financial and Quantitative Analysis*, *43*(4), 907–936.

DellaVigna, S., & Gentzkow, M. (2010). Persuasion: Empirical evidence. *Annual Review of Economics*, *2*(1), 643–669.

Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution*, *2*(4), 265–279.

Du, Q., Jiao, Y., Ye, P., & Fan, W. (2019). When mutual fund managers write confidently. *Available at SSRN*, *3513288*.

Eisele, A., Nefedova, T., Parise, G., & Peijnenburg, K. (2020). Trading out of sight: An analysis of cross-trading in mutual fund families. *Journal of Financial Economics*, *135*(2), 359–378.

Eliaz, K., & Schotter, A. (2010). Paying for confidence: An experimental study of the demand for non-instrumental information. *Games and Economic Behavior*, *70*(2), 304–324.

Ely, J., Frankel, A., & Kamenica, E. (2015). Suspense and surprise. *Journal of Political Economy*, *123*(1), 215–260.

Engelberg, J., & Parsons, C. A. (2016). Worrying about the stock market: Evidence from hospital admissions. *Journal of Finance*, *71*(3), 1227–1250.

Evans, R. B., Prado, M. P., & Zambrana, R. (2020). Competition and cooperation in mutual fund families. *Journal of Financial Economics*, *136*(1), 168–188.

Fechner, G. T. (1876). *Vorschule der aesthetik* (Vol. 1). Breitkopf & Härtel.

Fisman, R., & Khanna, T. (1999). Is trust a historical residue? Information flows and trust levels. *Journal of Economic Behavior & Organization*, *38*(1), 79–92.

Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. Free Press.

Gabaix, X., & Koijen, R. S. (2021). In search of the origins of financial fluctuations: The inelastic markets hypothesis. *National Bureau of Economic Research Working Paper*, *28967*.

Gallaher, S., Kaniel, R., & Starks, L. T. (2015). Advertising and mutual funds: From families to individual funds. *CEPR Discussion Paper*, (10329).

Gaspar, J.-M., Massa, M., & Matos, P. (2006). Favoritism in mutual fund families? Evidence on strategic cross-fund subsidization. *Journal of Finance*, *61*(1), 73–104.

Gennaioli, G. N., Shleifer, A., & Vishny, R. (2015). Money doctors. *Journal of Finance*, *70*(1), 91–114.

Giannetti, M., & Wang, T. Y. (2016). Corporate scandals and household stock market participation. *Journal of Finance*, *71*(6), 2591–2636.

Gil-Bazo, J., & Imbet, J. F. (2022). Tweeting for money: Social media and mutual fund flows. *Available at SSRN*, *3719169*.

Giorgetta, C., Grecucci, A., Zuanon, S., Perini, L., Balestrieri, M., Bonini, N., Sanfey, A. G., & Brambilla, P. (2012). Reduced risk-taking behavior as a trait feature of anxiety. *Emotion*, *12*(6), 1373.

Greenwood, R., & Shleifer, A. (2014). Expectations of returns and expected returns. *Review of Financial Studies*, *27*(3), 714–746.

Grice, R., & Gueciloeur, A. (2023). Mutual fund market structure and company fee competition: Theory and evidence. *Available at SSRN*, *4449026*.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, *70*(3), 393–408.

Guiso, L., Sapienza, P., & Zingales, L. (2004). The role of social capital in financial development. *American Economic Review*, *94*(3), 526–556.

Guiso, L., Sapienza, P., & Zingales, L. (2008). Trusting the stock market. *Journal of Finance*, *63*(6), 2557–2600.

Guiso, L., Sapienza, P., & Zingales, L. (2018). Time varying risk aversion. *Journal of Financial Economics*, *128*(3), 403–421.

Gurun, U. G., Stoffman, N., & Yonker, S. E. (2018). Trust busting: The effect of fraud on investor behavior. *Review of Financial Studies*, *31*(4), 1341–1376.

Haliassos, M., & Bertaut, C. C. (1995). Why do so few hold stocks? *The Economic Journal*, *105*(432), 1110–1129.

Harmon-Jones, E., & Allen, J. J. (2001). The role of affect in the mere exposure effect: Evidence from psychophysiological and individual differences approaches. *Personality and Social Psychology Bulletin*, *27*(7), 889–898.

Hartzmark, S. M., & Solomon, D. H. (2023). Marketwide predictable price pressure. *Available at SSRN*, *3853096*.

Hassan, T. A., Hollander, S., van Lent, L., & Tahoun, A. (2019). Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics*, *134*(4), 2135–2202.

Hayward, J. (1975). Information – a prescription against pain. *The Study of Nursing Care Project Reports*, (5). Royal College of Nursing, London.

Heath, D., Macciocchi, D., Michaely, R., & Ringgenberg, M. C. (2022). Do index funds monitor? *Review of Financial Studies*, *35*(1), 91–131.

Hillert, A., Niessen-Ruenzi, A., & Ruenzi, S. (2021). Mutual fund shareholder letters: Flows, performance, and managerial behavior. *Available at SSRN*, *2524610*.

Hortaçsu, A., & Syverson, C. (2004). Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds. *Quarterly Journal of Economics*, *119*(2), 403–456.

Huang, J., Wei, K. D., & Yan, H. (2007). Participation costs and the sensitivity of fund flows to past performance. *Journal of Finance*, *62*(3), 1273–1311.

Iliev, P., Kalodimos, J., & Lowry, M. (2021). Investors' attention to corporate governance. *Review of Financial Studies*, *34*(12), 5581–5628.

Jain, P. C., & Wu, J. S. (2000). Truth in mutual fund advertising: Evidence on future performance and fund flows. *Journal of Finance*, *55*(2), 937–958.

Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, *101*(6), 2590–2615.

Karpoff, J. M., Koester, A., Lee, D. S., & Martin, G. S. (2017). Proxies and databases in financial misconduct research. *The Accounting Review*, *92*(6), 129–163.

Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *Quarterly Journal of Economics*, *112*(4), 1251–1288.

Koijen, R. S., & Yogo, M. (2019). A demand system approach to asset pricing. *Journal of Political Economy*, *127*(4), 1475–1515.

Kostovetsky, L. (2016). Whom do you trust?: Investor-advisor relationships and mutual fund flows. *Review of Financial Studies*, *29*(4), 898–936.

Kuchler, T., & Stroebel, J. (2021). Social finance. *Annual Review of Financial Economics*, *13*, 37–55.

Kuhnen, C. M., & Knutson, B. (2011). The influence of affect on beliefs, preferences, and financial decisions. *Journal of Financial and Quantitative Analysis*, *46*(3), 605–626.

Lee, C. M., Ma, P., & Wang, C. C. (2015). Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics*, *116*(2), 410–431.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*(1), 146.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*(2), 267.

Loomis, J. L. (1959). Communication, the development of trust, and cooperative behavior. *Human Relations*, *12*(4), 305–315.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, *66*(1), 35–65.

Maenhout, P. J. (2004). Robust portfolio rules and asset pricing. *Review of Financial studies*, *17*(4), 951–983.

Maner, J. K., & Schmidt, N. B. (2006). The role of risk avoidance in anxiety. *Behavior Therapy*, *37*(2), 181–189.

Mankiw, N. G., & Zeldes, S. P. (1991). The consumption of stockholders and nonstockholders. *Journal of Financial Economics*, *29*(1), 97–112.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & McClosky, D. The Stanford CoreNLP natural language processing toolkit. In: In *Association for computational linguistics (acl) system demonstrations*. 2014, 55–60.

Massa, M. (2003). How do family strategies affect fund performance? When performance-maximization is not the only game in town. *Journal of Financial Economics*, *67*(2), 249–304.

Massa, M., Wang, C., Zhang, H., & Zhang, J. (2022). Investing in low-trust countries: On the role of social trust in the global mutual fund industry. *Journal of Financial and Quantitative Analysis*, *57*(1), 240–290.

McCloskey, D., & Klamer, A. (1995). One quarter of GDP is persuasion. *American Economic Review (Papers and Proceedings)*, *85*(2), 191–195.

Melcangi, D., & Sterk, V. (2020). Stock market participation, inequality, and monetary policy. *FRB of New York Staff Report*, (932).

