When Mutual Fund Managers Write Confidently

Qianzhou Du, Yawen Jiao, Pengfei Ye, and Weiguo Fan*

This draft: September 2019

Abstract

This paper studies the information content of the confidence level implied in the text. We use unsupervised machine learning techniques to generate a new lexicon for a writer's level of certainty in expressing opinions. Using it to capture mutual fund managers' textual confidence in letters to shareholders, we show confidence contains important information about the manager's skill: underperforming fund managers who write confidently significantly outperform other underperforming managers in the next six months. Further analyses reveal that our confidence measure is informationally distinct from other textual characteristics such as tone and overstatement, and outperforms human-based confidence measures in predicting performance. Underscoring the informational value of machine learning vs. human judgments, we also show capital flows do not respond to fund managers' confidence levels.

Keywords: Textual Analysis, Confidence, Mutual Funds, Skill

JEL Codes: G23, G41

* Du and Ye are at Virginia Tech. Jiao is at the University of California, Riverside. Fan is at the University of Iowa. Their email addresses are qiand12@vt.edu, yawenj@ucr.edu, pye@vt.edu, and weiguo-fan@uiowa.edu, respectively. We thank seminar participants at Virginia Tech and Nanjing University for helpful comments and discussion. All remaining errors are our own.

1. Introduction

Writing conveys not only facts but also the writer's belief. Using various textual analysis techniques, recent studies show human beliefs embedded in writing contain valuable forward-looking information. These studies, however, primarily focus on one specific belief element – the sentiment, typically measured by textual tone.[1] The information value of other elements, particularly the level of confidence (the level of certainty) expressed in the text, remains largely unexamined. In this study, we seek to fill this gap using mutual funds' letters to shareholders as a laboratory.

We propose a new lexicon to capture textual confidence using unsupervised machine learning methods. In the context of financial disclosures, the most relevant dimension of confidence is the expression of certainty for future states or events.[2] We use the *Over- and Under-statement* wordlists in the Harvard IV-4 Dictionary as our lexicon formation base – among all available wordlists, these lists are the most comprehensive in covering words describing the writer's level of certainty.[3] In spite of their comprehensiveness, these wordlists cannot be used directly to measure textual confidence for two reasons. First, like most IV-4 wordlists, they misclassify many business and finance-related words (e.g., "account"; "billion"; "capital") as words of textual characteristics (Loughran and McDonald, 2011). We carefully remove all business and finance-related words misclassified into these wordlists to minimize the confounding effects.

[1] Li (2010a), Das (2014), and Loughran and McDonald (2016) review textual analysis studies in finance and accounting. See, among others, Tetlock (2007), Engelberg (2008), Tetlock et al. (2008), Li (2010b), Loughran and McDonald (2011), and Jegadeesh and Wu (2012) for researches on the information content of tone. Kearney and Liu (2014) review the textual sentiment literature.

[2] See *New Webster's Dictionary*, *Oxford Dictionary*, and *Thesaurus of the English Language* (*NWDT*) for definitions of confidence.

[3] These wordlists can be found at http://www.wjh.harvard.edu/~inquirer/homecat.htm, under the Harvard IV-4 category No. 4: words indicating overstatement and understatement. See Section 2 for details.

Second, even after being carefully cleaned, the IV-4 over- and under-statement wordlists contain many words unrelated to the expression of certainty (or lack of certainty). In particular, they include words emphasizing the presence or lack of other semantic elements such as "*exceptionality, intensity, and extremity*". To overcome this multi-semantic problem, we employ an unsupervised machine-learning algorithm, the K-Means Clustering method, to categorize words in the cleaned IV-4 overstatement wordlist into three statistically distinct semantic clusters in the semantic space.[4] One group contains words such as "absolute", "exact", "assure", "confirm", "prove", "verify", etc. that can be used to express strong certainty, which we adopt as our confidence-indicating lexicon. Since this lexicon is generated by a machine learning algorithm without any human intervention, we refer to it as the *machine-based* confidence lexicon.[5] Using the same method, we also obtain a *machine-based* unconfidence lexicon from the cleaned IV-4 understatement wordlist.

We apply our machine-based confidence and unconfidence lexicons to capture the net level of confidence, i.e., the difference between the levels of confidence and unconfidence, expressed in mutual fund managers' letters to shareholders. Funds' shareholder letters typically cover two topics: past performance and future market conditions. Since the content tends to be homogeneous, *how* it is presented, as reflected in the textual characteristics of the letter, could be especially important. Hand-collecting data for 10,813 shareholder letters written by actively managed equity mutual funds for the period from 2006-2016, we examine whether the fund manager's level of

---

[4] Several unsupervised machine learning methods, including K-Means clustering, EM clustering with Gaussian mixture models, density based clustering, and hierarchical clustering, can be used to cluster data. We do not use the last two methods because the density based clustering method is more suitable for graphic data and the hierarchical clustering method is more suitable for merging clusters using a bottom-up approach. The K-Means and EM clustering methods lead to almost identical results; for brevity we only report the results from the former.

[5] For the other two groups, one group contains words of *emotional overstatement*, such as "absurd", "awful", "fantastic", and "incredible", and the other contains words related to the *intensity* of expression, such as "almost", "enough", "every", "important", and "most". We also examine the information contents of these two word groups.

textual confidence can predict future performance. Net confidence is measured using both the proportional method and the tf.idf weighting scheme (Jurafsky and Martin, 2009). The latter adjusts for the frequency of each word used in a letter and across all sample letters, thereby mitigating the distortions caused by high-frequency words.[6]

Ex-ante, conflicting motives for expressing confidence in shareholder letters could lead to weak predictive power for future performance: fund managers could either inform or mislead investors through writing confidently. For a fund manager who currently performs poorly, if the bad performance results from bad luck, she would anticipate better future performance. Writing confidently might help her signal this private knowledge – the behavior might not even be intentional, as the ability to remain confident and seek improvements when facing temporary difficulties can be an important dimension of true skill. A manager who currently performs well because of luck (rather than skill), on the other hand, might write confidently either because she misjudges her own capabilities or intends to exploit the fleeting luck to attract greater capital flows. When both types of fund managers write similarly, textual confidence can lose the power to forecast future performance. Indeed, we find that the net confidence itself cannot predict future alphas when funds are pooled together.

Once the fund's past performance is taken into account, the predictive power of textual confidence for future performance improves sharply. We partition funds into three groups: the bottom- (top-) performers include funds whose six-month returns before the letter release month are ranked in the bottom (top) cross-sectional quartile and the remaining funds are classified as medium-performers. We use the Fama-French-Carhart four-factor alpha (Carhart, 1997) net of expense in the six-month period following the letter release to measure a fund's future performance.

[6] The two measures generate similar results in most cases and we rely on the tf.idf score method when the results defer.

Among bottom-performers, greater net confidence in shareholder letters firmly predicts greater future alphas – a two-standard-deviation increase in it is associated with a 1% increase in annualized future alphas. This result is more pronounced for annual shareholder letters, and remains after we control for funds' past performance to account for performance reversals, fund characteristics such as size, age, expense ratio, and portfolio turnover ratio, and fund fixed effects.

