Deep Learning Fund Managers' Narratives:

Risk Assessment and Fund Performance*

Sean Cao
Robert H. Smith School of Business
University of Maryland, College Park
scao824@umd.edu

Baozhong Yang
J. Mack Robinson College of Business
Georgia State University
bzyang@gsu.edu

Alan L. Zhang
College of Business
Florida International University
alan.zhang@fiu.edu

This Version: September 2022

First Version: December 2019

* The authors have benefited from discussions with Asaf Manela, Lingling Zheng, and Dexin Zhou, and comments and suggestions from conference and seminar participants at 2021 CICF, 2021 China International Risk Forum, 2022 China FinTech Research Conference, University of Maryland, Georgia State University, Xiamen University and HKUST.

Deep Learning Fund Managers' Narratives:

Risk Assessment and Fund Performance

Abstract

We use a deep learning model to extract syntax and context information from mutual fund managers' narrative discussions and measure their risk assessment. We validate the forward-looking nature of the risk measure by showing that more negative (positive) risk assessment in managers' narratives leads to a reduction (increase) of portfolio risk in the subsequent period. The forward-looking risk assessment measure also reflects managerial skills: managers who are conscious of negative risk generate superior risk-adjusted returns and higher Sharpe ratios and are more likely to have higher intra-quarter trading skills and higher Morningstar ratings. Interestingly, not all market participants respond to this narrative-based measure except sophisticated investors. The forward-looking nature of our new measure can thus inform investors and researchers about fund managers' risk management and performance.

JEL Classification: C45, G11, G14, G23

Key Words: Deep Learning, Machine Learning, Mutual Fund, Disclosure, Textual Analysis, Sentiment, Risk Assessment, Fund Performance, Fund Skill

1. Introduction

In 2021, mutual funds in the U.S. manage about $34 trillion worth of assets. Many investors have their pension plans and life savings invested in mutual funds. Studying the risk-taking behavior of mutual funds is thus of prime importance for investors to understand the risk in their portfolios. For fund managers, the ability to correctly evaluate and manage risk is directly relevant to their goal of achieving superior returns. Regulators are also interested in fund managers' risk-taking behavior to prevent excessive risk taking, which may exacerbate market volatility during crisis periods. The literature has long studied the risk-taking behavior of mutual fund managers to better inform investors and regulators (Brown, Harlow, and Starks, 1996; Chevalier and Ellison, 1997; Pool, Stoffman, Yonker, and Zhang, 2018; Ma and Tang, 2019). Despite the importance of this subject, little is known about fund managers' forward-looking assessment of risk and how they act upon their opinions. The challenge lies largely in limitations in data, as it is challenging to identify how mutual fund managers contemplate their future risk management and investment plans.

In this study, we use deep learning to extract information relevant to risk management from mutual fund managers' portfolio discussion in their mandatory filings to SEC (i.e., shareholder reports in N-CSR/N-CSRS filings). This information allows us to measure managers' assessment of *forward-looking risk*. We study the following questions: 1) Can we use deep learning on managers' language to capture their risk assessment that reflects their future risk taking? 2) Is forward-looking risk assessment associated with fund skill and superior performance? 3) Do institutional or retail investors respond to managers' narrative risk assessment?

Conventional risk-taking measures are calculated based on historical numerical data of returns or holdings. Despite being easy to process and analyze, numerical data do not contain forward-looking information of managers' risk assessment. In contrast, while managers' qualitative discussion of portfolio decisions in mandatory SEC filings (shareholder reports) convey rich information, it is challenging to process such textual data due to their unstructured and high-dimensional nature. Managers can discuss their risk assessment in a variety of linguistic styles, making it difficult to identify patterns. Consider a hypothetical discussion: "Risk is beneficial, but our flow is constrained in the next period." In this example, the traditional "bag-of-words" approach of counting the appearances of the keyword "risk" does not indicate the manager's assessment regarding risk. One can also devise a rule-based approach (similar to Hassan, Hollander, van Lent, and Tahoun, 2019), e.g., counting words with a positive or negative tone (Loughran and McDonald, 2011) within ten words of "risk." Both "beneficial" and "constrained" are captured. Hence, this rule-based risk measure will generate a neutral risk assessment (one positive word and one negative word). However, the true interpretation of the sentence is that it expresses a positive risk assessment given that "constrained" is used to modify "flow" instead of "risk." Another limitation of the traditional approach is that it cannot capture double negations such as in "not averse to risk," which actually carries a positive tone.

Textual information consists of two dimensions: 1) lexical meanings of words; 2) syntactical interactions among words. The bag-of-words approach does not work well above because it only captures the meanings of words but ignores the syntactical structures (Loughran and McDonald, 2016). To tackle this issue, in this study we apply state-of-the-art deep learning models for natural language processing developed by Chen and Manning (2014) to parse texts and extract syntactical relations among words.

We employ the above deep learning model to parse all textual contents in mutual fund shareholder reports and construct directional risk assessment measures. In particular, we identify and count all instances of risk-related keywords syntactically modified by a word with positive or negative assessment. We first validate these measures by exploring their predictability of forward-looking risk taking after controlling for current risk taking. We find negative (positive) risk assessment strongly predicts managers' reduction (increase) of their risk taking in the subsequent period. We also compare our deep-learning-based measure with the traditional bag-of-word measures that simply count the occurrences of risk-related keywords. Our measure yields superior results than the bag-of-words measures in predicting the future risk-taking behavior of fund managers.

We further find that fund managers who are conscious about negative risk, i.e., having a negative risk assessment, are able to generate higher future fund performance, measured by Fama-French Carhart four-factor alpha (Fama and French, 1993; Carhart, 1997) and Sharpe ratio (Sharpe, 1994), controlling for past performance, risk and other fund characteristics. On the other hand, we do not find any predictability in subsequent performance from positive risk assessment. We further test whether risk-consciousness reflects managerial skill. Following Kacperczyk, Sialm, and Zheng (2008), we use the return gap of mutual funds to capture managers' intra-quarter trading skill. We find that funds with negative risk assessment are also more likely to have higher future intra-quarter trading skill, after controlling for past return gap.

The asymmetric effects of risk assessment on performance may arise for the following reasons. First, firms tend to withhold negative news and present positive tones in their disclosure (Loughran and McDonald, 2011; Kothari, Shu, and Wysocki, 2009). Fund managers thus need to possess skills and expend extra efforts to acquire and comprehend negative news. Second, positive risk assessment can reflect managers' overconfidence. Overconfident fund managers are likely to overinvest in positive news and take more risk (Palomino and Sadrieh, 2011) while they cannot short sell based on bad news. Overconfident managers are also more likely to hold overvalued stocks, resulting in lower subsequent performance (Adebambo and Yan, 2018). Overall, managers with negative risk assessment are likely to be skilled in identifying and parsing negative news, and are less subject overconfidence biases, both consistent with their superior performance.

Next, we examine the potential mechanism through which fund managers with negative risk assessment exhibit skill. One possibility is that these managers are able to anticipate future market uncertainty or declines and avoid investing in risky stocks in such episodes. To test this hypothesis, we consider how funds manage their downside beta, which measures funds' exposure to downside risk of the market. We find that managers actively reduce their exposure to downside risk when they have negative risk assessment. Furthermore, the reduction of downside beta is stronger when future market risk premium is negative, suggesting that funds anticipate future market conditions. These findings suggest that fund managers with negative risk assessments are those who see the "dark clouds in the sky" and thus manage their portfolio with precaution.

Since there are no fixed templates for the textual discussions in the shareholder reports, managers enjoy a large degree of freedom in what to convey to investors. We next explore if their risk assessment receives responses from the investment community. We find that risk-conscious managers who report negative risk assessment receive higher Morningstar ratings in the future. The increased Morningstar rating is primarily driven by the managers' future reduction in risk taking, suggesting that the market/rating agency rewards managers who possess and share a prudent view of risk with higher ratings.

Along this vein, we investigate whether our negative risk assessment measure provides any implications yet to be discovered by investors. Our results indicate that only sophisticated investors recognize managers' risk assessment skill and reward funds with greater capital flows when fund managers disclose negative risk assessment. Because only a group of investors can observe the public signals from mutual funds, it accentuates the uniqueness and usefulness to investors, particularly retail investors and less sophisticated investors.

Our findings contribute to several strands of literature. First, we contribute to the literature on mutual fund risk taking (e.g., Brown, Harlow, and Starks, 1996; Chevalier and Ellison, 1997; Huang, Sialm, and Zhang, 2011; Pool, Stoffman, Yonker, and Zhang, 2018; Ma and Tang, 2019) by constructing a forward-looking risk assessment measure and studying its implications for future risk taking and fund performance. The empirical evidence shows that not all market participants are responding to this narrative-based measure except sophisticated investors. Our new risk assessment measure provides an inside look at fund managers' reasoning process and allows investors and researchers to better understand and predict the risk management and investment decisions of fund managers.

Second, we contribute to the growing literature on textual analysis by introducing new deep learning methods. Textual analysis has generated many fruitful applications in finance, economics, and accounting (see, for example, Loughran and McDonald, 2016; Gentzkow, Kelly, and Taddy, 2019). However, traditional bag-of-word approaches used in much of the literature only capture lexical features of single words. Despite the success of the traditional approach, the context, the order or sequence of words, and the relations among words are lost in the process. Some recent studies use more recent developments in machine learning and natural language processing for textual analysis (e.g., Li, Mai, Shen, and Yan, 2019; Cong, Liang, Yang, and Zhang, 2020; Zhang, 2021; Abis, 2022; Cao, Jiang, Yang, and Zhang, 2022). Our study employs a deep neural network that helps capture the higher-order textual features. We show that such interactive syntactical features can precisely identify managers' tones specific to risk topics. Our method also has the benefit of being perfectly interpretable (despite utilizing deep neural networks) since the syntactic relations are derived from grammar and based on human understanding of texts. In general, this method can facilitate researchers in studying other complicated interactive relationships among words in future studies.

Our study is also related to several contemporaneous papers on mutual fund disclosures. Several studies focus on another type of mutual fund disclosure, prospectus. Abis (2022) and Abis and Lines (2022) apply random forest and k-means clustering on mutual fund prospectuses to categorize mutual funds, and Abis, Buffa, Javadekar, and Lines (2022) study how mutual funds strategically disclose due to investors learning. Sheng, Xu, and Zheng (2022) examine funds' incentive to disclose risk in summary prospectus. Distinct from these studies, we measure managers' risk assessment using their narratives in the more frequently disclosed shareholder reports, which provide more freedom for managers to express their views. Du, Jiao, Ye, and Fan (2020) and Hillert, Niessen-Ruenzi, and Ruenzi (2021) also use shareholder reports to study how writing styles (e.g., confidence and tones) of mutual fund managers affect flows and performance. We differ from these two studies by using state-of-the-art deep learning models to extract precise, qualitative information about managers' forward-looking risk sentiment, which reflects another dimension of their skill.