Milgrom, P., & Roberts, J. (1986). Relying on the information of interested parties. *RAND Journal of Economics*, 18–32.

Mitali, S. F. (2022). How do mutual fund investors react to text-based uncertainty? *Available at SSRN*, *4259075*.

Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing*, *58*(3), 20–38.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204.

Raghunathan, R., & Pham, M. T. (1999). All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes*, *79*(1), 56–77.

Roussanov, N., Ruan, H., & Wei, Y. (2021). Marketing mutual funds. *Review of Financial Studies*, *34*(6), 3045–3094.

Schildberg-Hörisch, H. (2018). Are risk preferences stable? *Journal of Economic Perspectives*, *32*(2), 135–154.

Sergeyev, D., Lian, C., & Gorodnichenko, Y. (2023). The economics of financial stress. *National Bureau of Economic Research Working Paper*, *31285*.

Sheng, J., Xu, N., & Zheng, L. (2022). Do mutual funds walk the talk? A textual analysis of risk disclosure by mutual funds. *Available on SSRN*, *3757077*.

Sicherman, N., Loewenstein, G., Seppi, D. J., & Utkus, S. P. (2016). Financial attention. *Review of Financial Studies*, *29*(4), 863–897.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *Journal of Finance*, *53*(5), 1589–1622.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Tom, G., Nelson, C., Srzentic, T., & King, R. (2007). Mere exposure and the endowment effect on consumer decision making. *Journal of Psychology*, *141*(2), 117–125.

Warner, J. B., & Wu, J. S. (2011). Why do mutual fund advisory contracts change? Performance, growth, and spillover effects. *Journal of Finance*, *66*(1), 271–306.

Wei, C., & Zhang, L. (2020). Trust and local bias. *Journal of Financial and Quantitative Analysis*, *55*(7), 2211–2245.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2p2), 1.

Internet Appendix

A Validation of Data Sources . . . 2
A.1 Validation of Textual Measures . . . 2
A.2 Validation of Fund Letter Readership Measure . . . 5
A.3 Validation of Prior Geographic Optimism Measure . . . 7
B Further Identification Details . . . 9
B.1 Instrumental Variables Strategy . . . 9
B.2 Corner Bunching Strategy . . . 13
C Ruling Out More Alternative Explanations . . . 14
C.1 Are Investors Being Educated Rather Than Reassured? . . . 14
C.2 Do Managers Transmit Signals That Pander to Investor Beliefs? . . . 16
C.3 Are Investors Expressing a Preference for Early Resolution of Uncertainty? 18
C.4 Are Investors Inferring Index Funds' Tracking Errors? . . . 19
D Robustness Checks . . . 21
D.1 Fund-Level Timeseries Variation in Communication . . . 21
D.2 Measuring Information Spillovers Over an Extended Sample . . . 22
D.3 Alternative Measure of the Level of Detail Communicated . . . 23
D.4 Further Tests to Rule Out Within-Family Rebalancing . . . 24
E Regular Expressions to Replicate Textual Measures . . . 26
E.1 Extracting Fund Letters From Reports . . . 26
E.2 Classifying Sentences Within Fund Letters . . . 28
References in Internet Appendix . . . 32

A Validation of Data Sources

A.1 Validation of Textual Measures

I now validate this study's methodology for detecting sentence types in the fund letter corpus. To do this, I compare my methodology – which has been manually adapted to the fund letter setting – with methodologies from other studies that are not intentionally adapted to my study's context. This includes well-accepted methods in the Psychology literature that suffer from a lack of transparency in how they operate to detect sentence types due to being implemented as "black-box" software packages.

I compare and contrast the power of each of these methodologies to detect forward-looking and risk-related statements in the context of fund letters. I find that the present paper's methodology is either comparable to, or exceeds, the ability of other methodologies employed in other parts of the Accounting and Psychology literatures to detect such statements. This paper's methodology thus succeeds in its goal, while also having the advantages of being (i) transparent, and (ii) intentionally designed for the fund letter corpus studied in this paper.

The first two alternative methodologies that I examine in this section (Li, 2010; Hassan, Hollander, van Lent, and Tahoun, 2019) were designed for detecting either forward-looking (only) or risk-related (only) statements in listed company filings.

Although neither methodology was designed with fund letters in mind, a study by Hillert, Niessen-Ruenzi, and Ruenzi (2021) used the Li (2010) methodology as part of an analysis of mutual fund shareholder letters, suggesting that this comparison will be informative. A statement is classified as forward-looking if it contains at least one forward-looking word, as defined by Hillert, Niessen-Ruenzi, and Ruenzi (2021) & Li (2010).

Similarly, using the methodology of Hassan, Hollander, van Lent, and Tahoun (2019), a statement is classified as risk-related if it contains at least one risk-related word, as defined by Hassan, Hollander, van Lent, and Tahoun (2019) & the Oxford English Dictionary.

The next alternative methodology I consider is the Linguistic Inquiry and Word Count (LIWC) software, of which I use the latest version, LIWC-22 (Boyd, Ashokkumar, Seraj, and Pennebaker, 2022). This software contains a number of proprietary dictionaries that are kept confidential by its vendor; it is likely these dictionaries comprise both individual word stems and regular expressions. It is operated by inputting a block of text, and returns the number of words in the input text that match a given dictionary. While the software is well accepted in the field of Psychology, its lack of transparency is a barrier to adopting this software for analyzing fund letters. I run the software to count the number of words matching two dictionaries – future focus in time orientation, and risk motives – in order to detect forward-looking and risk-related statements (respectively) in my fund letter corpus. To be precise, a statement is defined as being forward-looking if it contains at least one forward-looking word, as determined by

the LIWC software; similarly, a statement is defined as being risk-related if the LIWC software detects the presence of a risk-related word.

Table IA.1 compares the results of this paper's baseline methodology for detecting forward-looking statements to the two alternative measures. The baseline methodology detects the highest fraction of forward-looking sentences, and its classifications are positively correlated with both alternative methodologies'.

Table IA.2 compares the results of this paper's baseline methodology for detecting risk-related statements to the two alternative measures. The baseline methodology detects a higher fraction of risk-related statements than that of Hassan, Hollander, van Lent, and Tahoun (2019), and a comparable rate to the LIWC-22 software. The baseline method's classifications are positively correlated with both alternative methodologies'.

Table IA.1: Forward-looking sentence detection. Number & fraction of forward-looking statements detected in my fund letter corpus using three different methodologies, and (in the final column) the correlations of statement classifications of the two alternative methodologies with my own.

Methodology	**Description**	**Sentences detected**	**Fraction of corpus**	**Correlation to baseline**
This paper	Set of custom regular expressions designed for the fund letter corpus	1,529,823	31.41%	—
Hillert, Niessen-Ruenzi, and Ruenzi (2021), using the dictionary of Li (2010, Appendix B)	17 words used by Li (2010) to detect forward-looking statements in listed company 10-K & 10-Q filings	877,338	18.01%	0.3103
LIWC-22 (Boyd, Ashokkumar, Seraj, and Pennebaker, 2022)	Proprietary set of 138 words & word stems related to future focus in time orientation, designed for general-purpose psychology research	854,105	17.53%	0.4023

Table IA.2: Risk-related sentence detection. Number & fraction of risk-related statements detected in my fund letter corpus using three different methodologies, and (in the final column) the correlations of statement classifications of the two alternative methodologies with my own.

Methodology	**Description**	**Sentences detected**	**Fraction of corpus**	**Correlation to baseline**
This paper	Set of custom regular expressions designed for the fund letter corpus	678,338	13.93%	—
Hassan, Hollander, van Lent, and Tahoun (2019, Internet Appendix Table 3)	111 synonyms for the word "risk" according to the Oxford Dictionary, used to analyze listed company filings by Hassan, Hollander, van Lent, and Tahoun (2019)	380,268	7.81%	0.6341
LIWC-22 (Boyd, Ashokkumar, Seraj, and Pennebaker, 2022)	Proprietary set of 128 words & word stems related to risk motives, designed for general-purpose psychology research	689,979	14.17%	0.4684

A.2 Validation of Fund Letter Readership Measure

This section presents evidence that the individuals who downloaded annual & semi-annual reports from the SEC EDGAR system are broadly representative of the overall US population, based on their observable geographic dispersion. The locations of these readers are used to measure their characteristics, such as their anxiety attitudes, by matching fund reader locations with geographic local trust data.