A natural question is whether unskilled bottom-performers have the incentive or ability to mimic the skilled ones in expressing confidence in shareholder letters. First of all, unskilled managers' incentive to mimic depends on whether textual confidence can mitigate capital outflows, which we find not to be the case in later analysis. In other words, there is no clear benefit for mimicking. Second, writing confidently with accuracy and assurance, especially when facing difficulties related to poor performance, e.g., capital outflows and declines in compensation and reputation, requires superior knowledge and genuine faith in one's capability. Not all unskilled fund managers have these qualities, making mimicking difficult.[7]

Besides the level of certainty in expressing opinions, the human brain can interpret several other communication styles as indicating confidence, e.g., the feeling of being powerful or assured (Barbalet, 1993).[8] To explore other writing styles that might be perceived by humans as indicating confidence, we form alternative lexicons using Amazon's MTurk platform. For each word in the cleaned IV-4 over- (under-) statement wordlist, we invite five eligible participants to judge whether it indicates the fund manager's confidence (lack of confidence) in expressing opinions without being overly emotional – the word is included in the human-based confidence (unconfidence) lexicon if at least three participants strongly agree it does. The human-based

[7] Lucky top-performers might enjoy a temporary and artificial boost of confidence that can show in shareholder letters.
[8] People sometimes confuse confidence with optimism. In cognitive psychology, optimism is specifically related to the estimation level itself, i.e., the textual sentiment, which we treat as a separate textual characteristic in our study.

confidence lexicon turns out to be broader than the machine-based one: In addition to words related to certainty, MTurk participants also perceive many words about the intensity of expression, e.g., "important" and "significant", to be confidence-indicating. Indeed, more than half of the words in the human-based confidence lexicon are related to intensity rather than certainty. Despite its broadness, human-based net confidence is not as informative as the machined-based one in that it cannot predict future fund performance.

In an important paper, Loughran and McDonald (LM, 2011) compile a *Strong (Weak) Modal* wordlist. Some words in them, such as "always", "definitely", and "undisputed", can be used to express assertiveness or confidence. These wordlists are short with only 19 words in the *Strong Model* list and 27 in the *Weak Modal* list. Consequently, they might underestimate textual confidence in formal writing. Empirically, net confidence defined using these wordlists does not predict future fund performance.

We further find that the performance predictability of the machine-based net confidence measures is driven by *confidence*, while the informational value of *unconfidence* is weak. Our machine- and human-based methods generate two textual unconfidence lexicons and we also use the LM *Weak Modal* and *Uncertainty* wordlists to augment our tests. Regardless of the lexicon used, textual unconfidence cannot forecast future fund performance. This finding might stem from noises in unconfidence lexicons or pooling of skilled and unskilled fund managers who tend to write in a reserved manner – further research of them in future studies could be fruitful. In light of unconfidence measures' lack of performance predictability, we focus on confidence measures in the remaining tests.[9]

[9] The results are qualitatively the same if using net confidence measures (available upon request).

Some might misinterpret a hyperbolic writing style (e.g., the uses of words such as “awful”, “fantastic”, and “incredible”) as indicating confidence. We use the K-Means Clustering method to identify words of an emotionally overstating writing style. Our analyses show confidence and overstatement convey distinct signals about future fund performance: overstating funds, particularly the top performers, tend to have worse future performance, suggesting that lucky yet unskilled fund managers might hype their funds to exploit the fleeting luck.[10] We also find that an emotionally understating writing style is associated with better future performance for top performers. None of the results related to over- or under-statement affect the performance predictability of textual confidence. Moreover, we show confidence and tone/sentiment are informationally distinct from each other. While tone contains useful information about funds’ future performance, it does not affect the predictive power of confidence.

In spite of its predictive power for funds’ future performance, machine-based textual confidence is ineffective in influencing investors. Using capital flows as a proxy for investors’ reactions to shareholder letters, we find that neither machine-based nor human-based confidence measures lead to abnormal flows, suggesting that investors have limited ability to detect the information embedded fund managers’ expressions of confidence. These findings are consistent with the advantage of machine-based methods vs. human judgments in identifying non-public information in funds’ shareholder letters. The lack of investor reaction also lowers the likelihood for fund managers to strategically show confidence.

To the best of our knowledge, this paper is the first to examine the relevance and informational role of textual confidence (certainty) in financial disclosures. The existing textual

[10] Note that our “over-” and “under-statement” lexicons are subsets of the IV-4 wordlists that we refer to as the “IV-4 overstatement” and “IV-4 understatement” lexicons or their cleaned versions, which we refer to as the “cleaned IV-4 overstatement” and “cleaned IV-4 understatement” lexicons.

analysis literature primarily focuses on readability and tone – the former emphasizes ease of communication and the latter focuses on textual sentiment.[11] Compared to these textual characteristics, textual confidence reflects the writer's self-assessment of how accurately she can evaluate current events and forecast future. It also has closer relations to the writer's motives. By highlighting this new dimension of communication, we contribute to the growing literature on textual analysis and non-descriptive soft information.[12]

We are also the first to explore the differences in various machine- and human-based confidence lexicons. Our evidence shows not all forms of perceived confidence are equally informative: Among the three aspects we examine, specifically, certainty, intensity, and overstatement, both certainty and overstatement have predictive power for future performance, albeit in distinctive ways. Confidence is a key personality characteristic associated with leadership – our research contributes to the related literature by proposing new confidence measures and empirically evaluating and comparing them.

Last but not least, by examining mutual funds' shareholder letters, we also offer a novel angle to separate "skill" and "luck". Separating skill from luck is a central theme of the mutual fund performance literature (e.g., Fama and French, 2010).[13] Unlike corporate executives who routinely engage in written and verbal communications with investors, analysts, and even the general public, fund managers' communications occur primarily through annual or semi-annual reports. By presenting evidence on the relations between skill and a comprehensive set of textual characteristics, we identify writing style as a useful new tool for differentiating skill from luck.

---

[11] See, e.g., Li (2007), Lehavy et al. (2011), Loughran and McDonald (2014a), Lundholm et al. (2014), and Hwang and Kim (2017) for studies examining readability.

[12] See Liberti and Petersen (2018) for a review of the hard and soft information literature.

[13] A large literature shows fund/manager characteristics can reveal skill. For example, Kacperczyk, Sialm, and Zheng (2005, 2008), Kacperczyk and Seru (2007), Cremers and Petajisto (2009), among others, show funds' portfolio composition or portfolio-inferred trading offer insights into outperforming funds. Chevalier and Ellison (1999) show fund manager characteristics such as the SAT score can predict performance.

The remainder of the paper is organized as follows. Section 2 discusses confidence lexicons and measures of textual confidence. Section 3 describes data, sample selection, and key variables. Section 4 presents evidence on the relationship between textual confidence in mutual funds' shareholder letters and future fund performance. Section 5 examines fund investors' reactions to textual confidence. Section 6 concludes.