2. Parsing Syntax Relations from Texts

In this section, we describe the methodology we used to parse word dependencies. Given the technical nature of this procedure, we provide a general discussion and leave some details in the Appendix B.

The words in a sentence are organized based on grammatical rules and have complicated interdependency. One way to represent the inherent grammatical structure in a sentence is through a tree structure. For example, in the following diagram, the grammatical tree starts with the root word "moving" and grows the tree by adding words one by one according to their modifying role with a word already on the tree. The word "faster," for example, modifies "moving." This modifying relationship is identified by a directed arc (or arrow) going from "moving" to "faster." Further, the arrow is labeled as *advmod*, meaning that "faster" is an adverb modifier. We explain the dependency parsing procedure in more detail in Appendix B.

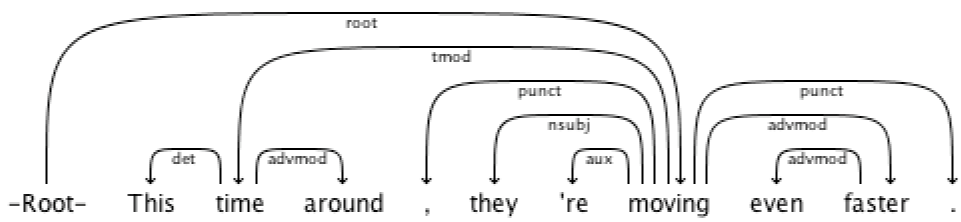


Figure 1. The tree representation of the grammatical structure of an example sentence (Source: Stanford NLP Website)

A major challenge in dependency parsing is the abundance of possible features and relations to consider. For example, there are usually tens of thousands of unique words and dozens of parts of speeches (POS) and labels. A parsing algorithm also needs to consider potential combinations and relations among pairs and triples of words, leading to millions of possible features. Extracting and analyzing all these features is very challenging and time-consuming. In a breakthrough paper, Chen and Manning (2014) develop a novel deep neural network algorithm to parse dependency trees of sentences with higher accuracy and efficiency than previous algorithms. We employ their parsing model and provide the description below.

First, we obtain two types of features for each word in a sentence: POS tags (from the Stanford POS tagger) and labels (from the Basic Stanford Parser). Each unique word is then represented as a vector $vec_w(w_i) \in \mathbb{R}^d$ with $d"N_w$, where N_w is the dimension of the dictionary, i.e., the number of unique words. This process is called *word embedding* and allows a dense representation of the sparse word vectors while maintaining many linguistic and semantic structures. For example, words that are close in meaning to each other are embedded into vectors that are close in the space $\mathbb{R}^d$. Further, semantic patterns often are translated into linear relations, e.g., $vec("Madrid") - vec("Spain") + vec("France")$ is close to $vec("Paris")$. One widely used word embedding model is the *word2vec* model developed by Google researchers (Mikolov, Chen, Corrado, and Dean, 2013). Several recent studies have applied word embedding techniques in financial economics, e.g., Li, Mai, Shen, and Yan (2018), Cong, Liang, and Zhang (2019), and Hanley and Hoberg (2019). In this study, word embedding is used as an intermediate yet essential step in the parsing model.

In addition, we represent POS tags and labels as d-dimensional vectors: each tag t_j and label l_k are mapped to vectors $vec_t(t_j), vec_l(l_k) \in \mathbb{R}^d$. Although POS tags and labels take values in relatively small discrete sets, there are semantic structures among the values, e.g., *NN* (singular noun) is close to *NNS* (plural noun) and *amod* (adjective modifier) is close to *num* (numeric modifier). Vector representations can help to capture these inherent structures.

The main parsing model is a three-layer feedforward neural network that predicts the next step (transition) in growing a dependency tree (see Appendix B for details about transitions). The first layer (the input layer) has d_{in} nodes and consists of the vector representations $(x^w, x^t, x^l) \in \mathbb{R}^{d \times d_{in}}$ of the words, POS tags, and labels that are used in the parsing step. The second layer (the hidden layer) has d_h nodes and a *cube activation function*,

$$h = \left(W_1^w x^w + W_1^t x^t + W_1^l x^l + b_1\right)^3$$

where W_1^w, W_1^t and W_1^l are matrices that transform the inputs to d_h-dimensional vectors and $b_1 \in \mathbb{R}^d$. The third layer (output layer) is a standard Softmax layer that has $|\mathcal{T}|$ nodes, where $|\mathcal{T}|$ is to the number of possible choices for the current transition and each node represents the probability of making a particular transition choice. The objective function minimizes cross-entropy loss for the probabilities with a l_2-regularization term,

$$L(\theta) = -\sum_{i=1}^{|\mathcal{T}|} \log(p_{t_i}) + \frac{\lambda}{2} \|\theta\|^2.$$

where θ is the set of parameters for the neural network.

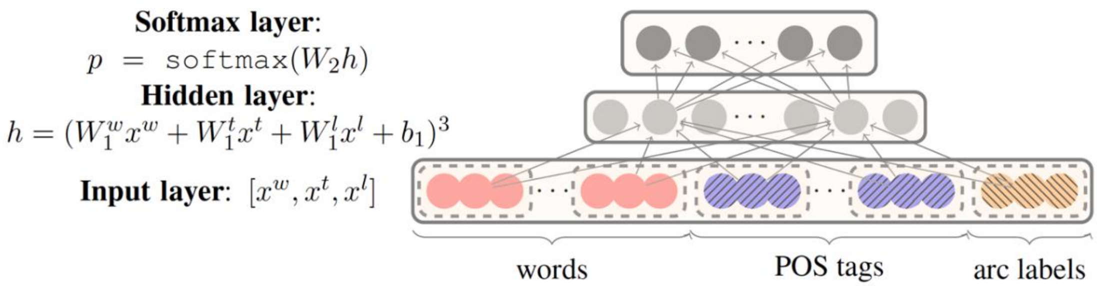


Figure 2. Neural network model for dependency parsing (Source: Chen and Manning (2014))

The cube activation function captures non-linear three-way interactions among words, POS tags, and labels, which are important in determining the dependency structure and therefore generates better performance than other commonly used activation functions (such as ReLU, logistic, or tanh functions).

The neural network is pre-trained on the English Penn Treebank database, which is a large database of English texts manually tagged and parsed by linguists. The database contains more than 3 million words from a wide range of sources such as Wall Street Journal articles, IBM computer manuals, nursing notes, and transcribed telephone conversations, etc. The model achieves a state-of-the-art out-of-sample parsing accuracy of 92%. We apply the trained model to all sentences in the shareholder reports in our sample to generate predicted dependency tree structures. We then extract from the parsed structures all pairs of words with a head word and a modifying word, together with POS tags and labels for the words.

3. Data and sample

3.1 Mutual fund shareholder reports

We retrieve mutual funds' shareholder reports from N-CSR (certified annual shareholder reports for management investment companies) and N-CSRS (certified semi-annual shareholder reports for management investment companies) filings on the SEC EDGAR website. A registered investment company (e.g., mutual fund companies) must electronically file Form N-CSR (We use N-CSR to represent both N-CSR and N-CSRS filings thereafter) to the SEC within 10 days of sending the corresponding reports to shareholders, unless it files for a hardship exemption. In general, an N-CSR filing includes the following items: a report to shareholders (Item 1), the company's code of ethics (Item 2), the names of the financial experts in the company's audit committee (Item 3), the disclosure of principal accountant fees and services for the previous two fiscal years (Item 4), the disclosure of listed registrants or reason for exemption from the audit committee (Item 5), the firm's security holdings (Item 6), and the disclosure of proxy voting policies (Item 7). In our analyses, we focus on Item 1 of N-CSR—the shareholder report.

3.2 Parsing N-CSR filings and constructing text-based variables

We download the plain text files for all N-CSR filings from EDGAR from 2006 to 2018 (see Section 3.3 for more description of the sample). A significant portion of an N-CSR filing's content consists of markup tags, ASCII-encoded graphics and tables, and other artifacts. Therefore, the complete filing tends to have a large size and can be computationally cumbersome to process. To the extent that our study focuses on the textual content of the document, we first use a computer program (in Python) to parse all filings and remove markup tags, graphics and tables. We then extract the contents of Item 1 from all filings. The above process substantially reduces the size of all N-CSR filings over our sample period from 239 gigabytes to 17.9 gigabytes (a reduction of more than 92.5%) and generates our input files for the next processing step. For each filing, we extract header information, including the Central Index Key (CIK, an investment company identifier), Series ID (a fund identifier), ticker, and the filing date.

We define *Length* as the number of words in each parsed N-CSR document after the above processing steps. In our study, we need to measure risk and tone separately. We first create a Risk dictionary of risk-related words (387 words) by combining the list of synonyms of "risk" in Hassan, Hollander, Lend, and Tahoun (2019) and the Loughran-McDonald ("LM" hereafter) dictionary of risk words. For the reader's convenience, we list all words in our Risk dictionary that appear in our sample in Table A1. We use the LM dictionaries of positive and negative words (2,355 words and 354 words, respectively) to measure word tones.[1]

Next, we construct pairs of dependent words using the neural-network dependency parsing model described in Section 2. Specifically, the neural network parser extracts a list of parsed dependency pairs from each sentence in an N-CSR document. Each pair consists of a head word and a modifying word. Since we focus on the tones in the managerial discussion of risk faced by the funds, we keep only pairs for which one word is in our Risk dictionary, and the other word in the LM Positive or Negative dictionary.[2] After applying these filters, we obtain a list of risk assessment pairs from each document and each pair contains a (positive or negative) "LM" word

[1] The LM dictionaries are available at the website https://sraf.nd.edu/textual-analysis/resources/.
[2] We lemmatize words and regard a word to be in a dictionary as long as its lemma is in it. Since there is a small overlap of the LM dictionary and our Risk dictionary, when both words in a pair are from LM dictionaries, we remove the pair from our analysis.

and a "Risk" word.[3] We provide several example excerpts of shareholder reports that contain risk assessment pairs in Appendix C. In addition, we consider the cases of negation. For example, "not averse to risk" which contains a pair of Risk word and negative word in fact carries a positive tone. We check all LM words and words paired with them. If a negative (positive) LM sentiment word is also paired with negations such as no, not, and never, such a LM word will be treated as a positive (negative) word.