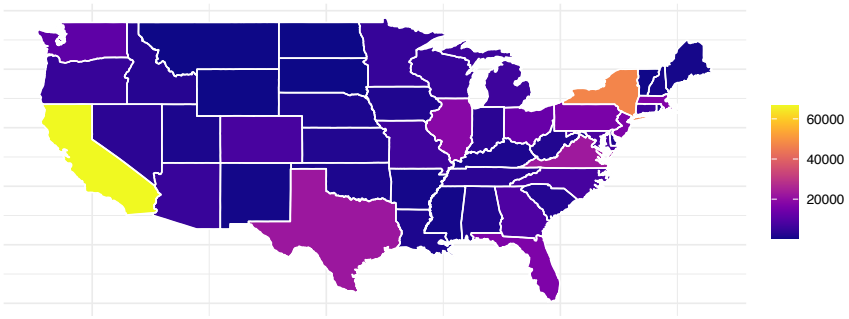
Figure IA.1(a) charts the state-level counts of annual and semi-annual report downloads from the SEC EDGAR system. Readers of funds' annual & semi-annual reports – which contain the fund letters used in this study – are located throughout the United States, rather than being concentrated in states that are home to financial institutions and professional investors, such as New York and Massachusetts (Kim, Wang, and Wang, 2022).

In fact, the correlation between state-level reader counts and the total state populations is 0.90 over the full sample used in this study, which runs from 2006 to the end of the SEC EDGAR log files in June 2017. At a more fine-grained unit of analysis, the correlation between county-level reader counts and the total county populations is also high, at 0.72. Figure IA.1(b) visualizes how county-level correlations between unique readers and the total population vary from state to state. All such correlations are positive, and for the most part uniformly high.

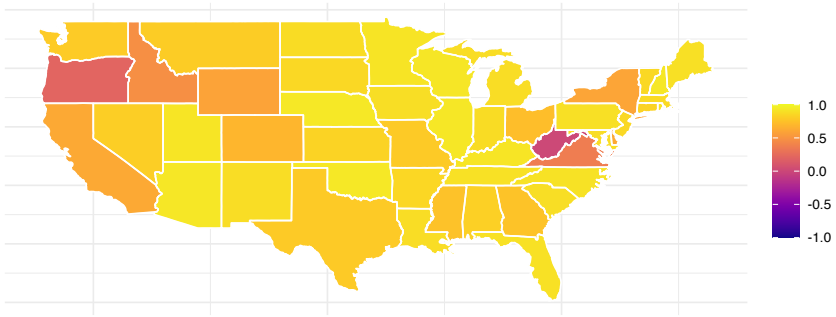
It is worth highlighting that similarly high correlations were found by Grice and Gueciouer (2023), who used a related (but separate) sample of fund *prospectus* readerships from SEC EDGAR, and found that model-derived competition measures based on this prospectus readership dataset predict mutual fund fee dispersion. Their study's findings validate SEC EDGAR as a way to measure investor behavior that affects mutual fund outcomes.

Figure IA.1: Annual & semi-annual report readership counts are broadly representative. Both charts visualize statistics derived from unique EDGAR user counts over the sample period 2006-2017. Users are geolocated based on their IP addresses. Sub-figure (a) displays the number of unique EDGAR annual & semi-annual report downloaders per state; lighter states have more unique readers, colored according to the legend. These reader counts are highly correlated with total population counts. Sub-figure (b) displays the correlations of unique reader counts with county population counts, for each state; lighter states exhibit higher correlations, colored according to the legend.

(a) State-level unique reader counts.



(b) Within-state reader-to-population correlations.



A.3 Validation of Prior Geographic Optimism Measure

In Equation (16), state s-month t Google search volume indices $\text{SVI}_{s,t}$ were used to define a measure of geographic prior optimism about stock market risk. This appendix examines the aggregate United States-wide version of the index, SVI_t, and compares it to the VIX index of aggregate implied stock market volatility.

Columns 1-3 of Table IA.3 regress the aggregate SVI_t against the VIX, together with a time trend (t) and a quadratic trend (t^2) to account for the growth in the use of Google searches over the entire sample period. In each specification, the SVI_t and VIX are positively and significantly associated. This validates Google search interest in this topic as a measure of sentiment with regard to stock market risk, at an aggregate level. Therefore geographic variation in Google search interest should also be a valid measure of geographic optimism about stock market risk.

Columns 4-5 of Table IA.3 regress the log of the SVI_t against the log of the VIX. Interpreting the log-log coefficient in column (4), a 10% increase in the level of the VIX is associated with an approximately 3% increase in Google search volume about the topic of stock market crashes. The positive association between the growths of the two variables thus parallels the positive association between their levels.

Table IA.3: Comparing aggregate Google search interest to the VIX. SVI_t is the monthly Google Search Volume Index across the United States about the topic of a stock market crash, as defined by Google themselves. The controls t and t^2 denote the presence of time trends and a quadratic growth term, respectively. The sample period of this analysis runs from 2004-2022.

Dependent Variables:	SVI_t			Log SVI_t	
	(1)	(2)	(3)	(4)	(5)
VIX_t	0.5228***	0.4938***	0.4896***		
	(0.1723)	(0.1669)	(0.1643)		
Log VIX_t				0.3076**	0.2638**
				(0.1291)	(0.1180)
Constant	10.25***	5.148	11.65***	2.020***	1.884***
	(3.022)	(3.265)	(3.165)	(0.3597)	(0.3464)
t		0.0501***	-0.1197***		0.0023***
		(0.0109)	(0.0342)		(0.0007)
t^2			0.0008***		
			(0.0002)		
N	225	225	225	225	225
R^2	0.15	0.23	0.29	0.07	0.19

Newey-West standard errors (in parentheses) use automatically-selected lags.
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

B Further Identification Details

B.1 Instrumental Variables Strategy

This appendix supplements Section 4.2.2 in the main text with further discussions of the underlying assumptions, as well as additional empirical evidence.

Evidence of Information Diffusion Within Fund Families The high and significantly non-zero estimates in columns (1) and (3) of Table IA.4 show that the level of detail provided in S&P 500 index fund letters is positively correlated with the average level of detail in similar fund letters that are written by domestic active equity funds, who are likely to conduct fundamental research. This is consistent with information spillovers within fund families. I provide a second piece of evidence by linking these spillovers to fund family incentives, which may encourage cooperation between fund managers. For this additional test, I include a family-level index of the extent to which families engage in cooperative behavior, in columns (2) and (4) of Table IA.4. This index was computed by Evans, Prado, and Zambrana (2020) based on a variety of incentives faced by managers, such as the structure of fund management and the design of compensation contracts. If the observed information spillovers are due to cooperation within fund families, one would expect a positive interaction term. That is indeed the case: the positive and significant interaction terms in columns (2) and (4) suggest that information spillovers are stronger when funds within the family have a greater incentive to cooperate.

Table IA.4 also makes clear that the instrument predicts the (log of the) number of words within the fund letters of S&P 500 funds in the same fund families. The results are robust to different fixed effects specifications, and to including a host of controls that will be employed in future analyses. The coefficients in columns (1) and (3) are positive and significantly different to zero: the instrument is relevant. Furthermore, F-stats around the 300 level suggest the instrument is also strong enough to test for coefficient significance even at the 99% level.[19]

Source of Variation Variation in w_A arises due to variation based on the signals uncovered by fund managers as part of their day-to-day operations. Different fund families would have a different mix of fund strategies; even within the same class of strategies, it is clear from the high volume of trade in financial markets that considerable heterogeneity in signal structures (and their realizations) must exist. S&P 500 indexing is a common fund strategy for fund families to offer, and therefore my sample of S&P 500 funds will be subject to heterogeneous

[19] According to Lee, McCrary, Moreira, and Porter (2022, Table 3), a first-stage F-stat of 104.67 is necessary to test for significance at the 95% level without adjusting standard errors. Similarly, an F-stat of 252.34 is necessary to test at the 99% level.

information spillovers from the other funds within those fund families, driving variation in the instrument.