2. Measuring Confidence in Business and Finance Writing

We describe the methods we use to measure the level of textual confidence in the business and finance writing context in this section. Drawing on prior studies (e.g., Tetlock, 2007; Tetlock et al., 2008; Loughran and McDonald, 2011), we employ a "bag of words" approach that ignores the grammar or order of words in the text and relies on a pre-determined lexicon to capture qualitative text characteristics. We examine several lexicons that potentially can capture the writer's level of confidence in expressing opinions and discuss two methods to compute textual confidence based on these lexicons.

2.1 Existing Lexicons Related to Confidence/Unconfidence

Several existing lexicons have the potential to capture a writer's confidence or unconfidence in writing. For example, Loughran and McDonald (2011) compile a *Strong Modal* wordlist and a *Weak Modal* wordlist. The *Strong Modal* wordlist includes words such as "always", "definitely", "unambiguously", etc. that can indicate a writer's assertiveness or confidence. The *Weak Modal* wordlist includes words such as "appear", "depend", "suggest", etc. that can indicate unconfidence. Both wordlists contain words with a strong sense of confidence/unconfidence, yet they are very short. The strong (weak) modal wordlist includes 19 (27, 21 if only counting root words) words. Consequently, they could underestimate the actual confidence level in the text and weaken its informational value. We propose two sets of confidence/unconfidence lexicons that are more comprehensive and discuss them below.

2.2 Cleaning the IV-4 Over- and Under-statement Wordlists

We posit that the "*Overstatement*" and "*Understatement*" wordlists in the Harvard IV-4 (IV-4 hereafter) Dictionary can be a useful starting point for expanding the confidence/unconfidence lexicons. These wordlists include 576 overstatement words and 275 understatement words.[14] According to the IV-4, the *Overstatement* wordlist includes words indicating an "*emphasis in realms of speed, frequency, causality, inclusiveness, quantity or quasi-quantity, accuracy, validity, scope, size, clarity, exceptionality, intensity, likelihood, certainty, and extremity*." In financial disclosures, words emphasizing *accuracy*, *causality*, *certainty*, *clarity*, and *validity* can all be used to indicate a writer's sense of certainty, a key element of confidence, in assessing situations or expressing opinions. Words emphasizing *exceptionality*, *inclusiveness*, *intensity*, and *extremity* might also be used to project an impression of confidence. The *Understatement* wordlist includes words that indicate "*de-emphasis and caution*," which can be interpreted as indicating a lack of confidence.

Misclassification problems make the IV-4 wordlists unsuitable to use without refinements in our context. As noted by Loughran and McDonald (2011), designed for sociology and psychology research, these wordlists suffer from misclassifications of business- and finance-related words. For example, in tone-related wordlists, words such as "liabilities" and "tax" are classified as negative tone words, though their meanings are neutral in business writing. Panel A

[14] The full wordlists can be found at http://www.wjh.harvard.edu/~inquirer/homecat.htm.

of Table 1 reports the top 30 IV-4 over- and under-statement words in funds' shareholder letters in our sample. Words such as "total", "capital", "security", "account", "due", "point", and "basis", which are common neutral terms in financial disclosures, are all classified as overstatement words; similarly, words such as "time", "number", and "event" are classified as understatement words, although they do not convey any sense of de-emphasis or caution in business writing.

We carefully identify the misclassified words and remove them from the IV-4 over- and under-statement wordlists. In addition to the above examples, we also remove "commission", "fundamental", "patent", "principal", "reserve", "save", "worth", and "wealth". Words related to quantity and frequency such as "hundred", "thousand", "million", "billion", "trillion", "once", "daily", and "everyday", are also removed because they are a common yet neutral part of financial disclosures. Twelve words included in both IV-4 over- and the under-statement wordlists, including "all", "doubt", "few", "large", "least", "matter", "part", "short", "simple", "rare", "rather" and "usual", are removed as well to avoid confusion. Finally, we remove words that are 1) common phrases in general discussions, such as "above", "after", "away", "do", "etc", and "what", 2) conceptually misclassified by IV-4: e.g., "permanent" and "wondrous" are misclassified as understatement words, and 3) irrelevant, such as "cohort", "earth", "hand", "laugh", "roundabout", and "snack".[15] The complete lists of removed words are in Appendices B and C and the removed top 30 words are highlighted in grey in Panel A of Table 1.

2.3 The Machine-Based Confidence Lexicon

[15] 17 and 12 common and irrelevant words are removed from the IV-4 over- and under-statement wordlists, respectively. We are particularly cautious in removing irrelevant words. Irrelevant words that are rarely used, such as "anarchy", "anti-social", and "heritage", are kept to reduce biases associated with human judgments. They are highlighted in *italic* in Appendices. The results are unchanged if we remove the rarely used irrelevant words and if we do not remove any irrelevant words.

Another problem of using the IV-4 over- and under-statement wordlists to measure textual confidence/unconfidence is of semantics. IV-4 assigns over 15 semantic dimensions (from *speed* to *extremity*) to overstatement words – some of them are unrelated to confidence while others correspond to different dimensions of confidence reflecting different motives and traits of the writer. For example, a letter with words indicating certainty and assurance could communicate very different information than a letter with words of extremity, yet both can leave the reader with an impression of confidence. Mixing all words in one lexicon could introduce noise in measuring textual confidence, thereby confounding our empirical analyses.

We resort to unsupervised machine-learning techniques to solve the multi-semantic problem. These techniques are transparent, replicable, and suffer from fewer biases related to human judgments than the traditional textual clustering methods. Machine learning research has developed several methods for clustering data, with some of the widely used ones being the K-Means clustering method, the expectation maximization (EM) clustering method with Gaussian mixture models, the density based clustering method, and the hierarchical clustering method. We omit the last two because the density based clustering method is more suitable for graphic data and the hierarchical clustering method is more suitable for merging clusters using a bottom-up approach. The K-Means and EM clustering methods have their own benefits and limitations, respectively. For example, the K-Means method is fast to implement but assumes spherical variances for clusters, while the EM method can output the probability of belonging to a cluster for each data point but needs the Gaussian assumption. Empirically, we find that the two methods lead to almost identical results in our setting. In the rest of the paper, we only report the results from the K-Means clustering method for brevity and because this method has a longer application history – the results from the EM clustering method are available upon request.

The K-Means Clustering method follows the same mathematical principle of the statistical cluster analysis to reclassify words into subgroups based on their mutual distances in the semantic space, with words in the same subgroup (cluster) having similar semantic values compared to words in other subgroups. In doing so, words with diverse semantic dimensions can be grouped into multiple homogeneous subgroups, a task humans would find challenging in many cases. To implement this method, we use the GloVe algorithm of Pennington, Socher, and Manning (2014) to vectorize each over- or under-statement word in a 100-dimension semantic space, and run the K-Means Clustering analysis for the cleaned IV-4 over- and under-statement wordlists separately.[16] We are agnostic about the value of K (i.e., the number of clusters) and try K=2, 3, 4, 5, and 6. Subsequently, we examine the resulting clusters under each K value and identify the one that makes the most semantic sense in business writing, which occurs for K=3.[17] In other words, the cleaned IV-4 over- (under-) statement lexicon is divided into three word groups that are statistically different from each other in the semantic space.