We define *#NegRiskPair* (*#PosRiskPair*) as the number of pairs with a negative (positive) LM word and a Risk word in a processed N-CSR document and *Length* as the number of words in the document. We then rank *#NegRiskPair/Length* (*#PosRiskPair/Length*) in each document for all filings in a year and scale the rank to [0, 1] to construct our risk assessment measures[4] as follows:

$$NegRisk = Rank(\#NegRiskPair/Length) \quad (1)$$

$$PosRisk = Rank(\#PosRiskPair/Length) \quad (2)$$

While our measure relies on sophisticated natural language processing tools, it is useful to compare our measure with count-based measures of risk and tones commonly used in the prior literature. We construct two sets of count-based measures as follows. The first measure is *RiskWord*, defined as the number of words from our Risk dictionary in a processed N-CSR document scaled by document *Length*. The second set of measures is *NegLM (PosLM)*, which is the number of words from the LM Negative (Positive) dictionary in the document scaled by *Length*. We note that these two sets of measures are constructed based on the count of single words in a document, without

---

[3] In an alternative sample, we require the risk-related word in the pair to serve as a noun in the sentence. In particular, the SNNDP-parsed part of speech of the word should be in the following categories: "NN", "NNP", "NNS", or "NNPS". Our results are qualitatively similar using this alternative sample (untabulated).

[4] We follow Loughran and McDonald (2011) to scale the raw measures *#NegRiskPair* and *#PosRiskPair* by the length of the document. Our results are qualitatively similar if we use the raw measures or construct the measures as *log(#NegRiskPair)* and *log(#PosRiskPair)* where we take the natural logarithms for *#NegRiskPair* and *#PosRiskPair*.

considering word dependencies or pairs. We sort these measures—*RiskWord*, *NegLM*, and *PosLM*—into ranks by year and scale the ranks to [0, 1], respectively, in our empirical analyses. In the untabulated analysis, we construct two additional risk assessment measures (*NegBW* and *PosBW*) using a rule-based Bag-of-Words approach and find qualitatively similar results. Specifically, we count the number of negative (positive) LM words that appear within 10 words of a Risk word and scale by *Length*, which is then ranked and scaled to [0, 1].

3.3 Mutual fund data

We obtain fund return data and fund characteristics such as expense ratio, turnover ratio, total net assets (TNA), and fund age from the Center for Research in Security Prices (CRSP) Survivorship Bias Free Mutual Fund database and fund portfolio holdings from the Thomson Reuters Mutual Fund Holdings database. We use the MFLINKS tables provided by Wharton Research Data Services (WRDS) to merge CRSP Mutual Fund database and the Thomson Reuters Mutual Fund Holdings database.

To merge the N-CSR shareholder reports and the mutual fund databases, we construct a link between Series ID (fund identifier in N-CSR) and the WFICN (Wharton Financial Institution Code Number; the identifier for fund portfolios in MFLINKS). Beginning on February 6, 2006, all open-ended mutual funds are required by SEC to report series (fund portfolio) and class (share class) identification information in their N-CSR filings. For each Series ID, mutual fund companies also report the information related to the underlying share class, including Class ID, Class Name and Class Ticker. We use the Class Ticker to match with the ticker symbol in CRSP Mutual Fund database. When a share class is matched by ticker symbol, we consider the associated Series ID and WFICN as matched. Since Series ID and WFICN are both fund portfolio level identifiers, we drop the cases in which one Series ID is matched to multiple WFICNs. At the portfolio level, we are able to match N-CSR filings with CRSP data for 3,349 domestic equity mutual funds[5].

Although mutual funds start to file N-CSR in 2003, the series and class identification information is not mandatory until 2006. Therefore, our sample period begins in January 2006 and ends in December 2018. Over the 13-year span, our final sample consists of 26,094 N-CSR filings by domestic equity funds. [6] Table 1 shows how our data filters and requirements impact the original sample of N-CSR filings.

[Insert Table 1 Here]

Since we are interested in manager's active investment and risk-taking decisions, we drop ETFs, annuities, and index funds and focus on actively managed funds. In addition, we follow the conventional selection criteria in Kacperczyk, Sialm, and Zheng (2008) to identify domestic equity funds. We aggregate all share classes at the fund level. *TNA* is aggregate total net assets ($mm) across all share classes one month before a filing date. *Age* is the number of years since the fund's oldest share class is launched. We use the natural logarithms of *TNA* and *Age* in our empirical analyses. Return-based variables, turnover ratio (*Turnover*), and expense ratio (*Expense*) are the *TNA*-weighted average across all fund share classes and scaled to percentage points. For the analyses that examine Morningstar fund rating and Morningstar risk component of the rating, we harvest the data from Morningstar.

In general, a fund files shareholder reports semi-annually. We, therefore, calculate fund's risk taking and performance in the 180-day period after the filing of a shareholder report. We measure

[5] Note that N-CSR is filed at the company level and a company may use the same filing for several funds within the fund family. In untabulated results, we find that the returns of funds under the same CIK are highly correlated. We also find qualitatively similar results in untabulated analyses when we exclude filings that contain reports for multiple funds in the same company.

[6] We keep only original filings and exclude amendments in our study because amendments can have identical shareholder reports with original filings and thus provide redundant information. Nevertheless, in untabulated results, we find our results are qualitatively similar if we keep both original and amended filings.

a fund's total risk taking (*Risk*) as the annualized standard deviation of the daily returns in the 180-day period.[7] We measure fund performance (*Alpha*) as the intercept from the regression of daily excess returns on the Fama-French-Carhart four factors, annualized by multiplying with 253. In addition, we compute *SharpeRatio* as the average daily returns in the 180-day period divided by the standard deviation of the daily returns in the same period. To control for past risk taking and past performance, we use the daily returns in the 180 days prior to the filing of a shareholder report to construct the *PastRisk*, *PastAlpha* and *PastSharpeRatio* measures.

Because Morningstar fund rating (*MS Rating*) and its risk component (*MS Risk*) are calculated monthly by Morningstar, we use the 6-month ahead (~ 180 days) values for each filing. Our variable of interest *ΔRisk* is the difference between *Risk* and *PastRisk* and measures the change in the risk-taking behavior between pre-filing 180-day period and post-filing 180-day period. The summary statistics of fund characteristics and text-based variables are reported in Table 2.

[Insert Table 2 Here]

We compute *DownsideBeta* as the average stock level downside beta based on the portfolios that funds disclosure one quarter after the filing. The downside beta of each stock is computed quarterly by regressing the daily excess return within the quarter on the market risk premium conditional on the latter being negative. We follow the extant literature to identify the unobservable skill of fund managers by the return gap measure of Kacperczyk, Sialm, and Zheng (2008). The monthly return gap is the difference between a fund's realized gross return and the hypothetical return on its most recently disclosed portfolio holdings. A higher return gap has been shown to

[7] We annualize the standard deviation of daily returns by multiplying it with the square root of 253, since there are on average 252.75 trading days in a year.

predict better future performance and thus proxy better unobservable skills. We define *RetGap* and *PastRetGap* as the monthly average return gap post and prior to a filing date, respectively.

4. Empirical Analyses

4.1 Post-Filing Fund Risk Taking

Our *NegRisk* and *PosRisk* measures indicate the risk assessment of fund managers. More importantly, the uniqueness of our measures is their ability to identify whether a manager's opinion on risk is favorable or not. For instance, suppose a manager states "While risk taking can sometimes be beneficial, we need to be aware of the potentially detrimental effects of risk." If we only count the number of appearances for the word "risk", we would end up with a measure of 2 as the risk assessment of the manager. However, our measures consider the sentence to have one "positive risk" opinion and one "negative risk" opinion, and enable us to example how opposing opinions lead to subsequent actions. One may also simply use the negative word counts and positive word counts to proxy the risk assessment. However, this approach is also less accurate because these word counts may represent sentiment about other subjects rather than about risk.

We first run a horse race among our risk assessment measures, *NegRisk* and *PosRisk*, three simple count-based sentiment measures, *RiskWord*, *NegLM*, and *PosLM*, and examine whether these measures can predict the risk-taking behavior in the post-filing period. In this study, we focus on domestic equity funds given that they are the predominant class of mutual funds and their performance has well-defined benchmarks, such as the Fama-French-Carhart four-factor model. In Table 3, we estimate the following regression

$$\Delta Risk_{i,t+1} = \alpha_i + \beta RiskAssessment_{i,t} + \gamma' X_{i,t} + \delta' FE_{i,t} + \epsilon_{i,t} \tag{3}$$

where i and t indicate the mutual fund and 180-day reporting period of the N-CSR filing. *RiskAssessment* may include *NegRisk*, *PosRisk*, *RiskWord*, *NegLM*, and *PosLM*. The control variables $X_{i,t}$ include *PastRisk*, *PastAlpha*, *log(TNA)*, *log(Age)*, *Expense*, and *Turnover*. $FE_{i,t}$ may include various fixed effects, including year fixed effects, fund fixed effects, and company fixed effects.[8] We first individually use the assessment measures in the regression and report the results in Columns (1) to (6) of Table 3 and then include all five measures in Columns (7) and (8). We find that the change in fund risk taking in the next period ($\Delta Risk$) depends negatively on *NegRisk* and positively on *PosRisk*, regardless of various fixed effects. The predictive power is virtually unchanged when adding count-based sentiment measures, as shown in the last two columns. On the other hand, the sign of the coefficients on *RiskWord*, *NegLM*, and *PosLM* is flipped when fixed effects are included.

[Insert Table 3 Here]

The evidence suggests that our deep-learning-based measures, *NegRisk* and *PosRisk*, capture managerial risk assessment more accurately than count-based measures. Such evidence serves as a validation test to show that a greater *NegRisk* (*PosRisk*) indicates a greater aversion (loving) to risk. The results suggest that when a manager has a negative (positive) assessment on risk, she would subsequently reduce (increase) the level of risk that her fund takes. The economic magnitude is also substantial. Overall, the investment behavior of fund managers is consistent with what they disclose to the fund investors, and our measures extract managers' forward-looking opinions from the textual disclosure.

4.2 Post-Filing Fund Performance

After establishing that mutual fund managers adjust risk taking based on their risk assessment, a natural question is whether their behavior brings benefits or costs to their fund investors. For

[8] Some companies may file a single shareholder report for several funds within the fund family. *NegRisk* and *PosRisk* may be therefore the same for these funds. We add company fixed effects to control for the common unobserved characteristics within the company.

instance, what should investors expect when their fund manager has a negative assessment on risk and plans to reduce the fund risk?

We answer the above question by exploiting the post-filing performance of the fund. Although the debate on whether actively managed mutual funds are skilled (i.e., whether funds have positive alphas) still exists, a substantial literature (see, for example, Jensen, 1968; Malkiel, 1995; Fama and French, 2010) shows that an average fund does not provide positive alpha to its fund investors. Therefore, if the discussion in the shareholder report reflects only the manager's belief and assessment but does not have informational value, the investors of the fund should not expect a positive alpha. In contrast, if the shareholder report contains valuable information or insight, the investors should expect superior performance when their fund manager acts consistently with the risk assessment disclosed in the shareholder report, especially when the manager is skilled.