Supplementary Results This appendix concludes with a number of additional empirical results: Tables IA.5 and IA.6 apply the instrumental variables strategy to analyze the response of gross inflows and gross outflows (respectively) to the level of detail provided. The results are similar to the un-instrumented results presented in Section 4.1.

Table IA.4: Instrument relevance, and evidence of information spillovers. The regressions in this table document a high correlation between the level of detail about forward-looking risks contained in S&P 500 index fund letters (dependent variable) written by fund i at time t, and the corresponding level of detail across active domestic equity sibling funds (only) within the same fund family, aggregated as a simple mean. *Fund controls* (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. Columns (1) & (3) show the unconditional strength of this relationship, across different fixed effects specifications, for the same sample used in future analyses. Columns (2) & (4) additionally include an interaction with an index for the strength of cooperation between sibling funds in the focal fund's overall family, as measured by Evans, Prado, and Zambrana (2020); their dataset ends in December 2015, leading to a smaller sample size in those columns.

Dependent Variable:	Log word count$_{i,t}$			
	(1)	(2)	(3)	(4)
Sibling Mean Detail$_{i,t}$	0.8492***	0.6102***	0.8120***	0.6379***
	(0.0334)	(0.0858)	(0.0405)	(0.0945)
Sibling Mean Detail$_{i,t}$		0.3308**		0.3791**
× Family Cooperation Index$_{i,t}$		(0.1323)		(0.1448)
Family Cooperation Index$_{i,t}$		-1.801**		-2.310***
		(0.7932)		(0.8524)
Fund Controls$_{i,t}$	✓	✓	✓	✓
Year-month FEs	✓	✓		
Fund FEs			✓	✓
N	1,155	640	1,155	640
R^2	0.94	0.94	0.95	0.95
F-stat	289.5	105.7	335.2	131.2

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table IA.5: Instrumented effect of the level of detail communicated about future risks on the inflows of S&P 500 index funds. The main covariate measures the level of detail contained in the letter written by fund i at time t about forward-looking risks, using the log of the number of words devoted to that topic.The instrument for the main covariate used in the 2SLS specifications is constructed from the log word counts (about the same topic of forward-looking risk) of the letters written by active domestic equity sibling funds within the focal index fund's overall family. *Fund controls* (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. *Market controls* (not shown) are for the (log) VIX level and for aggregate expected returns (proxied by the AAII survey employed by Greenwood and Shleifer (2014)); note these are not compatible with time effects. The sample consists of S&P 500 index funds only. Columns 1-3 show un-instrumented panel specifications. Columns 4-6 show 2SLS specifications, with F-stats corresponding to the first stage; the first-stage regressions are shown in Table IA.4.

Dependent Variable:	$\text{Inflow}_{i,t\to t+1}$ (%)							
Regression:	OLS				2SLS IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\text{Log word count}_{i,t}$	0.2124	0.4950***	0.4823***	0.4777**	0.4648*	0.5813***	0.5692***	0.6734***
	(0.2361)	(0.1756)	(0.1734)	(0.2008)	(0.2579)	(0.1968)	(0.1943)	(0.2471)
$\text{Fund Controls}_{i,t}$	✓	✓	✓	✓	✓	✓	✓	✓
$\text{Market Controls}_{t}$			✓				✓	
Year-month FEs	✓			✓	✓			✓
Fund FEs		✓	✓	✓		✓	✓	✓
N	1,155	1,155	1,155	1,155	1,155	1,155	1,155	1,155
R^2	0.25	0.74	0.74	0.79	0.25	0.74	0.74	0.79
F-stat (IV 1st-stage)					289.5	335.2	261.9	412.7

Clustered (Fund & Year-month) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table IA.6: Instrumented effect of the level of detail communicated about future risks on the gross outflows of S&P 500 index funds. The main covariate measures the level of detail contained in the letter written by fund i at time t about forward-looking risks, using the log of the number of words devoted to that topic. The instrument for the main covariate used in the 2SLS specifications is constructed from the log word counts (about the same topic of forward-looking risk) of the letters written by active domestic equity sibling funds within the focal index fund's overall family. *Fund controls* (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. *Market controls* (not shown) are for the (log) VIX level and for aggregate expected returns (proxied by the AAII survey employed by Greenwood and Shleifer (2014)); note these are not compatible with time effects. The sample consists of S&P 500 index funds only. Columns 1-3 show un-instrumented panel specifications. Columns 4-6 show 2SLS specifications, with F-stats corresponding to the first stage; the first-stage regressions are shown in Table IA.4.

Dependent Variable:	$\text{Outflow}_{i,t\to t+1}$ (%)							
Regression:	OLS				2SLS IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log word $\text{count}_{i,t}$	-0.0866	0.0236	0.0037	0.0524	0.0117	-0.0476	-0.0691	0.0311
	(0.1768)	(0.0825)	(0.0797)	(0.0983)	(0.2024)	(0.0993)	(0.0975)	(0.1228)
Fund $\text{Controls}_{i,t}$	✓	✓	✓	✓	✓	✓	✓	✓
Market Controls_t			✓				✓	
Year-month FEs	✓			✓	✓			✓
Fund FEs		✓	✓	✓		✓	✓	✓
N	1,155	1,155	1,155	1,155	1,155	1,155	1,155	1,155
R^2	0.24	0.83	0.83	0.85	0.24	0.83	0.83	0.85
F-stat (IV 1st-stage)					289.5	335.2	261.9	412.7

Clustered (Fund & Year-month) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

B.2 Corner Bunching Strategy

TODO: additional details of the shape restriction approach by Caetano, Caetano, and Nielsen (forthcoming) and Caetano, Caetano, Nielsen, and Sanfelice (2023)

C Ruling Out More Alternative Explanations

Section 6 rules out a number of potential alternative explanations for the paper's main results. This appendix continues to evaluate and rule out additional potential explanations.

C.1 Are Investors Being Educated Rather Than Reassured?

I now assess whether the communication related to future risks actually educates investors rather than simply reassuring them.

In my first test, I use state-level financial literacy scores that have been computed by Lusardi, Bumcrot, and Lin (2013) based upon geographical surveys. I aggregate these state-level scores to fund letter readership measures. Then I assess how the strength of the main detail-flow regressions vary with the financial literacy of fund letter readers.

My results are displayed in Table IA.7

And then additional results based on college graduation rates are presented in Table IA.8; these mostly confirm that the effect operates separately to education, although it should be noted college graduation rates are only a proxy for financial literacy.

Table IA.7: TODO: add to the caption

Dependent Variables:	Net flow	Inflow	Outflow
	(1)	(2)	(3)
$\text{Detail}_{i,t}$	0.5089***	0.6008**	-0.0267
	(0.1511)	(0.2507)	(0.0909)
$\text{Detail}_{i,t}$	0.0453	-0.1037	0.0169
$\times \mathbb{1}\{\text{Low Financial Literacy}\}_{i,t}$	(0.1369)	(0.1695)	(0.0962)
$\text{Fund Controls}_{i,t}$	✓	✓	✓
$\text{Market Controls}_{t}$	✓	✓	✓
Fund FEs	✓	✓	✓
N	901	901	901
R^2	0.29	0.68	0.80

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table IA.8: TODO: add to the caption

Dependent Variables:	Net flow	Inflow	Outflow
	(1)	(2)	(3)
$\text{Detail}_{i,t}$	0.3682**	0.3265*	-0.0246
	(0.1831)	(0.1776)	(0.1182)
$\text{Detail}_{i,t}$	0.2134	0.2691*	0.0467
$\times \mathbb{1}\{\text{Low College Graduation}\}_{i,t}$	(0.1384)	(0.1535)	(0.0950)
$\text{Fund Controls}_{i,t}$	✓	✓	✓
$\text{Market Controls}_{t}$	✓	✓	✓
Fund FEs	✓	✓	✓
N	1,155	1,155	1,155
R^2	0.31	0.74	0.83

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C.2 Do Managers Transmit Signals That Pander to Investor Beliefs?