Among the three subgroups generated by this algorithm for the overstatement wordlist, one is closely related to the sense of certainty, a key dimension of confidence, especially for financial disclosures in that certainty reflects a manager's strong faith in her ability to evaluate or influence events and situations. This word group includes 119 words, e.g., "assure", "absolute", "accurate", "exact", "admit", "confirm", and "prove" (see Appendices B and D for the complete list), and we use it as our new confidence lexicon. Because it is generated without human intervention, we call this lexicon the *machined-based* confidence lexicon. Panel B of Table 1 reports the top 30 *machine-based* confidence words used in shareholder letters of funds in our sample, which account for 71% of all machine-based confidence words used. The top five words are "confidence",

---

[16] Details and applications of this algorithm in textual analysis can be found at https://nlp.stanford.edu/projects/glove/.
[17] The K-Means Clustering results for K=2, 4, 5, and 6 are available upon request.

"absolute", "determine", "prove", and "exceed", accounting for 26% of all machine-based confidence words used.

The K-Means Clustering method also generates two other word groups. One includes 221 words that convey a sense of emotional exaggeration (overstatement), such as "absurd", "aghast", "exaggeration", "extravagant", "glorious", "fantastic", and "monstrous". We refer to them as the *Overstatement* lexicon. Note that our over- and under-statement lexicons are subsets of the cleaned IV-4 wordlists that we refer to as the "IV-4" over- and under-statement lexicons. The other group includes 182 words related to the *intensity* one can have in describing events or expressing opinions, such as "abundant", "almost", "always", "entire", "complete", "great", "high", "intense", "significant", and "strong". We refer to them as the *Intensity* lexicon. See Appendix B for the complete lists of the three lexicons described above. We also use the K-Means Clustering method to partition the cleaned understatement wordlist into three groups – the machine-based unconfidence lexicon, the lack-of-intensity lexicon, and the emotional-understatement lexicon.[18]

Unsupervised machine learning techniques are transparent and rely little on assumptions and human judgments, but are by no means perfect. For example, the K-Means Clustering method includes the word "complexity" in the confidence lexicon and "complex" in the intensity lexicon, even though the two words have very similar meanings. Moreover, this method is somewhat ineffective in reclassifying understatement words, which we will discuss more in Section 2.5. Note that these noises tend to bias against finding strong predictability using the machine-based lexicons. In spite of its drawbacks, we believe this method is an innovative and promising approach to address the multi-semantic problem in textual analysis.

[18] The machined-based unconfidence lexicon includes words such as "appear", "maybe", and "seem". The lack-of-intensity lexicon includes words such as "adequate", "gradual", "unfavorable", "unspecified", and "weak". The emotional-understating lexicon includes words such as "accustom", "conceivable", and "disputable". See Appendix C for the complete lists.

2.4 The Human-Based Confidence Lexicon

We also create new sets of human judgment based confidence and unconfidence lexicons. To this end, we conduct a human intelligence task (HIT) using Amazon's Mechanical Turk Services (MTurk).[19] MTurk invites eligible participants to perform tasks for pre-specified payments and has become a popular platform to conduct research requiring a large number of human inputs. Our HIT consists of 20 words randomly selected from either the cleaned IV-4 overstatement or the cleaned IV-4 understatement wordlist. After a brief review of the background of the project, we ask each qualified MTurk participant whether he/she agrees that each of the 20 words can indicate a fund manager's confidence (or lack of confidence) in expressing personal opinions without sounding too emotional. Figure 1 provides a snippet of our HIT.

To ensure the highest quality, we only invite MTurk Masters, a group of elite certified participants who have high approval ratings based on their track records to join our project.[20] Each participant receives $1 for successfully finishing the task. To minimize noise, we require each word to be judged by five participants. A word is included in our human-based confidence (unconfidence) lexicon if at least three participants *strongly agree* that it can indicate confidence (unconfidence) in mutual funds' shareholder letters. The human-based confidence lexicon includes 110 words and the unconfidence lexicon includes 49 words. The complete lists can be found in Appendices D and E.

[19] See www.mturk.com. IRB approvals were obtained before conducting the task.
[20] According to the self-reported demographic information, 43% of participants are female and 57% are male. 19% of participants are 20-30 years old, 41% are 30-40 years old, 24% are 40-50 years old, and 15% are more than 50 years old. All participants are native English speakers and all have at least a college degree. 82% have investment experience and 69% have experience in investing in mutual funds.

Words related to certainty can strongly indicate confidence, as 58 (out of 110) words in the human-based confidence lexicon overlap with the machine-based confidence/certainty lexicon. Human perception of confidence also goes beyond the sense of certainty and slightly tilts toward the *intensity* aspect – the human-based confidence lexicon includes many intensity related words such as “abundant”, “complete”, “entire”, “every”, “huge”, and “significant”. Panel B of Table 1 lists the 30 most used human-based confidence words in our sample, accounting for 79% of all human-based confidence words used. The top five words are “lead”, “significant”, “important”, “confidence”, and “major”; among them, only “confidence” is included in the machine-based confidence lexicon and all the other four are in the machine-based intensity lexicon.

2.5 Machine- and Human-based Unconfidence Lexicons

Using the cleaned understatement wordlist, we repeat our previous procedures to obtain machine- and human-based unconfidence lexicons. The machine-based unconfidence lexicon appears to be noisy. For example, words indicating the writer’s lack of certainty, such as “cautious”, “relative”, and “suggest”, get classified into other word groups. Words related to the lack of intensity, such as “less”, “small”, and “little”, are classified into the unconfidence (uncertainty) lexicon. As can be seen from Panel C of Table 1, among the top 30 most used machine-based unconfidence words in our sample, many words, such as “well”, “small”, “less”, “care”, “several”, and “slow”, are related to the lack of intensity rather than uncertainty.

In contrast, humans appear to be rather good at identifying unconfidence-related words. The human-based unconfidence lexicon, reported in Appendices D and E, contains a set of words that are focused and relevant. Panel C of Table 1 reports the top 30 most used human-based unconfidence words in our sample; the top 5 words are “relative”, “less”, “appear”, “uncertainty”, and "seem". These words are not only quite different from those in the machine-based lexicon (only 15 overlapping words) but also are more related to uncertainty.

To ensure our results related to unconfidence are not caused by the poor quality of machine-based unconfidence lexicon, in addition to the human-based unconfidence lexicon, we also use Loughran and McDonald's (2011) *Uncertainty* wordlist as an alternative lexicon. Intending to capture the level of (objective) uncertainty *facing* the company/fund, many words in this wordlist can also be used to express the writer's lack of confidence.