[Insert Table 4 Here]

We estimate the following regression

$$Alpha_{i,t+1} = \alpha_i + \beta RiskSentiment_{i,t} + \gamma' X_{i,t} + \delta' FE_{i,t} + \epsilon_{i,t} \quad (4)$$

where *Alpha* is the Fama-French-Carhart four-factor alpha of net-of-fee returns over the 180-day period after the filing of shareholder reports. Table 4 shows that the coefficients on *PosRisk* are statistically indifferent from zero, suggesting positive risk assessment on average adds little value to fund performance. However, funds achieve superior performance after they exhibit negative risk assessment in their shareholder report: the coefficients on *NegRisk* are positive and statistically significant. The results are robust to including various fixed effects and controlling for the count-based sentiment measures. Overall, Table 4 supports that the shareholder report contains material information, and such information is concentrated in negative risk assessment.

To further examine the subsequent performance associated with managers' risk assessment, we replace *Alpha* in equation (4) with *SharpeRatio* to explore how managers adjust the trade-off between risk taking and expected return:

$$SharpeRatio_{i,t+1} = \alpha_i + \beta RiskAssessment_{i,t} + \gamma' X_{i,t} + \delta' FE_{i,t} + \epsilon_{i,t} \quad (5)$$

Table 5 shows that managers with negative risk assessment gain higher Sharpe ratio in the future. Funds with positive risk assessment obtain zero to negative Sharpe ratio, likely because they could not generate better performance despite taking higher risk. In addition to the results in the previous section that managers reduce fund risk when they have negative risk assessment, the results in Table 5 suggest that when managers behave consistently with their negative risk assessment, their investors stand to benefit in terms of superior fund performance.

[Insert Table 5 Here]

4.3 Managerial Skill and Shareholder Report

The previous results suggest that one can predict fund behavior and performance based on the informational content in the shareholder reports. Although shareholder reports are mandatorily disclosed, the exact discussion and information are entirely within the managers' discretion. In other words, managers can strategically choose to disclose certain information to their investors. Next, we explore whether the disclosed information represents the quality of the manager.

[Insert Table 6 Here]

We follow the investment literature and consider return gap (*RetGap*), proposed in Kacperczyk, Sialm and Zheng (2008), as a proxy for managerial skill. We first explore whether our risk assessment measures can help investors select skilled managers. Specifically, we exploit the following regression:

$$RetGap_{i,t+1} = \alpha_i + \beta RiskSentiment_{i,t} + \gamma' X_{i,t} + \delta' FE_{i,t} + \epsilon_{i,t} \quad (6)$$

Results in Table 6 show that managers with negative risk assessment are more likely to generate larger return gap in the subsequent six-month and one-year periods. The positive assessment, on the other hand, appears to provide little informational value to fund investors. The overall evidence suggests that negative risk assessment is more informative and helps investors identify skilled managers.

To understand how managerial skill is associated with managers' risk assessment, and why negative risk assessment contains privileged information while positive risk assessment does not, we study how skilled managers exhibit their risk assessment in the disclosure. We group funds into two groups based on whether the average return gap in the past six months (*Past RetGap*) is greater than zero or not, and then re-estimate equations (3), (4) and (5) for the two groups. A greater-than-zero average return gap indicates that the unobserved actions of a fund manager leads to greater returns and thus is associated with a high-quality or skilled manager.

Table 7 Panel A shows that both groups reduce (increase) their total risk when the managers have negative (positive) risk assessment. However, the group of skilled funds (with *Past RetGap* > 0) engages in more dramatic reductions in risk taking when having negative assessment: the difference in the coefficients for the two groups is statistically significant at the 1% level. In addition, Table 7 Panel B suggests that only the skilled group can translate their behavior into positive risk-adjusted returns, measured by four-factor alpha. The coefficient on *NegRisk* is positive and statistically significant for the skilled managers while the coefficient is indistinguishable from zero for their unskilled counterparties. The difference in coefficients for the two groups is also statistically significant at the 10% level, suggesting that skilled managers achieve superior performance for investors by reducing fund risk based on their negative risk assessment. Panel C produces similar but weaker results for Sharpe ratio. Although the coefficient on *NegRisk* is for skilled managers is larger than that for unskilled managers, the difference is not statistically significant.

[Insert Table 7 Here]

Overall, we provide support that shareholder reports contain important information from which investors can infer fund managers' subsequent behavior. In addition, such information is particularly valuable when disclosed by skilled managers, who can take actions and translate their information into superior future fund performance.

4.4 The Source of Superior Performance

So far we find that risk-conscious managers tend to be more skilled and deliver better performance than other managers. A natural question is: how do they achieve it? We thus consider the reasons that managers want to reduce risk-taking to explore the source of superior future performance. One explanation is the downside risk of the stock market. If a manager anticipates the market to perform poorly, e.g., to be at the brink of a bear market, she could reduce her fund's exposure to the market risk.

In this subsection, we explore whether managers reduce the downside beta of fund portfolios and, if so, whether they actively change it. We start the analysis with relating the downside beta of a fund's portfolio after a fund's filing to its manager's risk assessment. Each quarter, we calculate the downside beta for each stock by using excess daily returns of the stock within the quarter and taking the beta coefficient of the market model conditional on negative market risk premium. We then aggregate the downside beta to fund level using the average stock downside beta across a fund's portfolio post filing. Table 8 column (1) provides supportive evidence that a fund with greater negative risk assessment has lower downside beta subsequent to the fund's filing.

[Insert Table 8 Here]

To rule out the possibility that the reduction in a fund's downside beta is purely a change in underlying stocks' downside betas rather than the manager actively trading, we further compute the average downside betas of the stocks purchased and of the stocks sold, respectively, by the fund over the past filing period. We consider the difference between the two betas as the manager's action in changing the fund's downside beta. The second column in Table 8 confirms that it is a manager's active trading that results in low downside beta of the fund when her negative risk assessment is high. In other words, the manager buys stocks with low downside beta and sell stocks with high downside beta, leading to a reduced downside beta for the entire portfolio.

An underlying assumption for a manager to lower the downside beta and deliver greater fund performance is that the market performs poorly subsequently so that reducing the exposure to market risk helps avoid potential losses. We validate such an assumption by looking at the number of days with negative market risk premium post a fund's filing. In column (3) of Table 8, we find that when a manager has greater negative risk assessment, she is likely to experience more days with negative market returns. Overall, the results in Table 8 provides one channel through which risk-conscious managers are able to perform well – by reducing the exposure to downside risk.

4.5 Response to and Recognition of Risk Assessment

Investors' decision to invest in a fund or not plays a vital role in managerial compensations. Therefore, an essential incentive for managers to disclose useful information in the shareholder reports is to attract capital flows. Meanwhile, each fund manager competes with her peers for these flows. In this section, we explore if the investment community responds to fund managers' risk assessment and what types of investors can recognize the risk assessment skill.

Recent studies (Cheng, Lu, and Zhang, 2021; Evans and Sun, 2021; Ben-David, Li, Rossi, and Song, 2022) find that funds with Morningstar rating is an important indicator for investors to invest their capital into funds.. Because Morningstar rating is calculated with consideration for investors' risk aversion, the risk component is an important part of the rating. It is thus plausible that managers with different risk assessments adjust the risk component and, in respond to the adjustment in risk taking, Morningstar assigns higher rating for funds with better risk assessment skill.[9]

To examine this conjecture, we explore whether our risk assessment measures are predictors of future Morningstar rating. Specifically, we analyze the following regression,

$$MS\ Rating_{i,t+1} = \alpha_i + \beta RiskAssessment_{i,t} + \gamma' X_{i,t} + \delta' FE_{i,t} + \epsilon_{i,t} \tag{6}$$

where i and t indicate the mutual fund and 180-day reporting period of the N-CSR filing. *RiskAssessment* includes *NegRisk*, *PosRisk*, *RiskWord*, *NegLM* and *PosLM*. The control variables $X_{i,t}$ include *PastRisk*, *PastAlpha*, *PastRating*, *log(TNA)*, *log(Age)*, *Expense* and *Turnover*. $FE_{i,t}$ may include various fixed effects, including year fixed effects, fund fixed effects, and company fixed effects. Because Morningstar rating is updated monthly, we use the 6-month ahead rating after date t as the $MS\ Rating_{i,t+1}$. We also replace *MS Rating* by *MS Risk* to study the risk component of the rating.

[Insert Table 9 Here]

Columns 1 and 2 of Table 9 show clear evidence that managers with negative risk sentiment reduce *MS Risk*. In sharp contrast, managers with positive risk assessment change little on *MS Risk*. Columns 3 and 4 provide weak but supportive evidence that those managers who reduce *MS Risk* are able to obtain a higher Morningstar rating. Therefore, for managers, behaving consistently with

[9] For the methodology of Morningstar rating and its risk component, see: https://www.morningstar.com/content/dam/marketing/shared/research/methodology/771945_Morningstar_Rating_for_Funds_Methodology.pdf

negative risk assessment not only generates superior performance, but also leads to a better Morningstar rating that can help attract capital flows.

Next, we examine whether investors can recognize the negative risk assessment and, if so, what type of investors are able to discover the risk assessment skill. Although we have shown that qualitative disclosure in shareholder reports contains a unique value, it provides less direct information than quantitative disclosure, such as past fund performance, especially to investors who are less sophisticated. In other words, sophisticated investors are more likely to parse out the valuable information from the unstructured text disclosed by fund managers. Chalmers and Reuter (2012), Del Guercio and Reuter (2014), and Barber, Huang, and Odean (2016) find that investors of direct-sold funds are more sophisticated than investors in broker-sold funds. We thus define funds that sell shares through brokers as funds with naïve investors (*Sophisticated* = 0) and those that sell directly to investors as funds with sophisticated investors (*Sophisticated* = 1). We examine how flows respond to managers' risk assessment disclosure differentially for naïve and sophisticated investors.

[Insert Table 10 Here]

Table 10 presents the findings. The positive coefficient on the interaction term between *NegRisk* and *Sophisticated* is statistically significant, while the coefficients on the interaction terms between other risk assessment measures and *Sophisticated* are all indistinguishable from zero. The results suggest that funds with sophisticated investors attract more flows when they disclose greater negative risk assessment, relative to their peers with naïve investors. This is consistent with our prediction that sophisticated investors are likely to recognize the risk assessment skill.

Overall, we find that the large degree of freedom on managers' textual disclosure offer them an effective way to communicate with current as well as potential investors by providing quality information. Managers with negative risk assessment are able to reduce their risk taking to acquire higher Morningstar rating, suggesting that the investment community responds to managers efforts to share their unique information Furthermore, only funds with sophisticated investors are able to attract more capital inflows in disclosing negative risk assessment as their investors are more capable of understanding qualitative information contained in their disclosure. It also suggests that only a group of investors recognize the risk assessment skill, emphasizing the uniqueness and usefulness of our risk assessment measure to the general public, especially to retail investors and less sophisticated investors.