One strand of the literature on behavioral persuasion posits that communication that is targeted to agree with a recipient's existing beliefs can persuade her to take an action favored by the sender of a message (Mullainathan and Shleifer, 2005). To test whether such a mechanism is currently at play, I examine how the distance between prior sentiment and conveyed sentiment interacts with aggregate flows.

I repeat the previous regressions of Table 10, but now take the absolute value of the sentiment update, $|\Delta \text{ Sentiment}_{i,t}|$, as a measure of distance between prior and message sentiment. Table IA.9 presents the results of these panel regressions.

Under a mechanism in which distance between the recipient's beliefs and the message induces inflows to the fund – the desired action of the fund manager who sends the message – one would expect be a negative and significant coefficient on the distance to beliefs term $|\Delta \text{ Sentiment}_{i,t}|$. Instead, the coefficients displayed in Table IA.9 columns 1-2 are positive in magnitude. Furthermore, none of the coefficients (including the additional interactions in columns 3 & 5) are significantly different to zero. These results are not consistent with the proposed pandering mechanism.[20]

[20] GSV discuss that fund managers may pander to investor beliefs when designing securities to attract investor clients in the first place. Such pandering in security design is distinct from pandering in persuasive communication.

Table IA.9: Interaction between communication detail & sentiment distance about future risks, and the relation with S&P 500 index fund flows. This table shows the relationship between net flows, communication detail, the distance of the communicated sentiment from the inferred prior, and their interaction. The main covariate measures the level of detail contained in the letter written by fund i at time t about forward-looking risks, using the log of the number of words devoted to that topic. The sentiment measure is calculated on forward-looking risk-related statements. As shown in Eqn. (17), the change in sentiment Δ Sentiment$_{i,t}$ is calculated as the difference between the (standardized) net sentiment of the forward-looking risk-related communication contained in the letter, and the (standardized) prior sentiment described in Section 6.2.2. The standardization is performed across the whole sample for columns 2-3, and within-fund for columns 4-5. The absolute value of the change in sentiment, $|\Delta$ Sentiment$_{i,t}|$, measures the distance between the aggregate prior sentiment of letter readers and the conveyed sentiment of the letter itself. Controls (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. The sample consists of S&P 500 index funds only.

Dependent Variable:	Net Flow$_{i,t\to t+1}$ (%)				
	(1)	(2)	(3)	(4)	(5)
Detail$_{i,t}$	0.5228***	0.5302***	0.5312**	0.5098***	0.4892**
	(0.1846)	(0.1852)	(0.2269)	(0.1754)	(0.2038)
$\|\Delta$ Sentiment$_{i,t}\|$		0.0321	0.0362	-0.0746	-0.1638
		(0.1014)	(0.2739)	(0.1358)	(0.3347)
Detail$_{i,t}$ $\times$ $\|\Delta$ Sentiment$_{i,t}\|$			-0.0009		0.0175
			(0.0621)		(0.0588)
Z-score Calculation		Unconditional		Within-fund	
Fund Controls$_{i,t}$	✓	✓	✓	✓	✓
Year-month FEs	✓	✓	✓	✓	✓
Fund FEs	✓	✓	✓	✓	✓
N	901	901	901	898	898
R^2	0.45	0.45	0.45	0.45	0.45

Clustered (Fund & Year-month) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C.3 Are Investors Expressing a Preference for Early Resolution of Uncertainty?

TODO: reference Table IA.10 and the paper by Schlag, Thimme, and Weber (2021), and discuss how (i) this doesn't make sense unless there's also some way that this leads them to increase their portfolio allocation (which isn't clear), and (ii) the results continue to hold anyway

Table IA.10: Effect of communication detail when the market prefers a late resolution of uncertainty. TODO: add to the caption

Dependent Variables:	Net flow (1)	Inflow (2)	Outflow (3)
$\text{Detail}_{i,t}$	0.6621*	0.6690*	-0.0760
	(0.3645)	(0.3663)	(0.2477)
$\text{Fund Controls}_{i,t}$	✓	✓	✓
$\text{Market Controls}_{t}$	✓	✓	✓
Fund FEs	✓	✓	✓
N	276	276	276
R^2	0.42	0.80	0.91

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C.4 Are Investors Inferring Index Funds' Tracking Errors?

Studying S&P 500 index-tracking funds is useful for identification, since this is an objectively homogeneous set of funds (Hortaçsu and Syverson, 2004). In particular, it is unlikely that investors are learning about fund managers' skill, in the style of Berk and Green (2004), since no active management is involved. To the extent that manager skill is involved, though, it is in the reduction of tracking error compared to the funds' common index. I therefore directly tackle the question of whether investors' are learning about future tracking errors from the level of detail provided about future risks.

In the literature and in practice, the tracking error is defined as the standard deviation of the difference between monthly fund returns and the S&P 500 index returns (e.g. Chevalier and Ellison, 1997; Grinold and Kahn, 1999; Cremers and Petajisto, 2009). Likewise, I define the future tracking error over some future horizon h as

$$\text{Tracking Error}_{i,t+1\to t+h} = \text{StdDev}\Big\{ r_{i,t+k} - r_{t+k}^{S\&P\,500} \;\Big|\; 1 \le k \le h \Big\}. \tag{21}$$

The assumption underlying the test is as follows: if investors respond to the level of detail because it signals a low tracking error, I assume it should have predictive power for the realized tracking error in the period following the dissemination of the fund letters. The placebo test in Table IA.11 confirms that this is *not* actually the case: the level of detail conveyed about future risks is not a signal for how well a fund will track its S&P 500 index going forward.

Table IA.11: Placebo test for tracking error prediction. TODO: add to the caption

Dependent Variables:	Tracking Error$_{i,t+1\to t+h}$			
	(1)	(2)	(3)	(4)
Detail$_{i,t}$	0.0001	-0.0011	0.0006	0.0005
	(0.0005)	(0.0008)	(0.0008)	(0.0010)
Total Length$_{i,t}$	0.0005	0.0016*	-0.0009	-0.0028**
	(0.0008)	(0.0010)	(0.0014)	(0.0014)
Horizon h (months)	6	12	24	36
Fund FEs	✓	✓	✓	✓
Year-month FEs	✓	✓	✓	✓
N	1,159	1,153	964	899
R^2	0.83	0.75	0.59	0.58

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

D Robustness Checks

This appendix contains extended and alternative specifications to check the robustness of the paper's results.

D.1 Fund-Level Timeseries Variation in Communication

Table IA.12 analyzes timeseries variation at an individual fund level, rather than at an index level, as in the main text. The VIX continues to be primary variable associated with fund-level communication measures.

Table IA.12: Timeseries variation in communication about future risks contained in S&P 500 index fund letters. This table displays the results of panel regressions of the level of detail of communication (columns 1-2) about forward-looking risk-related terms and the sentiment expressed therein (columns 3-4) against the VIX, Brenner and Izhakian's (2018) measure of market-level ambiguity, aggregate expected returns (proxied by the AAII survey employed by Greenwood and Shleifer (2014)) and past realized returns. The sample consists of S&P 500 index funds only.

Dependent Variables:	Log word count$_{i,t}$		Sentiment$_{i,t}$ (%)	
	(1)	(2)	(3)	(4)
Log VIX$_t$	0.1982**	0.1845**	-0.6176*	-0.6674**
	(0.0855)	(0.0889)	(0.3244)	(0.3321)
Log Ambiguity$_t$	0.0390	0.0477	0.2058	0.2129
	(0.0427)	(0.0444)	(0.1547)	(0.1581)
AAII Expected Return$_t$ (%)		-0.0038		-0.0242***
		(0.0024)		(0.0088)
Return$_{i,t-6\to t-1}$ (%)		0.0041		0.0394***
		(0.0030)		(0.0117)
Fund FEs	✓	✓	✓	✓
N	1,662	1,662	1,662	1,662
R^2	0.63	0.63	0.24	0.25

Clustered (Fund) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

D.2 Measuring Information Spillovers Over an Extended Sample

Table IA.13 extends the sample of fund family cooperation indices and verifies the baseline results on information spillovers continue to hold.