2.6 The Confidence/Unconfidence Measures and Other Textual Variables

We use both the proportional ratio and the tf.idf score to measure the confidence/unconfidence level in a fund manager's letter to shareholders. Both are based on the number of words identified by our confidence (unconfidence) lexicons. Formally, for the fund letter *i* disclosed at the report release month *t*, its *proportional confidence* level is defined as:

$$Confidence_{i,t} = \frac{\#\,Words\ Indicating\ Confidence\ in\ Letter_{i,t}}{\#\,Non\text{-}stop\ Words\ in\ Letter_{i,t}}. \tag{1A}$$

Its *proportional unconfidence* level is defined as:

$$Unconfidence_{i,t} = \frac{\#\,Words\ Indicating\ a\ Lack\ of\ Confidence\ in\ Letter_{i,t}}{\#\,Non\text{-}stop\ Words\ in\ Letter_{i,t}}. \tag{1B}$$

Its *proportional net confidence* level is defined as:

$$Net\ Confidence_{i,t} = Confidence_{i,t} - Unconfidence_{i,t}. \tag{1C}$$

The proportional measure could be biased because it implicitly assumes words in the same lexicon are equally important in terms of information content. It also gives high-frequency words much greater weights, whereas the information content of a word used 100 times cannot be a hundred times greater than that of a word used just once. To mitigate these biases, we use a weighting scheme developed by Jurafsky and Martin (2009), the tf.idf weighing scheme, to refine the confidence/unconfidence measures. Specifically, for each confidence or unconfidence word j, its weighted value in the letter i at time t, or its tf.idf score, $W_{i,j,t}$, is defined as:

$$W_{j,i,t} = \frac{1+\log(tf_{j,i,t})}{1+\log(a_{i,t})}\log\frac{N}{df_j}. \tag{2A}$$

In equation (2A), $tf_{j,i,t}$ is the count of the word j in letter i, a_i is the number of non-stop words in the letter, N is the total number of shareholder letters in the sample, and df_j is the number of sample letters that contain the word j. The tf.idf confidence, unconfidence, and net confidence scores of the letter i disclosed at month t are defined as:

$$tf.idf\ Confidence_{i,t} = \sum_{j\in confidence} W_{j,i,t}; \tag{2B}$$

$$tf.idf\ Unconfidence_{i,t} = \sum_{j\in unconfidence} W_{j,i,t}; \tag{2C}$$

$$tf.idf\ \ Net\ Confidence_{i,t} = tf.idf\ Confidence_{i,t} - tf.idf\ Unconfidence_{i,t}. \tag{2D}$$

Because of the logarithmic transformation in the tf.idf weighting scheme, the scale of the tf.idf score is very different from that of the proportional construction. More importantly, to the extent that the tf.idf measures correct the drawbacks of the proportional measures, we trust the tf.idf measures more when the results based on the two construction methods differ.

In addition to confidence, unconfidence, and net confidence, we also construct several other lexicon-based textual characteristic variables including modal (strong, weak, and net), uncertainty (as an alternative measure of unconfidence), overstatement (overstatement, understatement, and net), and tone (positive, negative, and net), in similar manners. To measure modal, uncertainty, and tone, we use the lexicons provided by Loughran and McDonald.[21] To measure overstatement, we use the lexicons related to emotional over/understatement identified by the K-Means Clustering method. We analyze these variables either because they can be

[21] We thank Tim Loughran and Bill McDonald for providing the tonal and modal wordlists. Our wordlists do not include all inflections for nouns and verbs. We use the lemmatization process to remove inflectional endings of nouns and verbs in lexicons and text, and apply the lemmatized lexicons to lemmatized letters to construct textual variables.

alternative measures of confidence (e.g., strong modal) or unconfidence (e.g., weak modal or uncertainty), or because they are of interest in our context (e.g., overstatement and tone).

Finally, we also construct two non-lexicon based textual characteristic variables: article length and readability. Article length is the number of non-stop words in the letter. Readability is measured using the Flesch–Kincaid Grade Level score (1948).[22] This score measures the number of education years generally required to understand the text. Because these two variables are not the focus of this study, we use them as control variables.

3. Data, Sample, and Key Variables

3.1 Mutual Funds' Shareholder Letters

All registered investment companies, including mutual funds, are required to file shareholder reports (Form N-CSR) with the Securities and Exchange Commission (SEC) every six months. These reports usually, although not always, start with a letter to shareholders written by fund managers. Some fund managers, instead of writing a formal letter, discuss similar topics in a designated section typically titled "Management's Discussion of Fund Performance."[23] Our shareholder letter sample includes both formal shareholder letters and management discussions. The latter is relatively uncommon, accounting for 5% of our sample letters. Our results are unaffected if the informal discussions are excluded.

We download all N-CSR forms for the period from 2006 to 2016 from the SEC's EDGAR database and use the characteristics of a common letter to extract shareholder letters.[24]. Specifically,

[22] Flesch-Kincaid score is computed as $= 0.39 * \frac{Total\ Words}{Total\ Sentences} + 11.8 * \frac{Total\ Syllables}{Total\ Words} - 15.97$. Results using alternative readability measures such as Fog Index or Flesch Index are very similar.
[23] For example, most Fidelity funds do not include formal shareholder letters in shareholder reports and discuss past performance in "Management's Discussion of Fund Performance" instead.
[24] Our sample period is subsequent to the SEC's 1998 rule requiring firms to use plain English in filings (to enhance the readability of disclosures), and therefore not subject to its confounding effects (Loughran and McDonald, 2014b).

we define a segment of the text as a shareholder letter if: (1) it starts with a formal addressee including terms such as "dear shareholders," "dear fellow," "dear stockholders," "to our shareholders," and "to our stockholders;" (2) it ends with terms such as "respectfully submitted," "sincerely," "cordially," "best regards," "respectfully," "thank you," "regards," or "very truly yours." For reports without a letter defined in the above manner, we examine whether they contain a separate section with a title containing the keywords "management discussion"; the text between this section and the next section is treated as the shareholder letter if available. When the above procedure does not yield any outcome, the file is regarded as not containing a shareholder letter and excluded from the sample. From the 200,201 N-CSR forms downloaded from the SEC's Edgar database, we obtain 121,545 well-defined shareholder letters and management discussions. Each letter is identified by the fund ticker and release date (i.e., the date on which the N-CSR form is filed with the SEC).

We focus on shareholder letters written by actively managed domestic equity mutual funds covered by both the CRSP Survivor-Bias-Free Mutual Fund Database and Thomson Reuters' S12 Mutual Fund Portfolio Holdings Database.[25] We also require that each sample fund must be at least two years old to mitigate the incubation bias documented by Evans (2010) and have more than $15 million in total net assets. This sample selection procedure yields 10,813 shareholder letters of 1,477 actively managed equity funds from 2006-2016, approximately 1,000 letters per year.

3.2 Contents of Mutual Funds' Shareholder Letters

[25] Specifically, we rely on CRSP's Objective Code (*crsp_obj_cd*) to identify actively managed domestic equity funds – funds included in this study are those with objective codes of EDY or EDC. Sector funds, index funds, and exchange-traded funds are excluded.

Mutual funds are considerably more homogeneous than corporations in business model and accounting rules. Consequently, the contents of funds' shareholder letters in our sample turn out to be very similar, in that fund manager almost always focus on reviewing past performance and selectively discussing winning or losing bets, and expressing opinions about future economic and market conditions. The letter must be certified by the mutual fund company's principal executive and financial officers.