6. Conclusion

We use a deep learning model to extract syntactic structures from textual data of mutual fund disclosure and construct forward-looking risk assessment measures, which capture the manager's assessment and belief about the risks facing a mutual fund. Managers with a more negative (positive) risk assessment are more likely to reduce (increase) their portfolio risk in the following period. Although managers adjust their risk taking consistently with their risk assessment, only negative risk assessment in the disclosure contains useful information to investors, and managers with negative risk assessment generate superior risk-adjust return, higher Sharpe ratio and larger return gap. Skilled managers are more likely to change their fund risk in accordance with their own risk assessment and obtain higher alphas as a result.

The investment community responds to managers who disclose quality information with a prudent view. Managers with negative risk assessment are more likely to be assigned a higher Morningstar rating. Given the forward-looking nature, our new measures can inform investors and researchers about fund managers' risk management and investment decisions. Interestingly, not all market participants can recognize this narrative-based measure except sophisticated investors. The forward-looking nature of our new measure can thus inform investors and researchers about fund managers' risk management and performance.

Because our deep-learning-based measures capture higher-order syntactic interactions among words, they generate superior results than measures constructed with more traditional bag-of-word approaches. We also note that transfer learning, or building special-purpose models based on pre-trained deep learning models that utilized large-scale, general labeled data, can be time-saving and solve the challenge of lack of training data for machine learning models. Overall, we believe it is promising to develop more applications of deep learning models in textual analytics that can reveal and analyze linguistic features previously inaccessible to researchers.

Appendix A: Definitions of Variables

Variable	Definition
Alpha	The Fama-French-Carhart four-factor alpha using daily returns during days [0, 180] for a fund's filing on day 0.
DownsideBeta	Each quarter, each stock's daily excess returns are regressed on the daily market risk premium when it is negative to obtain the downside beta of a stock. *DownsideBeta* of a fund is the average stock downside beta based on a fund's portfolio disclosed one quarter after its shareholder report. An alternative definition of *DownsideBeta* is the average downside beta of purchased stocks minus the average downside beta of sold stocks.
Expense	The most recent expense ratio prior to filing month *t*.
Flow6m/Flow12m	The future 6-month flow for the filing of fund *i* in month *t*, expressed in percentage points. $Flow6m = (TNA_{i,t+5} - TNA_{i,t-1} \times R_{i,t-1,t+5}) / TNA_{i,t-1}$ *Flow12m* is calculated analogously.
log(Age)	The natural logarithm of a fund's age.
log(TNA)	The natural logarithm of a fund's total net assets (TNA) in month $t - 1$.
MS Rating	Morningstar rating in month $t + 5$ where *t* is the filing month.[10]
MS Risk	The risk component of Morningstar rating (*MS Rating)*.
NegLM	The number of words from the Loughran-McDonald Negative dictionary in the document scaled by the number of words in the document.
NegRisk	*NegRisk = Rank(#NegRiskPair/Length)*. *#NegRiskPair* is the number of pairs with a negative LM word and a Risk word (See Table A1) in a processed N-CSR document and *Length* is the number of words in the document. We then rank #*NegRiskPair/Length* for each document every year and scale the rank to [0, 1].
PastAlpha	The Fama-French-Carhart four-factor alpha using daily returns during days [-180, -1] for a fund's filing on day 0.
PastFlow	The past 6-month flow during months [$t - 6$, $t - 1$] for a fund's filing in month *t*.
PastMSRating	Morningstar rating in month $t - 1$ where *t* is the filing month.
PastMSRisk	The risk component of *MS Risk.*
PastRetGap	The average return gap in the past 6 months.
PastRisk	The annualized standard deviation of the daily returns during days [-180, -1] for a fund's filing on day 0.
PastSharpeRatio	The average daily returns during days [-180, -1] divided by the standard deviation of the returns in the same period for a fund's filing on day 0.
PosLM	The number of words from the Loughran-McDonald Positive dictionary in the document scaled by the number of words in the document.
PosRisk	*PosRisk = Rank(#PosRiskPair/Length)*, analogous to *NegRisk*.
ΔRisk	*ΔRisk = Risk – PastRisk.* A fund's total risk taking (*Risk*) is the annualized standard deviation of the daily returns during days [0, 180] for a fund's filing on day 0.
RetGap6m/RetGap12m	The monthly return gap is the difference between a fund's realized gross return and the hypothetical return on its most recently disclosed portfolio

[10] For the complete methodology by Morningstar to calculate fund rating, see: https://www.morningstar.com/content/dam/marketing/shared/research/methodology/771945_Morningstar_Rating_for_Funds_Methodology.pdf

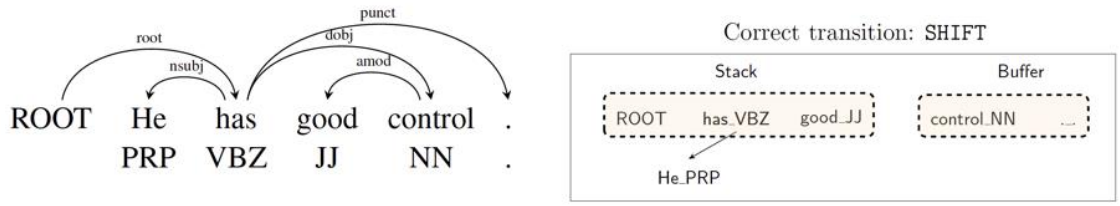
	holdings. *RetGap6m (RetGap12m)* is the average return gap in the future 6 months (12 months).
RiskWord	The number of words from our Risk dictionary (See Table A1) in a document scaled by the number of words in the document.
SharpeRatio	The average daily returns during days [0, 180] divided by the standard deviation of the returns in the same period for a fund's filing on day 0.
Sophisticated	An indicator variable equal to one if a fund sells directly to investors, i.e., a fund with sophisticated investors, and zero if a fund sells shares through brokers, i.e., a fund with naïve investors.
Turnover	The most recent turnover ratio prior to filing month *t*.

Appendix B. Details for Dependency Tree Parsing

A natural language parser is a program that works out the grammatical structure of sentences which is usually referred as the tree-structure of sentences. The goal of using tree-structure to represent a sentence is that the flat, streaming and structure features can all be extracted for further content analysis.

To illustrate the parsing procedure, we first introduce configurations and transitions. A configuration for a given sentence to be parsed consists of three components: a stack s, a buffer b, and a set of dependency arcs A. The stack and buffer are used to store words for processing. Each arc describes the relationship between a head word and a modifying word and is labeled by the nature of the modifying relationship, e.g., *advmod*, *adjmod*, and *nmod*. Given any sentence with n words, $w_1, w_2, \cdots, w_n$, its initial configuration is $C_0 = \{s = [ROOT], b = [w_1, w_2, \cdots, w_n], A = \emptyset\}$, and typically its terminal configuration is achieved when the buffer b becomes empty, denoted as C_*. The terminal configuration contains a complete set of arcs that determine the target dependency tree structure of the sentence.

Each step of the parsing process is called a transition. There three types of transitions, $LEFT - ARC(l)$, $RIGHT - ARC(l)$, and $SHIFT$. Given any configuration $\{s = [s_1, s_2, \cdots], b = [b_1, b_2, \cdots], A\}$, a $LEFT - ARC(l)$ transition adds an arc $s_1 \rightarrow s_2$ with label l and remove s_2 from the stack, a $RIGHT - ARC(l)$ transition adds an arc $s_2 \rightarrow s_1$ with label and removes s_1 from the stack, and a $SHIFT$ moves b_1 from the buffer to the stack. This procedure continues until there are no words left in both the buffer and the stack and the tree-structure of the sentence $w_1, w_2, \cdots, w_n$ can be derived from the set of dependency arcs A. Figure 3 shows an example of such a parsing procedure.



Transition	Stack	Buffer	A
	[ROOT]	[He has good control .]	$\emptyset$
`SHIFT`	[ROOT He]	[has good control .]	
`SHIFT`	[ROOT He has]	[good control .]	
`LEFT-ARC(nsubj)`	[ROOT has]	[good control .]	$A\cup$ nsubj(has,He)
`SHIFT`	[ROOT has good]	[control .]	
`SHIFT`	[ROOT has good control]	[.]	
`LEFT-ARC(amod)`	[ROOT has control]	[.]	$A\cup$amod(control,good)
`RIGHT-ARC(dobj)`	[ROOT has]	[.]	$A\cup$ dobj(has,control)
...	...	...	...
`RIGHT-ARC(root)`	[ROOT]	[]	$A\cup$ root(ROOT,has)

Figure 3. The transition steps in the parsing of an example sentence

Appendix C. Examples of Shareholder Reports with Risk Assessment Statements

This appendix provides excerpts of shareholder reports in which risk assessment statements, i.e., dependency pairs of risk and positive/negative words, appear. The fund name, filing date, and subsequent change in fund risk are reported.

Excerpts	Fund and filing information
My team and I have continually worked, tirelessly, to improve the strategy while **reducing risk** each and every day. JFK once said, "Great accomplishments are not achieved by extraordinary men doing extraordinary things extraordinarily well, but by ordinary men doing ordinary things extraordinarily well." I believe the team at IPS Strategic Capital is a group of very hard-working professionals that look to achieve extraordinary things.	IPS Strategic Capital Absolute Return Fund, May 07, 2018 $\Delta Risk$ = -0.63%
Although diversification doesn't eliminate the risk of loss or guarantee a profit, a careful selection of complementary asset classes may cushion your portfolio against **excessive volatility**.	AIM Funds Group, March 06, 2009, $\Delta Risk$ = -1.08%
As bottom-up, fundamental stock pickers, we maintain our focus on identifying businesses with idiosyncratic growth drivers that should power through a variety of economic or market scenarios and whose stocks present **attractive risk**/reward opportunities. We believe that if we can identify and invest in high-quality companies with more durable growth opportunities than the market expects, investors in the stocks of those companies should be rewarded.	Vanguard U.S. Growth Fund, October 25, 2017, $\Delta Risk$ = 7.88%

Reference

Abis, Simona, 2022, Man vs. machine: Quantitative and discretionary equity management, working paper.

Abis, Simona, Andrea M. Buffa, Apoorva Javadekar, and Anton Lines, 2022, Learning from prospectuses, working paper.

Abis, Simona, and Anton Lines, 2022, Do mutual funds keep their promises? working paper.

Adebambo, Biljana N., and Xuemin Yan, 2018, Investor overconfidence, firm valuation, and corporate decisions, *Management Science* 64, 5349-5369.

Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which factors matter to investors? Evidence from mutual fund flows, *Review of Financial Studies* 29, 2600-2642.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2022, What do mutual fund investors really care about, *Review of Financial Studies* 35, 1723-1774.

Brown, Keith C., W. V. Harlow, and Laura T. Starks, 1996, Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry, *Journal of Finance* 51, 85-110.

Cao, Sean S., Guang Ma, Jennifer Wu Tucker, and Chi Wan, 2018, Technological peer pressure and product disclosure, *The Accounting Review* 93, 95-126.