Table IA.13: Information spillovers within fund families, using forward-filled cooperation index values. The regressions in this table document a high correlation between the level of detail about forward-looking risks contained in S&P 500 index fund letters (dependent variable) written by fund i at time t, and the corresponding level of detail across active domestic equity sibling funds (only) within the same fund family, aggregated as a simple mean. *Fund controls* (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. Columns (1) & (3) show the unconditional strength of this relationship, across different fixed effects specifications. Columns (2) & (4) additionally include an interaction with an index for the strength of cooperation between sibling funds in the focal fund's overall family, as measured by Evans, Prado, and Zambrana (2020) – since their dataset ends in December 2015, the last observation for each family has been forward-filled to the end of my sample, where possible.

Dependent Variable:	Log word count$_{i,t}$			
	(1)	(2)	(3)	(4)
Sibling Mean Log word count$_{i,t}$	0.8492***	0.5969***	0.8120***	0.6595***
	(0.0334)	(0.0740)	(0.0405)	(0.0930)
Sibling Mean Log word count$_{i,t}$		0.3953***		0.3321**
× Family Cooperation Index$_{i,t}$		(0.1099)		(0.1465)
Family Cooperation Index$_{i,t}$		-2.101***		-1.540*
		(0.6582)		(0.7885)
Fund Controls$_{i,t}$	✓	✓	✓	✓
Year-month FEs	✓	✓		
Fund FEs			✓	✓
N	1,155	1,083	1,155	1,083
R^2	0.94	0.93	0.95	0.93
F-stat	289.5	114.6	335.2	125.7

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

D.3 Alternative Measure of the Level of Detail Communicated

Table IA.14 provides evidence on the effect of the level of detail communicated on various fund flows and under various fixed effect specifications, using an alternative measure of the level of detail communicated: the fraction of the fund's letter devoted to forward-looking risk-related communication.

The estimated directions of the effect remain unchanged when compared to the baseline regressions in the main paper. The interpretation of the magnitude of the effect differs.

Table IA.14: Effect of the fractions of fund letters devoted to forward-looking risks on S&P 500 index fund flows. The main covariate measures the level of detail contained in the letter written by fund i at time t about forward-looking risks, using the fraction of words devoted to that topic in the whole letter. Controls (not shown) are for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. The sample consists of S&P 500 index funds only. Each flow variable is the next month's flow, in percentage points of TNA..

Dependent Variables:	Net flow	Inflow	Outflow
	(1)	(2)	(3)
Fraction of letter$_{i,t}$ (%)	0.0282**	0.0268*	1.87×10^{-5}
	(0.0123)	(0.0145)	(0.0038)
Fund Controls$_{i,t}$	✓	✓	✓
Year-month FEs	✓	✓	✓
Fund FEs	✓	✓	✓
N	1,253	1,253	1,253
R^2	0.41	0.78	0.85

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

D.4 Further Tests to Rule Out Within-Family Rebalancing

Tables IA.15 and IA.16 split the baseline sample of active domestic equity funds into funds taking above-market and below-market risks, respectively, when testing whether within-family rebalancing can explain the paper's main effects. This classification is based funds' rolling CAPM β^{MKT} exposures.

Table IA.15: Testing for communication-driven rebalancing flows out of high-risk active funds. This table shows the relationship between outflows from active domestic equity funds and the level of detail about forward-looking risk-related statements contained in the letters of their sibling S&P index tracker funds (columns 1-2) or in their own letters (columns 3-4). The level of detail is measured as the log of the number of words devoted to forward-looking risk-related statements. Controls (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. The sample consists of active domestic equity funds with CAPM $\beta^{MKT} > 1$ and that belong to fund families that also contain an S&P 500 index tracker fund.

Dependent Variable:	Sibling Outflow$_{i,t\to t+1}$ (%)			
	(1)	(2)	(3)	(4)
S&P 500 tracker Detail$_{i,t}$	-0.0635	-0.1130		
	(0.2706)	(0.1249)		
Sibling Detail$_{i,t}$			-0.0645	-0.1759
			(0.2594)	(0.1309)
Controls$_{i,t}$	✓	✓	✓	✓
Year-month FEs	✓		✓	
Fund FEs		✓		✓
N	4,502	4,502	4,502	4,502
R^2	0.40	0.79	0.40	0.79

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table IA.16: Testing for communication-driven rebalancing flows out of low-risk active funds. This table shows the relationship between outflows from active domestic equity funds and the level of detail about forward-looking risk-related statements contained in the letters of their sibling S&P index tracker funds (columns 1-2) or in their own letters (columns 3-4). The level of detail is measured as the log of the number of words devoted to forward-looking risk-related statements. Controls (not shown) are for each fund letter's total (log) word count, for the fund's prior month returns, fees & (log) size, and for the overall fund family's (log) age and (log) size. The sample consists of active domestic equity funds with CAPM $\beta^{MKT} < 1$ and that belong to fund families that also contain an S&P 500 index tracker fund.

Dependent Variable:	Sibling Outflow$_{i,t\to t+1}$ (%)			
	(1)	(2)	(3)	(4)
S&P 500 tracker Detail$_{i,t}$	-0.1756	-0.1053		
	(0.1858)	(0.1247)		
Sibling Detail$_{i,t}$			-0.1019	-0.0854
			(0.2061)	(0.1135)
Controls$_{i,t}$	✓	✓	✓	✓
Year-month FEs	✓		✓	
Fund FEs		✓		✓
N	4,621	4,621	4,621	4,621
R^2	0.38	0.81	0.37	0.81

Clustered (Fund & Year-month) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

E Regular Expressions to Replicate Textual Measures

This appendix provides the regular expressions used to detect fund letters and to classify individual statements. I have two goals in doing so. First, I aim to facilitate the replication and extension of this work. Second, I hope to illustrate the simplicity and transparency of this approach, which requires no estimation.

Internet Appendix A.1 validates these manually-defined measures by comparing them to other approaches.

E.1 Extracting Fund Letters From Reports

I provide two regular expressions that can be applied to the extracted text of N-CSR(S) shareholder reports. The separate regular expressions may detect overlapping text (for example if one is subsumed by the other in the report), so the detected text output of each should be merged appropriately.