We use an unsupervised Bayesian machine learning method, the Latent Dirichlet Allocation method (LDA), to verify the above observation.[26] Developed by Blei, Ng, and Jordan (2003), this method is conceptually analogous to the statistical factor analysis in that its outcome contains multiple strings of words (textual factors), with each string representing a specific machine-recognized topic. Using the "perplexity" criterion, we identify two as the appropriate number of topics for fund letters in our sample, among the choices of one to five. Each topic is represented by 50 keywords. The first topic contains words such as "market", "sector", "rate", "equity", "economy", "economics", etc., which we refer to as "market discussion". The second topic contains words such as "performance", "return", "expense", "security", "holding", etc., which we refer to as "performance discussion". To summarize, the LDA analysis shows the contents of mutual funds' shareholder letters are rather focused and homogeneous. Compared to corporate disclosures that tend to cover a much wider range of topics (e.g., Dyer, Lang, and Stice-Lawrence, 2017), they are more suitable for our purpose - when contents are similar, *how* they are presented could be particularly important.

3.3 Characteristics of Sample Letters

[26] See Dyer, Lang, and Stice-Lawrence (2017) for more details about the LDA method.

Panel A of Table 2 reports the summary statistics of our sample shareholder letters. All variables in this study are winsorized at the 1st and 99th percentiles. On average, shareholder letters in our sample have 1,248 non-stop words with a median of 858. The average Flesch-Kincaid readability score is 13.7 with a median of 13.5, i.e., 14 years of formal education are needed to understand the letters.

We report two sets of summary statistics for textual confidence, modal, uncertainty, overstatement, and tone variables, one for the proportional measures and one for the tf.idf measures. Since the proportional measures have greater intuitive appeal and the two sets of measures have similar distributions, we mainly discuss the proportional measures for brevity.

Our sample shareholder letters on average have a machine-based confidence level (denoted by *confidence-M*) of 0.0068, a human-based confidence level (denoted by *confidence-H*) of 0.0157, and a strong-modal based confidence level (denoted by *LM strong modal*) of 0.0033. Among these confidence variables, the human-based one has the highest mean because its underlying lexicon contains many highly frequent words related to intensity, such as "significant", "important", etc. The modal-based confidence measure has the lowest mean because its underlying lexicon is very small. The correlations among these three variables are modest, as shown in Panel B of Table 2. The correlation between *confidence-M* and *confidence-H* is 0.61 and that between *confidence-M* and Strong Modal is just 0.38.

Table 2 also reports the summary statistics of the unconfidence level of our sample letters. The average unconfidence level is 0.0062 for the machine-based measure (denoted by *unconfidence-M*), 0.0071 for the human-based measure (denoted by *unconfidence-H*), and 0.0066 for the modal-based measure (denoted by *LM weak modal*). We also use the LM uncertainty wordlist to construct an alternative unconfidence measure, denoted by *LM uncertainty*, and its mean is 0.0212, which is greater than means of other unconfidence measures, reflecting the fact that this lexicon has the most words among all unconfidence lexicons. Panel C of Table 2 reports the correlations among the unconfidence variables. They are also relatively low: for example, the correlation is 0.55 between *unconfidence-H* and *unconfidence-M*, 0.17 between *unconfidence-H* and *LM weak modal*, and 0.31 between *unconfidence-H* and *LM uncertainty*.

The means of confidence and unconfidence variables of our sample shareholder letters appear to be small. Note that this is not surprising because as a legal document, a shareholder letter with too many confidence words can increase legal risk (Rogers, Van Buskirk, and Zechman, 2011). We will show in Section 4 that a small dose of confidence words can be powerful in predicting future fund performance.

Coming to over- and under-statement variables, the mean overstatement (understatement) level is 0.0005 (0.0003). Both are low, which is, again, not surprising, because we do not expect professional fund managers to be overly emotional in formal writing. In contrast, the means of positive and negative tone variables are much higher, both about 0.03, resulting in a mean close to zero for net tone. Panel D of Table 2 presents the pair-wise correlations between these variables and confidence variables. The correlation between net confidence and net overstatement is only 0.03 and that between net confidence and net tone is 0.17. These findings suggest that confidence, overstatement, and tone represent three distinct dimensions of mutual funds' shareholder letters.

3.4 Funds Characteristics

Panel E of Table 2 presents the summary statistics for our sample funds. CRSP Mutual Fund Database reports fund characteristics at the class level. For multi-class funds, we combine the class-level data into the fund level. In particular, a fund's TNA and flow are defined as the sum of each class's TNA and flow. Other fund characteristics, such as return, expense ratio, and portfolio turnover ratio, are defined as the average of class variables (weighted by the class TNA). Fund age is defined as the age of the oldest existing class in the fund.

All variables are winsorized at the 1st and 99th percentiles to reduce the effect of outliners. The average TNA of our sample funds is $1,906 million and the median is much more modest at $210 million. The average annual expense ratio is 1.17% and the average age is 16.8 years. Funds in our sample trade relatively frequently, with an average annual portfolio turnover rate of 0.81. The average annualized six-month return net of expense ratio is 5.24%. The risk-adjusted returns are much lower, with the average annualized four-factor alpha being -1.20%. The average monthly return volatility in the last six months is 3.94%. The mean six-month fund flow rate (TNA divided by total net flow) is 0.03. We compare the funds in our sample with the CRSP universe of actively managed equity funds in untabulated analyses and find that our sample funds have slightly larger TNAs. Both the performance metrics (return and alpha) and other fund characteristics, on the other hand, are quite comparable. We include fund size, annual expense ratio, portfolio turnover rate, and fund age as baseline control variables in regressions.

4. Textual Confidence and Mutual Fund Performance

4.1 Pooling Funds' Past Performance

As discussed in Section 1, the predictive power of textual confidence in a fund's shareholder letter for its future performance is expected to be related to its past performance. Our first step in testing this hypothesis is to examine whether textual confidence alone can predict future performance – we conjecture that it cannot because of the contrasting motives of skilled and lucky fund managers discussed before. We use the *net* confidence measure from the machine-based confidence lexicon as our key independent variable in this and next sections for brevity. More specifically, we conduct panel regressions of the following equation:

$$FF4Alpha_{i,t+1,t+6} = \alpha + \beta \times Net\ Confidence_{i,t} + \gamma \times X_{i,t} + \varepsilon_{i,t+1,t+6}. \tag{3}$$

The dependent variable is a fund's Fama-French-Carhart four-factor alpha for the period from t+1 to t+6, with *t* being the shareholder letter release month. More specifically, we use each fund's monthly returns in the 24 months before the letter release month to estimate loadings on the market, size, value, and momentum factors. Assuming these factor loadings remain unchanged in the next six months, we compute the fund's future alpha based on its monthly returns and monthly factor premiums in the next six months. Alphas are annualized and net of fund expense, while using gross returns and gross alphas leads to similar results as those reported in the paper.