Cao, Sean S., Wei Jiang, Baozhong Yang, and Alan L. Zhang, 2022, How to talk when a machine is listening: Corporate disclosure in the age of AI, working paper.

Cao, Sean S., Kai Du, Baozhong Yang, and Alan L. Zhang, 2021, Copycat skills and disclosure costs: Evidence from peer companies' digital footprints, *Journal of Accounting Research* 59, 1261-1302.

Cao, Zhiyan, and Ganapathi S. Narayanamoorthy, 2011, The effect of litigation risk on management earnings forecasts, *Contemporary Accounting Research* 28, 125-173.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

Chalmers, John, and Jonathan Reuter, 2012, How do retirees value life annuities? Evidence from public employees, *Review of Financial Studies* 25, 2601-2634.

Chen, Danqi, and Christopher Manning, 2014, A fast and accurate Dependency Parser using Neural Networks, *Proceedings of EMNLP*.

Cheng, Si, Ruichang Lu, and Xiaojun Zhang, 2021, What should investors care about? Mutual fund ratings by analysts vs. Machine learning technique, working paper.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167-1200.

Cong, Lin and Tengyuan Liang and Xiao Zhang, 2019, Textual factors: A scalable, interpretable, and data-driven approach to analyzing unstructured information, working paper.

Cong, Lin William, Tengyuan Liang, Baozhong Yang, and Xiao Zhang, 2020, Analyzing textual information at scale, Book Chapter in Information to Facilitate Efficient Decision Making: Big Data, Blockchain and Relevance, World Scientific Publishers.

Evans, Richard B., and Yang Sun, 2021, Models or stars: The role of asset pricing models and heuristics in investor risk adjustment, *Review of Financial Studies* 34, 67-107.

Du, Qianzhou, Yawen Jiao, Pengfei Ye, and Weiguo Fan, 2020, When mutual fund managers write confidently, working paper.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, Eugene F., and Kenneth R. French, 2010, Luck versus skill in the cross-section of mutual fund returns, *Journal of Finance* 65, 1915-1947.

Field, Laura, Michelle Lowry, and Susan Shu, 2005, Does disclosure deter or trigger litigation?, *Journal of Accounting and Economics* 39, 487-507.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy, 2019, Text as data, *Journal of Economic Literature* 57, 535-74.

Guercio, Diane Del, and Jonathan Reuter, 2014, Mutual fund performance and the incentive to generate alpha, *Journal of Finance* 69, 1673-1704.

Hanley, Kathleen Weiss, and Gerard Hoberg, 2019, Dynamic interpretation of emerging risks in the financial sector, *Review of Financial Studies* 32, 4543-4603.

Hassan, Tarek A, Stephan Hollander, Laurence van Lent, and Ahmed Tahoun, 2019, Firm-level political risk: Measurement and effects, *Quarterly Journal of Economics* 134, 2135-2202.

Hillert, Alexander, Alexandra Niessen-Ruenzi, and Stefan Ruenzi, 2021, Mutual fund shareholder letters: flows, performance, and managerial behavior, working paper.

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945–1964, *Journal of Finance* 23, 389-416.

Jiang, Fuwei, Joshua Lee, Xiumin Martin, and Guofu Zhou, 2019, Manager sentiment and stock returns, *Journal of Financial Economics* 132, 126-149.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2008, Unobserved actions of mutual funds, *Review of Financial Studies* 21, 2379-2416.

Kothari, Sabino P., Susan Shu, and Peter D. Wysocki, 2009, Do managers withhold bad news? *Journal of Accounting Research* 47, 241–276.

Li, Kai, Feng Mai, Rui Shen, and Xinyan Yan, 2021, Measuring corporate culture using machine learning, *Review of Financial Studies* 34, 3265-3315.

Loughran, Tim, and Bill Mcdonald, 2011, When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks, *Journal of Finance* 66, 35-65.

Loughran, Tim, and Bill Mcdonald, 2016, Textual analysis in accounting and finance: A survey, *Journal of Accounting Research* 54, 1187-1230.

Ma, Linlin, and Yuehua Tang, 2019, Portfolio manager ownership and mutual fund risk taking, *Management Science* 65, 5518-5534.

Malkiel, Burton G., 1995, Returns from investing in equity mutual funds 1971 to 1991, *Journal of Finance* 50, 549-572.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean, 2013, Efficient estimation of word representations in vector space, working paper, Google Inc.

Palomino, Frederic, and Abdolkarim Sadrieh, 2011, Overconfidence and delegated portfolio management, *Journal of Financial Intermediation* 20, 159-177.

Pool, Veronika K, Noah Stoffman, Scott E Yonker, and Hanjiang Zhang, 2018, Do shocks to personal wealth affect risk-taking in delegated portfolios?, *Review of Financial Studies* 32, 1457-1493.

Sheng, Jinfei, Nan Xu, and Lu Zheng, 2022, Do mutual funds walk the talk? A textual analysis of risk disclosure by mutual funds, working paper.

Table 1 N-CSR Sample Creation

This table reports the impact of the imposition of various filters on sample size in constructing our sample of mutual fund shareholder reports, returns, and characteristics data.

Source/Filter	Filing Sample Size	Filings Removed
EDGAR N-CSR/N-CSRS 2006-2018 complete sample	73,346	
Shareholder report (Item 1) can be extracted	71,117	2,229
Contain Series ID and Class Ticker	40,279	30,838
Matched to WFICN	32,498	7,781
Exclude index funds, annuities, and ETFs	30,072	2,426
Valid fund returns and characteristics around filing dates	28,551	1,521
Number of words in shareholder reports > 250	27,428	1,123
Select domestic equity funds	24,404	3,024
Final Sample	24,404	

Table 2 Summary Statistics

This table provides the summary statistics. The sample spans from 2006 to 2018, and the risk assessment measures are based on shareholder reports in Form N-CSR. All variables are defined in Appendix A.

Variables	(1) Mean	(2) Median	(3) Std	(4) P25	(5) P75	(6) N
NegRisk	0.476	0.476	0.279	0.235	0.717	131,201
PosRisk	0.391	0.391	0.315	0.052	0.669	131,201
RiskWord	0.480	0.479	0.279	0.239	0.720	131,201
NegLM	0.475	0.475	0.277	0.236	0.714	131,201
PosLM	0.476	0.475	0.276	0.237	0.714	131,201
ΔRisk	0.163	-0.073	8.382	-2.495	2.629	131,201
PastRisk	14.590	13.160	10.440	8.178	18.230	131,201
Alpha	-0.922	-0.581	8.863	-4.784	3.602	131,201
PastAlpha	-0.877	-0.608	8.868	-4.752	3.549	131,201
SharpeRatio	6.377	5.630	12.270	-2.701	14.280	131,198
PastSharpeRatio	6.494	5.954	12.120	-2.371	14.220	131,198
RetGap6m	-0.050	-0.036	0.734	-0.215	0.122	82,521
RetGap12m	-0.050	-0.037	0.534	-0.173	0.085	77,423
PastRetGap	-0.051	-0.035	0.731	-0.212	0.123	85,061
MS Rating	3.002	3.000	0.960	2.250	3.750	114,833
PastMSRating	3.009	3.000	0.957	2.333	3.750	113,840
MS Risk	0.020	0.013	0.022	0.005	0.029	131,201
PastMSRisk	0.019	0.012	0.085	0.004	0.025	129,086
Flow6m	1.735	-3.675	36.060	-8.904	3.265	130,040
Flow12m	4.429	-7.518	57.760	-17.020	6.017	128,221
PastFlow	3.063	-3.413	39.950	-8.747	3.942	129,818
Sophisticated	0.331	0.000	0.471	0.000	1.000	131,201
log(TNA)	5.678	5.765	1.852	4.458	6.962	131,201
log(Age)	2.654	2.803	0.807	2.359	3.128	131,201
Expense	1.084	1.070	0.381	0.830	1.320	131,201
Turnover	78.080	53.000	89.990	27.000	93.000	131,201

Table 3 Risk Assessment and Risk-Taking

This table reports the relation between risk assessment measures and fund risk-taking. *NegRisk (PosRisk)* captures managers' negative (positive) risk assessment in their narratives. Panel A shows the change in the risk-taking for groups sorted by negative risk assessment. Funds are grouped into quintiles and the average change in risk-taking is reported for each group. The last column reports a two-sample *t*-test between the lowest and highest quintiles. Panel B shows regressions of change in the risk-taking on risk assessment measures and fund characteristics. The dependent variable *ΔRisk* is the change in the total risk-taking from before to after a filing date where the total risk-taking is defined as the annualized standard deviation of the daily returns over a 180-day period. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by fund. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Univariate sort on risk assessment

	ΔRisk					
Quintile	Low	2	3	4	High	High - Low
NegRisk	0.43	0.24	0.07	0.29	-0.22	-0.65
						(8.88)

Panel B: Risk Assessment and Risk-Taking

Dependent Variable	ΔRisk							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NegRisk	-1.503***	-0.561***					-1.118***	-0.891***
	(-13.08)	(-7.05)					(-7.96)	(-8.81)
PosRisk	0.390***	0.272***					0.273***	0.312***
	(4.48)	(4.27)					(2.62)	(4.35)
RiskWord			-1.345***	0.451***			-0.799***	0.963***
			(-13.74)	(6.90)			(-5.89)	(12.64)
NegLM					-0.874***	-0.220***	0.089	-0.302***
					(-5.91)	(-2.66)	(0.44)	(-2.85)
PosLM					0.638***	-0.093	0.397**	-0.092
					(4.48)	(-1.31)	(2.47)	(-1.13)
PastRisk	-0.377***	-0.739***	-0.378***	-0.740***	-0.373***	-0.739***	-0.379***	-0.737***
	(-61.25)	(-126.47)	(-60.42)	(-126.25)	(-59.79)	(-126.10)	(-59.14)	(-126.29)
PastAlpha	-0.006	0.034***	-0.006	0.034***	-0.006	0.034***	-0.006	0.034***
	(-1.19)	(10.11)	(-1.28)	(10.10)	(-1.21)	(10.08)	(-1.34)	(10.07)
log(TNA)	0.498***	0.365***	0.523***	0.367***	0.518***	0.365***	0.508***	0.366***
	(15.44)	(6.98)	(16.38)	(6.99)	(16.17)	(6.95)	(15.99)	(7.01)
log(Age)	-1.308***	-1.128***	-1.316***	-1.147***	-1.327***	-1.122***	-1.297***	-1.106***
	(-15.46)	(-7.79)	(-15.74)	(-7.93)	(-15.88)	(-7.77)	(-15.52)	(-7.66)
Expense	3.584***	1.873***	3.568***	1.877***	3.590***	1.883***	3.552***	1.864***
	(21.15)	(6.90)	(21.07)	(6.89)	(21.23)	(6.91)	(20.94)	(6.85)
Turnover	-0.001	0.003***	-0.001	0.003***	-0.001	0.003***	-0.001	0.003***
	(-1.56)	(3.87)	(-1.64)	(3.79)	(-1.45)	(3.83)	(-1.36)	(3.86)
Observations	131,201	131,138	131,201	131,138	131,201	131,138	131,201	131,138
Year FE	No	Yes	No	Yes	No	Yes	No	Yes
Fund FE	No	Yes	No	Yes	No	Yes	No	Yes
Company FE	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R-squared	0.197	0.598	0.196	0.598	0.195	0.598	0.198	0.598