The first regular expression detects letter portions that are formatted as fund letters:

```
'(?:(?:Dear|Fellow|To ?(?:a|A)ll|To ?(?:a|A)ll ?(?:o|O)ur|To ?(?:o|O)ur)(?:
 ?(?:m|M)ost)?(?: ?(?:(?:a|A)ppreciated|(?:c|C)herished|(?:f|F)ellow|(?:g|G
)reat(?:est)?|(?:h|H)onored|(?:e|E)steemed|(?:p|P)rized|(?:r|R)espected|(?:
v|V)alued))? ?(?:(?:c|C)lient|(?:c|C)ontributor|(?:c|C)ustomer|(?:i|I)
nvestor|(?:o|O)wner|(?:p|P)artner|(?:s|S)hareholder|(?:s|S)tockholder|(?:s|
S)ubscriber|(?:s|S)ponsor|(?:u|U)nit.?(?:h|H)older)s?(?:,|\\:|\\n)?|(?:DEAR
|FELLOW|TO ?ALL|TO ?ALL ?OUR|TO ?OUR)(?: ?MOST)?(?: ?(?:APPRECIATED|
CHERISHED|FELLOW|GREAT(?:EST)?|HONORED|ESTEEMED|PRIZED|RESPECTED|VALUED))?
?(?:CLIENT|CONTRIBUTOR|CUSTOMER|INVESTOR|OWNER|PARTNER|SHAREHOLDER|
STOCKHOLDER|SUBSCRIBER|SPONSOR|UNIT.?HOLDER)S?(?:,|\\:|\\n)?|(?:Dear|Fellow
|To)(?:.{0,100})(?:(?:c|C)lient|(?:c|C)ontributor|(?:c|C)ustomer|(?:i|I)
nvestor|(?:o|O)wner|(?:p|P)artner|(?:s|S)hareholder|(?:s|S)tockholder|(?:s|
S)ubscriber|(?:s|S)ponsor|(?:u|U)nit.?(?:h|H)older)s?(?:,|\\:|\\n)|(?:DEAR|
FELLOW|TO)(?:.{0,100})(?:CLIENT|CONTRIBUTOR|CUSTOMER|INVESTOR|OWNER|PARTNER
|SHAREHOLDER|STOCKHOLDER|SUBSCRIBER|SPONSOR|UNIT.?HOLDER)S?(?:,|\\:|\\n)|
Welcome,|LETTER|Letter from|PHOTO|Photo|MESSAGE|Message from|CHATTER|
Chatter)(.*?)(?:(?:(?:Yours ?(?:(?:c|C)ordially|(?:f|F)aithfully|(?:r|R)
espectfully|(?:s|S)incerely|(?:t|T)ruly))|(?:(?:In|With) ?(?:(?:a|A)ll ?)
?(?:(?:o|O)ur ?)?(?:(?:a|A)ppreciation|(?:g|G)ratitude|(?:p|P)artnership
|(?:r|R)espect))|(?:(?:(?:b|B)est|(?:k|K)ind|(?:w|W)arm)(?:est)? ?(?:(?:r|R
)egards|(?:w|W)ishes))|Many ?(?:t|T)hanks|Stay ?(?:(?:s|S)afe|(?:h|H)ealthy
)|Until ?(?:n|N)ext ?(?:t|T)ime|(?:(?:(?:Very ?)?(?:(?:a|A)ppreciatively
|(?:c|C)ordially|(?:f|F)aithfully|(?:r|R)espectfully|(?:t|T)ruly|(?:s|S)
incerely))(?: ?(?:(?:(?:y|Y)ours)|(?:(?:s|S)ubmitted)))?)(?:,|\\n))|(?:(?:
```

```
Best|Yours|Truly|Sincerely|Respectfully|Cordially)(?:,|\\n))|CFA|PhD|Chief|
CHIEF|Officer|OFFICER|Founder|FOUNDER|PRESIDENT|CHAIRMAN|Head ?of|HEAD ?OF|
Chairman ?of|CHAIRMAN OF|Investor ?Relations|INVESTOR ?RELATIONS|Managing ?
Director|MANAGING ?DIRECTOR|Managing ?Partner|MANAGING ?PARTNER|Registered
?(?:r|R)epresentative|REGISTERED ?REPRESENTATIVE|UNDERSTANDING ?YOUR|
ANALYZING ?YOUR|VALUE ?OF|TOTAL ?RETURNS|FEE|SUMMARY|SCHEDULE|STATEMENT|
INVESTING|BENCHMARK|COMPARISON|DISCLOSURE|(?:p|P)rospectus|PROSPECTUS|
Important ?(?:i|I)nformation|IMPORTANT ?INFORMATION|Past ?performance|PAST
?PERFORMANCE|does not constitute)'
```

The second regular expression detects letter portions that are formatted as a section in the shareholder report:

```
'(?:(?:manage(?:ment|r).?s?.? ?)?discussion ?(?:(?:and ?|\\& ?|\\+ ?)
analysis ?)?of ?(?:(?:fund|portfolio) ?)?performance|manage(?:ment|r).?s?.?
 ?discussion ?(?:and ?|\\& ?|\\+ ?)analysis|(?:(?:fund|portfolio) ?)?
performance ?(?:review|discussion|analysis)|m ?a ?n ?a ?g ?e ?(?:m ?e ?n ?t
|r ?).{0,10}(?:a ?n ?a ?l ?y ?s ?i ?s|d ?i ?s ?c ?u ?s ?s ?i ?o ?n|r ?e ?v
?i ?e ?w)|(?:(?:fund|portfolio) ?)?(?:management|manager|managers|manager.s
|managers.) ?(?:analysis|discussion|review)|(?:questions? ?(?:and|\\&|\\+)
?answers?|q ?(?:and|\\&|\\+) ?a|interviews?)(?: ?with)?|markets? ?(?:
commentar|perspective|outlook|view)|(?:portfolio|fund)? ?:((?:management|
manager|managers|manager.s|managers.) ?)?(?:commentar|perspective|outlook|
view))(.*?)(?:(?:(?:Yours ?(?:(?:c|C)ordially|(?:f|F)aithfully|(?:r|R)
espectfully|(?:s|S)incerely|(?:t|T)ruly))|(?:(?:In|With) ?(?:(?:a|A)ll ?)
?(?:(?:o|O)ur ?)?(?:(?:a|A)ppreciation|(?:g|G)ratitude|(?:p|P)artnership
|(?:r|R)espect))|(?:(?:(?:b|B)est|(?:k|K)ind|(?:w|W)arm)(?:est)? ?(?:(?:r|R
)egards|(?:w|W)ishes))|Many ?(?:t|T)hanks|Stay ?(?:(?:s|S)afe|(?:h|H)ealthy
)|Until ?(?:n|N)ext ?(?:t|T)ime|(?:(?:(?:Very ?)?(?:(?:a|A)ppreciatively
|(?:c|C)ordially|(?:f|F)aithfully|(?:r|R)espectfully|(?:t|T)ruly|(?:s|S)
incerely))(?: ?(?:(?:(?:y|Y)ours)|(?:(?:s|S)ubmitted)))?)(?:,|\\n))|(?:(?:
Best|Yours|Truly|Sincerely|Respectfully|Cordially)(?:,|\\n))|u ?n ?d ?e ?r
?s ?t ?a ?n ?d ?i ?n ?g ? ?y ?o ?u ?r|c ?o ?m ?p ?a ?r ?i ?n ?g ? ?y ?o ?u
?r|p ?o ?r ?t ?f ?o ?l ?i ?o ? ?s ?u ?m ?m ?a ?r|p ?e ?r ?f ?o ?r ?m ?a ?n
?c ?e ? ?s ?u ?m ?m ?a ?r|s ?c ?h ?e ?d ?u ?l ?e ? ?o ?f|f ?e ?e ? ?s ?c ?h
 ?e ?d ?u ?l ?e|s ?t ?a ?t ?e ?m ?e ?n ?t ? ?o ?f|f ?e ?e ? ?s ?t ?a ?t ?e
?m ?e ?n ?t|v ?a ?l ?u ?e ? ?o ?f ? ?p ?o ?r ?t ?f ?o ?l ?i ?o|t ?o ?t ?a ?
l ? ?r ?e ?t ?u ?r ?n ?s|e ?x ?p ?e ?n ?s ?e ?s ? ?p ?a ?i ?d|c ?o ?m ?p ?a
 ?r ?i ?s ?o ?n ? ?o ?f|u ?n ?d ?e ?r ?s ?t ?a ?n ?d ?i ?n ?g ? ?y ?o ?u ?r
|a ?n ?a ?l ?y ?z ?i ?n ?g ? ?y ?o ?u ?r|f ?u ?n ?d ? ?e ?x ?p ?e ?n ?s ?e|
g ?l ?o ?s ?s ?a ?r ?y|n ?o ?t ?e ?s|d ?i ?s ?c ?l ?o ?s ?u ?r ?e|d ?o ?e ?
s ? ?n ?o ?t ? ?c ?o ?n ?s ?t ?i ?t ?u ?t ?e|p ?l ?e ?a ?s ?e ? ?n ?o ?t ?e
```

```
|i ?m ?p ?o ?r ?t ?a ?n ?t ? ?i ?n ?f ?o ?r ?m ?a ?t ?i ?o ?n|f ?o ?r ?e ?g
 ?o ?i ?n ?g ? ?i ?n ?f ?o ?r ?m ?a ?t ?i ?o ?n|integral ?part|return ?
includes|registered ?under|past ?performance|investment ?company ?act
|----|====)'
```

E.2 Classifying Sentences Within Fund Letters

The following regular expressions should be applied to individual sentences. Sentences are defined as grammatically-valid phrases, as detected using the open-source StanfordNLP toolkit (Manning et al., 2014).