The control variables (*X*) include the letter's length and readability score, and fund characteristics including total net assets (TNA), expense ratio, portfolio turnover rate, and fund age. Year and fund fixed effects are included in all regressions and omitted from reporting. Year fixed effects account for the time-series variations in textual confidence driven by the market conditions and fund fixed effects control for the time-invariant component of each fund manager's writing style, thereby allowing us to capture the effects of within-fund variations. Standard errors are clustered by fund.[27]

We report the results in columns 1 and 6 of Table 3. When all funds are pooled together, the net textual confidence itself has no predictive power for future performance: the coefficients on *Net Confidence* in both columns are small and statistically insignificant. The results remain unchanged when we use other lexicons to define textual confidence and when we include both textual confidence and unconfidence in the regressions. The lack of performance predictability by

[27] The results are robust to clustering standard errors by year and by both year and fund (Peterson, 2009).

the pooled textual confidence measures underscores the importance of taking funds' past returns into consideration, which, as we will show in the next section, can reveal important information about skill.

4.2 Performance Predictability Conditional on Funds' Past Returns

In this section, we test the hypothesis that textual confidence can reveal fund managers' skill vs. luck once funds' past performances are accounted for. We focus on the machine-based net confidence measure for now and will compare it with other confidence measures later. We partition funds into three groups: the "top performer" group includes funds whose returns in the six months before the letter release month are in the highest cross-sectional quartile, the "bottom performer" group includes funds whose past six-month returns are in the lowest cross-sectional quartile, and the "medium performer" group includes the remaining 50% of funds.[28] We use past returns instead of risk-adjusted returns for partition because returns can stem from both luck and skill and therefore allow us to capture fund managers' reactions (in writing styles) to both, whereas the risk-adjusted return is primarily a skill measure. We create dummy variables *Top*, *Mid*, and *Btm* to indicate top, medium, and bottom performers, respectively. These dummies are then interacted with the net confidence measure, which effectively partitions the confidence measure into three parts conditional on past returns. We estimate the following equation to examine the performance predictability of textual confidence for each fund group:

$$FF4Alpha_{i,t+1,t+6} = \alpha + \beta_1 \times Top_{i,t} \times Net\ Confidence_{i,t} + \beta_2 \times Mid_{i,t} \times Net\ Confidence_{i,t} + \beta_3 \times Btm_{i,t} \times Net\ Confidence_{i,t} + \delta_1 \times Top_{i,t} + \delta_2 \times Btm_{i,t} + \gamma \times X_{i,t} + \varepsilon_{i,t+1,t+6}. \tag{4}$$

[28] The results are robust to other partition thresholds, e.g., by terciles or by quantiles within each investment objective category.

Top and *Btm* are included in the regressions to control for the potential reversal relation between past return and future alpha (treating medium performers as the based group).

Columns 2 and 7 of Table 3 report the results. Among bottom performers, managers writing more confidently outperform others: the coefficient on $Btm \times Net\ Confidence$ is 0.6397 (t = 2.29) in column 2, suggesting that a two standard deviation change in net textual confidence is associated with a 1% change in the fund's future alpha. The result is even stronger under the tf.idf score approach (column 7). Moreover, when focusing on annual reports in columns 3 and 8, we find that the performance predictability of textual confidence becomes more pronounced than when annual and semi-annual reports are pooled together, consistent with the notion that fund managers tend to pay more attention to drafting annual reports, potentially because investors value annual reports more than semi-annual reports.

Coming to the top and medium performers, there is some evidence that some of these funds' managers write confidently to deceive investors. For example, the coefficient on $Mid \times Net\ Confidence$ is -0.3038 and significant at the 10% level in column 2, indicating that when medium performers write confidently, their funds' future performance tends to be worse than otherwise similar funds. However, this finding is not robust as it does not hold under the tf.idf score approach (column 7). Textual confidence cannot predict top performers' future performance either. To summarize, the results in this section show machine-based textual confidence has strong predictive power for the future performance of bottom performers, whereas its informational value is weaker for medium and top performers.

4.3 Alternative Confidence Lexicons

Thus far we have focused on the net confidence measure generated from the machine-base confidence lexicons in the regressions. Recall that we have two other confidence lexicons, one is human-based and the other is the LM strong/weak modal wordlists. Although the three share some common words, they also have some important differences. The machine-based lexicons mostly capture the sense (or lack) of certainty, which is an important dimension in communicating confidence; the human-based lexicons include many words indicating intensity (or lack of intensity) in communication; the strong/weak modal wordlists emphasize assertiveness (or lack of assertiveness) in communication and contain only a small number of words. In this section, we investigate whether human- and LM modal-based confidence lexicons have similar informational values as the machine-based lexicons or not.

We first examine the textual confidence measure based on human-based lexicons. As mentioned before, the human-based confidence lexicons tilt toward the intensity of expression. Although this is intuitively appealing in that elevated intensity might strengthen the writer's expression of certainty, ex-ante, it is unclear whether this aspect is related to the fund manager's skill. We re-estimate equation (4) using the human-based net textual confidence measure and report the results in columns 4 (proportional measure) and 9 (tf.idf score) of Table 3. Both columns show human-based measures do not have strong predictive power for funds' future performance. For example, the coefficient on none of the three fund groups' confidence measures is statistically significant in column 9. The results suggest that compared to certainty, the intensity aspect of confidence contains little information about fund managers' skill. Since the human-based confidence lexicons emphasize intensity, these results also suggest that investors are unlikely to be able to detect skill from shareholder letters without resorting to machine-based techniques.

Next, we examine whether the textual confidence measure constructed using the LM modal words can predict funds' future performance. The strong modal words convey the writer's assertiveness and the weak modal words convey reservation. Both wordlists are drastically shorter than their machine- and human-based counterparts. This could be an advantage if the informational value of the writing style is concentrated in a small set of words and a disadvantage if the briefness limits the lexicons' power to capture the writing style. The results from the LM modal based net textual confidence measures are reported in columns 5 (proportional measure) and 10 (tf.idf score) of Table 3. None of the coefficients on the confidence measures is significant, suggesting that the reduction of power related to the wordlists' briefness outweighs their simplicity. Another drawback of small lexicons is a few words can have disproportionally large influences. For example, Panel C of Table 1 show "best", "always", and "must" account for 83 % of all strong modal words in our sample letters, and "may", "could", and "appear" account for 72% of all weak modal words.

4.4 Confidence Vs. Unconfidence

The key variable of interest in our previous tests has been net confidence, the difference between textual confidence and unconfidence. In this section, we examine the confidence and unconfidence variables separately. This is of interest to us because the two can carry distinct informational values that cannot be revealed by studying net confidence alone. For example, a high net confidence score can result from a high count of confidence words or a low count of unconfidence words, yet the performance predictability of net confidence can differ for the two cases. We decompose the net confidence measures in equation (4) into confidence and unconfidence measures and report the estimation results in Table 4. Panel A (B) focuses on the confidence (unconfidence) measures and untabulated analyses show similar results when including both in the same regression.