Table 4. The Implications of Risk Management for Future Performance: Abnormal Returns

This table reports the relation between risk assessment measures and future performance, measured by Fama-French-Carhart Four-factor Alpha. *NegRisk (PosRisk)* captures managers' negative (positive) risk assessment in their narratives. Panel A shows the *Alpha* for groups sorted by negative risk assessment. Funds are grouped into quintiles and the average *Alpha* is reported for each group. The last column reports a two-sample *t*-test between the lowest and highest quintiles. Panel B shows regressions of future performance, measured by Fama-French-Carhart Four-factor Alpha, on risk assessment measures and fund characteristics. The dependent variable *Alpha* is the annualized intercept from a regression on Fama-French-Carhart four-factor model over a 180-day period after a filing date. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by fund. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Univariate sort on risk assessment

	Alpha					
Quintile	Low	2	3	4	High	High - Low
NegRisk	-1.25	-1.11	-0.72	-0.88	-0.65	-0.60
						(7.66)

Panel B: Risk Assessment and Abnormal Returns

Dependent Variable	*Alpha*			
	(1)	(2)	(3)	(4)
NegRisk	0.271**	0.301***	0.553***	0.616***
	(1.96)	(2.71)	(3.46)	(4.27)
PosRisk	0.018	0.071	-0.029	0.026
	(0.18)	(0.84)	(-0.27)	(0.29)
RiskWord			-0.658***	-0.676***
			(-4.48)	(-5.22)
NegLM			0.149	0.025
			(0.90)	(0.16)
PosLM			0.175	0.190
			(1.41)	(1.52)
PastRisk	-0.041***	-0.001	-0.043***	-0.002
	(-6.80)	(-0.14)	(-6.90)	(-0.19)
PastAlpha	0.080***	-0.031***	0.079***	-0.031***
	(12.56)	(-5.11)	(12.47)	(-5.10)
log(TNA)	-0.221***	-1.552***	-0.214***	-1.554***
	(-6.51)	(-16.93)	(-6.28)	(-16.91)
log(Age)	0.535***	0.932***	0.542***	0.926***
	(6.46)	(3.58)	(6.52)	(3.56)
Expense	-2.333***	-1.022**	-2.353***	-1.015**
	(-15.29)	(-2.52)	(-15.51)	(-2.50)
Turnover	-0.001	-0.004***	-0.001	-0.004***
	(-1.22)	(-3.70)	(-1.13)	(-3.70)
Observations	131,201	131,138	131,201	131,138
Year FE	No	Yes	No	Yes
Fund FE	No	Yes	No	Yes
Company FE	No	Yes	No	Yes
Adjusted R-squared	0.027	0.140	0.027	0.140

Table 5. The Implications of Risk Management for Future Performance: Sharpe Ratio

This table reports the relation between risk assessment measures and fund future performance, measured by Sharpe Ratio. *NegRisk (PosRisk)* captures managers' negative (positive) risk assessment in their narratives. Panel A shows the Sharpe Ratio for groups sorted by negative risk assessment. Funds are grouped into quintiles and the average Sharpe Ratio is reported for each group. The last column reports a two-sample *t*-test between the lowest and highest quintiles. Panel B shows regressions of future performance, measured by Sharpe Ratio, on risk assessment measures and fund characteristics. The dependent variable is the average daily returns over a 180-day period after a filing date divided by the standard deviation of the returns in the same period. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by fund. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Univariate sort on risk assessment

	SharpeRatio					
Quintile	Low	2	3	4	High	High - Low
NegRisk	0.76	0.76	0.85	0.80	0.89	0.13
						(9.86)

Panel B: Risk assessment and Sharpe Ratio

Dependent Variable	*SharpeRatio*			
	(1)	(2)	(3)	(4)
NegRisk	0.176***	0.144***	0.184***	0.186***
	(7.71)	(8.10)	(7.76)	(8.64)
PosRisk	-0.055***	-0.055***	-0.048**	-0.045***
	(-3.47)	(-4.43)	(-2.44)	(-3.20)
RiskWord			-0.034	-0.092***
			(-1.30)	(-4.61)
NegLM			0.021	0.009
			(0.58)	(0.32)
PosLM			-0.023	-0.029
			(-0.85)	(-1.34)
PastRisk	0.010***	0.016***	0.010***	0.016***
	(12.98)	(12.45)	(12.89)	(12.13)
PastAlpha	0.008***	-0.006***	0.008***	-0.006***
	(10.16)	(-9.45)	(10.15)	(-9.46)
log(TNA)	-0.046***	-0.151***	-0.046***	-0.151***
	(-9.12)	(-13.11)	(-9.12)	(-13.10)
log(Age)	0.162***	0.259***	0.162***	0.260***
	(14.74)	(9.35)	(14.84)	(9.33)
Expense	-0.377***	-0.439***	-0.377***	-0.437***
	(-16.69)	(-8.61)	(-16.71)	(-8.56)
Turnover	-0.000***	-0.000***	-0.000***	-0.000***
	(-3.87)	(-3.78)	(-3.90)	(-3.76)
Observations	131,156	131,093	131,156	131,093
Year FE	No	Yes	No	Yes
Fund FE	No	Yes	No	Yes
Company FE	No	Yes	No	Yes
Adjusted R-squared	0.016	0.298	0.016	0.298

Table 6. The Implications of Risk Management for Future Return Gap

This table shows regressions of future return gap on risk assessment measures and fund characteristics. *NegRisk (PosRisk)* captures managers' negative (positive) risk assessment in their narratives. The dependent variable *RetGap* is the monthly difference between a fund's realized gross return and the hypothetical return on its most recently disclosed portfolio holdings according to Kacperczyk, Sialm, and Zheng (2008). *RetGap6m/RetGap12m* is the average return gap in the next 6 months (12 months) after a filing date. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by fund. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	(1)	(2)	(3)	(4)
Dependent Variable	*RetGap6m*		*RetGap12m*	
NegRisk	0.020*	0.009	0.025**	0.027*
	(1.68)	(0.61)	(2.31)	(1.89)
PosRisk	-0.013	-0.019*	-0.017**	-0.019**
	(-1.36)	(-1.82)	(-2.44)	(-2.31)
RiskWord		0.019		-0.000
		(1.22)		(-0.03)
NegLM		0.004		-0.003
		(0.25)		(-0.19)
PosLM		0.020		0.008
		(1.39)		(0.68)
PastRisk	-0.002***	-0.002***	-0.000	-0.000
	(-2.83)	(-2.77)	(-0.08)	(-0.06)
PastAlpha	-0.002***	-0.002***	-0.002***	-0.002***
	(-2.65)	(-2.64)	(-4.42)	(-4.43)
log(TNA)	-0.039***	-0.039***	-0.037***	-0.037***
	(-4.44)	(-4.43)	(-4.32)	(-4.32)
log(Age)	0.038*	0.037	0.055**	0.055**
	(1.67)	(1.62)	(2.29)	(2.29)
Expense	-0.006	-0.007	-0.002	-0.002
	(-0.15)	(-0.17)	(-0.05)	(-0.06)
Turnover	0.000	0.000	0.000	0.000
	(1.12)	(1.11)	(0.31)	(0.31)
PastRetGap	-0.029**	-0.029**	-0.046**	-0.046**
	(-2.37)	(-2.37)	(-2.23)	(-2.23)
Observations	74,214	74,214	68,277	68,277
R-squared	0.110	0.110	0.198	0.198
Year FE	Yes	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Company FE	Yes	Yes	Yes	Yes
Adjusted R-squared	0.0673	0.0673	0.158	0.158

Table 7. Risk Assessment and Managerial Skill

This table shows regressions of change in the risk-taking, abnormal returns, and Sharpe ratio on risk assessment measures and fund characteristics for subsamples. *NegRisk (PosRisk)* captures managers' negative (positive) risk assessment in their narratives. Subsamples are partitioned based on whether *PastRetGap* is greater than zero, where *PastRetGap* is the past 6-month average difference between a fund's realized gross return and the hypothetical return on its most recently disclosed portfolio holdings according to Kacperczyk, Sialm, and Zheng (2008). All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by fund. In all panels, column (3) reports *z*-statistics in parentheses. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Panel A: Risk-taking

Dependent Variable	*ΔRisk*		
Subgroup	*PastRetGap* > 0	*PastRetGap* < 0	*Diff.of Coeff.*
	(1)	(2)	(3)
NegRisk	-1.279***	-0.565***	-0.714***
	(-7.10)	(-4.65)	(-3.29)
PosRisk	0.551***	0.307***	0.244
	(4.02)	(3.27)	(1.47)
RiskWord	1.326***	0.626***	
	(7.92)	(5.16)	
NegLM	-0.031	-0.036	
	(-0.17)	(-0.27)	
PosLM	-0.347**	-0.088	
	(-2.26)	(-0.86)	
PastRisk	-0.790***	-0.891***	
	(-96.98)	(-92.50)	
PastAlpha	0.025***	0.012**	
	(3.98)	(2.55)	
log(TNA)	0.205**	0.317***	
	(2.26)	(4.92)	
log(Age)	0.111	-0.574**	
	(0.32)	(-2.01)	
Expense	1.352***	1.566***	
	(2.97)	(4.36)	
Turnover	0.004***	0.002*	
	(2.84)	(1.78)	
Observations	36,753	47,912	
Year FE	Yes	Yes	
Fund FE	Yes	Yes	
Company FE	Yes	Yes	
Adjusted R-squared	0.670	0.689	

Panel B: Abnormal returns

Dependent Variable	*Alpha*		
Subgroup	*PastRetGap* > 0	*PastRetGap* < 0	*Diff.of Coeff.*
	(1)	(2)	(3)
NegRisk	0.814***	0.196	0.618*
	(2.81)	(1.01)	(1.77)
PosRisk	0.222	0.218*	0.004
	(1.14)	(1.67)	(0.02)
RiskWord	-0.271	-0.814***	
	(-1.02)	(-4.32)	
NegLM	-0.805***	-0.004	
	(-2.69)	(-0.02)	
PosLM	0.434*	0.092	
	(1.69)	(0.51)	
PastRisk	0.074***	0.050***	
	(5.08)	(3.14)	
PastAlpha	-0.039***	-0.041***	
	(-3.76)	(-4.26)	
log(TNA)	-2.144***	-2.007***	
	(-11.67)	(-13.29)	
log(Age)	1.805***	0.875*	
	(3.41)	(1.83)	
Expense	-1.532**	-0.907	
	(-2.06)	(-1.55)	
Turnover	-0.007***	-0.006***	
	(-2.92)	(-2.95)	
Observations	36,753	47,912	
Year FE	Yes	Yes	
Fund FE	Yes	Yes	
Company FE	Yes	Yes	
Adjusted R-squared	0.129	0.137	