Detecting Boilerplate Sentences

```
'40 act|\\.com|accompanying (note|statement)|account(ed|ing)|accrual|(
additional|detailed|further) information|(advised to|please) (contact|
consult|email|visit)|admin|adhere to|agent|board|officer|treasurer|trustee|
amortiz|(after|before|pre|post)-?tax|appendix|arrear|audit|accordance with|
accretion|additional charge|advisory|affiliat|aggregate cost|amount is|
annualize|arrangement|as follows|attached|benchmark|(beyond|outside) our
control|board considered|board of directors|books|calculated (based|by|on|
using)|capital contribution|cash equivalent|charged to|class shares are (
offered|sold) at|(cease|commence)d?( investment)? (operation|trading)|
collateral|committee|compliance|consolidated|contained herein|(contact|
consult|email|visit) y?our|contingent|contractual|control over|constitute|
correspondence|course of business|custodian|deferred|deem|denominated|
depreciation|derivative|determined in accordance|differ|disclosure|
distribut|effective rate|enter(ing)? into|examples? (above|below|is|are)|
exceeds the value|exempt from|exercise price|fair valu|filed|(feeder|master
) (fund|portfolio)|for example|foregoing|gaap|(gross|net|total) change in|
if any|in determining the|index is (a |an |comprised|based|designed|managed
|offers|unmanaged|provides)|include.? but are not limited to|illiquid asset
|in writing|incur|independent (consultant|director)|indemni(t|f)|
information (is provided on an|is not|on|presented|reflects|shown)|integral
 part of|input|invest up to|investment company act|involve risk|is
calculated|is shown in|lease|left blank|liabilit|lipper|may buy and sell|
may differ from|may (include|involve|result)|(may|should) not be (
considered|construed|relied|depended|taken)|may (also )?(invest in|hold)|
maximum (exposure|commitment)|means? the?|meaningful|methodology|
morningstar|must be (accompanied|preceded)|no significant change|not (
eliminate|indicative|intended|meant|reflected|warranted)|not? (assurance|
guarantee)s? (of|that|the)|not be relied|notes to|obligat|offered at|
```

```
organized as|payable|potential for gain|proprietary to|principal (loss|risk
)|purchases? (and|or) sale|principally traded|recognized on|records|redeem|
redemption|refers to|registered under|reimburse|repurchase|representative
of|required to|represents the|respectively|responsibility of|restricted
security|restrictions on|rights reserved|risk of loss|rounds to|schedules?
of|sales (charge|expense)|see (accompanying|note|the|page)|seeks to|settle|
share class|stated as|statements? of|structured as|subject to|(such|these|
valuation) models|tax.exempt|tax (advantage|benefit|expense|jurisdiction|
position|purpose|return)|thank you|the adviser|
the funds? (do|did|had no|have no|invest in|may|offer|value)|total
investment|transfer|unanticipated change|under (normal|typical) (circum|
condition|course of|market)|underlying security|up.to.date|valued (using|
based|by)|views expressed|waiver|we welcome your|when applicable|will seek
to|www\\.'
```

Detecting Forward-Looking Sentences

```
'anticipat|belie(f|ve)|conceiv|conjecture|consider|diagnos|expect|foresee|
estimat|guess|hope|interpret|look|plan|project|predict|think|\\b(develop|
unfold|regard|see|view)\\b|we (continue|remain|favou?r)|horizon|timeline|(
short|medium|near|long).term|(over the|next|coming|forthcoming|future|
subsequent) (period|quarter|year|month|meeting|date|release)s?|ahead|
attitude|forward|future|going|moving|opinion|optimis|outlook|perspectiv|
pessimis|prognosis|realis|stance|chance|danger|indicat|hazard|likelihood|
risk|possib|probab|prospect|signal|spell|threat|warn|(can|cannot|can't|will
|won't|not|should|shouldn't|would|wouldn't|to) (affect|arise|change|cascade
|cause|climb|come (about|off|to)|continue|crop up|(de|in)crease|(de|en)
velop|enlarge|emerge|ensue|expand|fall|feed|follow|guide|grow|happen|hold|
lead|lengthen|materiali|modify|occur|pan out|release|remain|pass|persist|
push|raise|rise|spill|stretch|surface|take place|transpire|turn out|widen|
update)|(ready|(position|prepar)(ed|ing)) (for|to)|bullish|bearish|(over|
under).?valu|cheap|expensive|richly'
```

Detecting Risk-Related Sentences

```
'alarm|alert|capitulat|catastroph|correction|choppy|collaps|crash|crisis|
danger|disaster|disorder|expose|forebod|hang over|impend|insecurity|
instability|jeopard|hazard|menace|overhang|precarious|rare event|risk|
uncertain|unpredictable|unstable|storm|stress|squall|systemic|threat|
tempest|trouble|tumult|turbulen|uneas|warn|wary|wild|worr|volatil|\\b(loom|
rough|tail)\\b|afraid|anxiety|anxious|apprehensi|caution|cautious|
```

```
consternation|daunt|dismay|distress|doom|fear|fright|gloom|nervous|rattle|
shake|shook|terror|trepid|defended|defenceless|immune to|in the firing line
|guarded|mercy of|open to attack|prey to|powerless|prone to|protected|safe|
sensitive to|shielded|vulnerable|wide open to|(black|white) swan|peso
problem'
```

Detecting Performance- and Return-Related Sentences

```
'performance|to.date|\\bytd\\b|((our|the (etf|fund|portfolio|trust|share
class)).*?(achieve|accrue|affected|assisted|cumulat|earned|experienced|gain
|fail|fell|helped|hit|hurt|impacted|loss|lost|perform|realize|return|showed
|success|succeed|strong|support|weak|win|won))'
```

References in Internet Appendix

Berk, J. B., & Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, *112*(6), 1269–1295.

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). *The development and psychometric properties of LIWC-22*.

Brenner, M., & Izhakian, Y. (2018). Asset pricing and ambiguity: Empirical evidence. *Journal of Financial Economics*, *130*(3), 503–531.

Caetano, C., Caetano, G., Nielsen, E., & Sanfelice, V. (2023). The effect of maternal labor supply on children: Evidence from bunching.

Caetano, C., Caetano, G., & Nielsen, E. R. (forthcoming). Correcting for endogeneity in models with bunching.

Chevalier, J., & Ellison, G. (1997). Risk taking by mutual funds as a response to incentives. *Journal of Political Economy*, *105*(6), 1167–1200.

Cremers, K. M., & Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies*, *22*(9), 3329–3365.

Evans, R. B., Prado, M. P., & Zambrana, R. (2020). Competition and cooperation in mutual fund families. *Journal of Financial Economics*, *136*(1), 168–188.

Gennaioli, G. N., Shleifer, A., & Vishny, R. (2015). Money doctors. *Journal of Finance*, *70*(1), 91–114.

Greenwood, R., & Shleifer, A. (2014). Expectations of returns and expected returns. *Review of Financial Studies*, *27*(3), 714–746.

Grice, R., & Gueciouer, A. (2023). Mutual fund market structure and company fee competition: Theory and evidence. *Available at SSRN*, *4449026*.

Grinold, R. C., & Kahn, R. N. (1999). *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk* (2nd ed.). McGraw Hill.

Hassan, T. A., Hollander, S., van Lent, L., & Tahoun, A. (2019). Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics*, *134*(4), 2135–2202.

Hillert, A., Niessen-Ruenzi, A., & Ruenzi, S. (2021). Mutual fund shareholder letters: Flows, performance, and managerial behavior. *Available at SSRN*, *2524610*.

Hortaçsu, A., & Syverson, C. (2004). Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds. *Quarterly Journal of Economics*, *119*(2), 403–456.

Kim, D., Wang, Q., & Wang, X. (2022). Geographic clustering of institutional investors. *Journal of Financial Economics*, *144*(2), 547–570.

Lee, D. S., McCrary, J., Moreira, M. J., & Porter, J. (2022). Valid t-ratio inference for iv. *American Economic Review*, *112*(10), 3260–3290.

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.

Lusardi, A., Bumcrot, C. B., & Lin, J. (2013). The geography of financial literacy. *Numeracy*, *6*.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & McClosky, D. The Stanford CoreNLP natural language processing toolkit. In: In *Association for computational linguistics (acl) system demonstrations*. 2014, 55–60.

Mullainathan, S., & Shleifer, A. (2005). Persuasion in finance.

Schlag, C., Thimme, J., & Weber, R. (2021). Implied volatility duration: A measure for the timing of uncertainty resolution. *Journal of Financial Economics*, *140*(1), 127–144.