Panel C: Sharpe ratio

Dependent Variable	*Sharpe Ratio*		
Subgroup	*PastRetGap* > 0	*PastRetGap* < 0	*Diff.of Coeff.*
	(1)	(2)	(3)
NegRisk	0.174***	0.117***	0.057
	(4.88)	(4.15)	(1.25)
PosRisk	-0.081***	-0.082***	-0.001
	(-3.35)	(-4.41)	(0.03)
RiskWord	0.082**	0.046	
	(2.49)	(1.56)	
NegLM	-0.156***	-0.088***	
	(-4.22)	(-2.62)	
PosLM	0.040	-0.041	
	(1.18)	(-1.48)	
PastRisk	0.038***	0.031***	
	(29.35)	(20.40)	
PastAlpha	-0.004***	-0.002**	
	(-4.54)	(-1.99)	
log(TNA)	-0.174***	-0.163***	
	(-10.52)	(-11.44)	
log(Age)	0.366***	0.301***	
	(6.71)	(5.24)	
Expense	-0.321***	-0.328***	
	(-4.49)	(-5.34)	
Turnover	-0.001**	-0.001***	
	(-2.44)	(-2.89)	
Observations	36,749	47,896	
Year FE	Yes	Yes	
Fund FE	Yes	Yes	
Company FE	Yes	Yes	
Adjusted R-squared	0.435	0.389	

Table 8. Risk Assessment and Exposure to Downside Risk

This table shows the relation between risk assessment measures and funds' exposure to the downside market. *NegRisk (PosRisk)* captures managers' negative (positive) risk assessment in their narratives. In the first two columns, the dependent variable is the downside beta of a fund's holdings after the disclosure. The first column uses the average stock-level downside beta of a fund's entire portfolio. The second column reports the downside beta of active portfolio change, measured as the average downside beta of purchased stocks minus the average downside beta of sold stocks. In the third column, the dependent variable is the percentage of days with negative market risk premium over the 180 days after a fund's disclosure. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by fund. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	(1)	(2)	(3)
Dependent Variable	*DownsideBeta*		% of Days with Negative Market Risk Premium
	Entire Portfolio	Buy - Sell	
NegRisk	-0.013**	-0.013*	-0.656***
	(-2.19)	(-1.68)	(-8.42)
PosRisk	-0.006*	-0.003	0.089
	(-1.69)	(-0.54)	(1.59)
RiskWord	0.003	-0.002	0.226***
	(0.68)	(-0.26)	(2.78)
NegLM	0.010	0.014*	0.445***
	(1.44)	(1.72)	(5.14)
PosLM	0.016***	-0.011	-0.671***
	(2.81)	(-1.40)	(-9.19)
PastRisk	0.006***	-0.001*	0.016***
	(23.77)	(-1.95)	(6.26)
PastAlpha	0.001***	0.000	-0.012***
	(6.12)	(0.01)	(-4.75)
log(TNA)	0.013***	-0.000	0.236***
	(3.60)	(-0.09)	(9.29)
log(Age)	-0.013	0.001	-0.774***
	(-1.19)	(0.11)	(-7.91)
Expense	-0.022	-0.026	0.670***
	(-1.30)	(-1.23)	(4.60)
Turnover	0.000**	-0.000	0.001***
	(2.03)	(-0.71)	(2.60)
Observations	68,145	68,145	68,145
Year FE	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes
Company FE	Yes	Yes	Yes
Adjusted R-squared	0.319	0.036	0.375

Table 9. The Implications of Risk Management for Future Morningstar Ratings and Morningstar Risk

This table shows regressions of Morningstar rating and its risk component on risk assessment measures and fund characteristics. *NegRisk (PosRisk)* captures managers' negative (positive) risk assessment in their narratives. *MS Rating* is the mutual fund rating published by Morningstar, and *MS Risk* is the risk component in the *MS* Rating. All variables are defined in Appendix A. *t*-statistics, in parentheses, are based on standard errors clustered by fund. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	(1)	(2)	(3)	(4)
Dependent Variable	*MS Risk*		*MS Rating*	
NegRisk	-0.002***	-0.001***	0.014	0.020*
	(-4.77)	(-4.15)	(1.42)	(1.71)
PosRisk	0.000	-0.000	0.027***	0.025***
	(0.74)	(-0.67)	(3.78)	(3.29)
RiskWord	0.002***	0.002***	0.002	-0.009
	(4.04)	(7.09)	(0.25)	(-0.90)
NegLM	-0.001**	-0.001**	-0.013	-0.018
	(-2.19)	(-2.44)	(-1.18)	(-1.36)
PosLM	0.001***	0.001***	-0.005	-0.035***
	(3.13)	(4.66)	(-0.60)	(-3.27)
PastRisk	0.001***	0.001***	0.001***	0.000
	(55.54)	(27.68)	(3.65)	(0.12)
PastAlpha	-0.000**	0.000	0.005***	0.004***
	(-2.38)	(0.21)	(13.05)	(12.11)
PastMSRisk	0.012*	0.001	-0.052	-0.023
	(1.66)	(0.71)	(-1.27)	(-0.56)
PastMSRating	-0.000	-0.000**	0.835***	0.710***
	(-1.60)	(-2.23)	(218.54)	(115.11)
log(TNA)	0.001***	0.001***	0.010***	-0.036***
	(8.76)	(3.77)	(4.64)	(-5.26)
log(Age)	-0.002***	-0.002**	0.004	-0.114***
	(-5.17)	(-2.44)	(0.69)	(-2.85)
Expense	0.006***	0.004***	-0.053***	-0.064*
	(16.57)	(4.03)	(-6.31)	(-1.75)
Turnover	0.000	0.000**	-0.000***	-0.000
	(1.20)	(2.34)	(-3.86)	(-0.97)
Observations	113,630	113,562	113,206	113,132
Year FE	No	Yes	No	Yes
Fund FE	No	Yes	No	Yes
Company FE	No	Yes	No	Yes
Adjusted R-squared	0.434	0.704	0.726	0.745

Table 10. Risk Assessment and Fund Flows

This table shows regressions of future fund flows on risk assessment measures and fund characteristics. *NegRisk (PosRisk)* captures managers' negative (positive) risk assessment in their narratives. *Sophisticated* is an indicator variable equal to one if a fund sells directly to investors, i.e., a fund with sophisticated investors, and zero if a fund sells shares through brokers, i.e., a fund with naïve investors. All variables are defined in Appendix A. Control variables include *PastRisk*, *PastAlpha*, *Log(TNA)*, *Log(Age)*, *Expense*, *Turnover*, and their interactions with *Sophisticated*. *t*-statistics, in parentheses, are based on standard errors clustered by fund. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	(1)	(2)	(3)	(4)
Dependent Variable	*Flow6m*		*Flow12m*	
NegRisk	-1.026	-0.755	-2.060	-1.348
	(-1.46)	(-0.91)	(-1.62)	(-0.90)
NegRisk × *Sophisticated*	2.325*	2.673**	4.113*	3.249
	(1.89)	(1.98)	(1.87)	(1.41)
PosRisk	-0.178	0.016	0.595	0.931
	(-0.38)	(0.03)	(0.75)	(1.08)
PosRisk × *Sophisticated*	0.647	0.362	-0.738	-1.134
	(0.84)	(0.40)	(-0.56)	(-0.74)
RiskWord		-1.160		-2.302**
		(-1.54)		(-2.03)
RiskWord × *Sophisticated*		1.069		2.700
		(0.85)		(1.39)
NegLM		0.696		0.999
		(0.77)		(0.63)
NegLM × *Sophisticated*		-2.027		-1.105
		(-1.18)		(-0.41)
PosLM		-0.910		-1.474
		(-1.19)		(-1.16)
PosLM × *Sophisticated*		1.699		1.683
		(1.17)		(0.70)
PastFlow	-0.061	-0.063	-0.197**	-0.202**
	(-1.20)	(-1.25)	(-2.51)	(-2.56)
PastFlow × *Sophisticated*	-0.106*	-0.103*	-0.287***	-0.286***
	(-1.76)	(-1.70)	(-2.64)	(-2.62)
Observations	129,866	129,866	128,068	128,068
Control variables	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Company FE	Yes	Yes	Yes	Yes
Adjusted R-squared	0.201	0.201	0.332	0.332

Table A1. List of Risk Words

This table lists all risk-related words from the Risk dictionary that appear in the shareholder reports in our sample.

List of Risk-related Words in Shareholder Reports				
abeyance	contingent	insecurity	probably	tentatively
almost	could	intangible	prospect	treacherous
alteration	crossroad	jeopardy	qualm	tricky
ambiguity	crossroads	likelihood	quandary	uncertain
ambiguous	debatable	may	query	uncertainly
ambivalence	depend	maybe	random	uncertainty
ambivalent	dependence	menace	randomly	unclear
anticipate	dependency	might	randomness	undefined
anticipated	dependent	misgiving	reassess	unfamiliar
anticipation	dicey	nearly	recalculate	unhedged
apparent	differ	niggle	recalculation	unknown
apparently	dilemma	occasionally	reconsider	unknowns
appear	disquiet	ordinarily	reexamination	unobservable
apprehension	dubious	pending	reexamine	unproven
approximate	exposure	perhaps	reservation	unquantifiable
approximately	exposures	perilous	revise	unquantified
approximation	fickleness	possibility	revised	unreliability
arbitrarily	fitful	possible	risk	unseasonably
arbitrary	fluctuate	possibly	riskiness	unsettled
assume	fluctuation	precarious	risks	unspecified
assumed	fluctuations	precaution	roughly	untested
assumption	gamble	precautionary	seldom	unusual
assumptions	halting	predict	skepticism	unusually
believe	hazy	predictability	sometime	vacillation
believes	hesitancy	prediction	sometimes	vague
bet	hesitant	predictive	somewhat	vaguely
cautious	hidden	predictor	somewhere	vagueness
cautiously	imprecise	preliminarily	speculate	variability
cautiousness	imprecision	preliminary	speculation	variable
chance	improbability	presumably	speculative	variance
changeable	improbable	presume	speculatively	variant
clarification	incertitude	presumed	sporadic	variation
conceivable	indecision	presumption	sticky	variations
conceivably	indecisive	probabilistic	sudden	varied
conditional	indefinite	probabilities	suddenly	vary
conditionally	indefinitely	probability	suggest	wager
contingency	indeterminate	probable	tentative	wariness