We find that among bottom performers, using machine-based confidence words heavily is positively related to future performance. The coefficient on *Btm* × *Confidence* is 0.9215 (t = 2.59) in column 1 of Panel A, translating into an 1% increase in future alpha when textual confidence increases by two standard deviations, and the results become even stronger when tf.idf scores are used (column 5) or when the sample is limited to annual reports (columns 2 and 6). Consistent with the previous findings that human- and LM modal-based net confidence measures contain little information about skill, we find that the coefficients on *Btm* × *Confidence* are all insignificant for confidence measures related to these lexicons (columns 3, 4, 7, and 8).

Panel B of Table 4 examines the performance predictability of various textual unconfidence measures and shows the predictive power of net confidence in Table 3 primarily stems from the prevalence of confidence words rather than the absence of unconfidence words. Regardless of whether it is constructed using the machine-based unconfidence, human-based unconfidence, LM weak modal, or LM uncertainty lexicons, the coefficients on unconfidence measures are all statistically insignificant.

4.5 Overstatement and Confidence

Recall that in addition to words related to certainty and intensity, the K-Means clustering method also identifies words related to overstatement and understatement.[29] These wordlists are interesting because they can be used to express strong or reserved emotions, which some investors might relate to confidence. For example, expressions of emotions might be perceived as a

[29] The top 30 over- and under-statement words used in our sample letters can be found in Internet Appendix A.

weakness indicating a lack of confidence. In this section, we examine whether textual overstatement contains information about fund managers' skill.

We define a shareholder letter's net overstatement level as the difference between its textual overstatement and understatement levels and re-estimate equation (4) using this net overstatement measure. The results are reported in Table 5. Overstating letters appear to send a negative signal about funds' future performance: columns 1 and 5 show the coefficient on the net overstatement measure is -1.7754 (t = -2.42) when the proportional measure is used and -0.0360 (t = -1.88) when the tf.idf score is used. In terms of the economic significance, a two standard deviation increase in net overstatement is associated with a 0.43% drop in alpha.

Interestingly, top performers are the driving force for the negative relation between net overstatement and future performance. In column 2 of Table 5, we decompose the net proportional overstatement measure into those for bottom, medium, and top performers, and find that the coefficient on $Top \times\ Net\ Overstatement$ is -7.5130 (t = -3.57). This result remains qualitatively unchanged when the tf.idf scores are used (column 6). To the extent that an overstating writing style might help influence investors' reactions, this finding suggests that some lucky yet unskilled fund managers, expecting the superior past performance to be transient, are motivated to capitalize on the fleeting luck. Alternatively, the overstating writing style might result from the manager's emotional reactions (e.g., ecstasy and excitement) to the unexpected superior past performance that does not continue into the future because it is driven by luck.

We examine textual overstatement and understatement separately in columns 3 and 7 of Table 5 and note two interesting findings. First, consistent with the results from net overstatement, we find that the coefficients on $Top \times Overstatement$ are all negative and significant at the 1% level, suggesting that overstating top performers tend to have declining future performance.

Second, the coefficients on $Top \times Understatement$ are all positive and significant at the 5% level, suggesting that understating top performers' superior performance will continue into the future. Both results contribute to the performance predictability of net overstatement.

When including the machine-based confidence measure and over- and under-statement variables together (columns 4 and 8 of Table 5), we find that the results for all of them remain qualitatively unchanged, indicating that confidence and overstatement represent two distinct information sources about mutual fund managers' skill. Note that we only report the results for the textual confidence measures but not the unconfidence or net confidence measures because the former contains little information about future fund performance (Table 4) and the latter leads to qualitatively similar results (untabulated).

Bochkay, Chava, and Hales (2018) investigate the information content of linguistic extremity (or hyperbole).[30] They show trading volume and stock prices react more strongly when corporate executives use hyperbolic languages in earnings conference calls, and such language contains information about the firm's future operating performance. Their hyperbole measure focuses on the degree of tone instead of the emotional expressiveness emphasized in our overstatement measures. Furthermore, they examine informal oral communications whereas we analyze the styles of formal writing.

4.6 Confidence and Tone

Tone is one of the most studied textual characteristics in accounting, economics, and finance. Two recent papers examine the tone of mutual funds' shareholder letters. Hillert, Niessen-

[30] Bochkay et al. (2018) rely on human judgment to form their lexicon. Specifically, they compile a list of words that either account for at least 1% of the conference call or are included in LM's tone wordlists. They ask MTurk users to assign an "extremity rating" ranging from -5 to 5 to each word and those having ratings greater than 3 or less than -3 are identified as extreme words (not publicly available).

Ruenzi, and Ruenzi (2016) do not find any significant relation between letter tone and future fund performance, whereas Chu and Kim (2018), focusing on a sample of closed-end funds, find a negative relation. In this section, we compare the informational value of our machine-based confidence measure and that of textual tone.

The tone lexicon has limited overlap with the machine-based confidence lexicon: the top 30 machine-based confidence words in our sample do not overlap with the top 30 positive-tone words at all.[31] Panel D of Table 2 shows the correlation between net confidence and net tone is 0.17, indicating a relatively low likelihood for the two types of words to appear in the same shareholder letter. We report further analyses in Table 6. In Column 1, we find that the coefficient on proportional net tone is -0.1826 (t = -3.24), suggesting that more pessimistic fund managers have better future performance. This result, although consistent with the findings of Chu and Kim (2018), is puzzling because most prior studies document a positive relation (Tetlock, 2007; Tetlock, Saar-Tsechansk and Macskassy, 2008; Loughran and McDonald, 2011). To gain deeper insights, in Column 2 of Table 6, we decompose net tone into those for bottom, medium, and top performers to examine their relations to future fund performance. It turns out that the negative relation in column 1 is driven by bottom performers. In other words, when bottom performers candidly discuss their performance with a negative tone, this acknowledgment and acceptance of reality is an indicator for future performance improvements. The results in Columns 3 further show this effect works in both ways: candidly discussing poor past performance (instead of trying to spin them) forecasts better future performance – the coefficients on $Btm \times Negative\_Tone$ are all positive and significant, while sugarcoating poor past performance with a positive tone forecasts

[31] The top 30 positive and negative tone words used in our sample letters can be found in Internet Appendix B.

deteriorating performance – the coefficients on $Btm \times Positive\_Tone$ are all negative and significant.

We add the machine-based confidence measure to the regressions in columns 4 and 8 of Table 6. We report the results for the textual confidence measures but not the unconfidence or net confidence measures because the former contains little information about future fund performance (Table 4) and the latter leads to qualitatively similar results (untabulated). The results in columns 4 and 8 show the performance predictability of textual confidence is not affected by the tone variables: the coefficients on $Btm \times Confidence$ remain positive and highly significant in both columns. These findings suggest that textual confidence and tone are complements rather than substitutes in predicting future fund performance and they convey distinct information about fund managers' skill.

5. Do Investors Respond to Textual Confidence?

We have found evidence for the performance predictability of (machine-based) textual confidence, while an important question related to it is the motives behind mutual fund managers' writing styles. Is the managers' use of confidence words and the ability to communicate confidence in shareholder letters an intrinsic trait correlated with skill, i.e., something hard for unskilled managers to do? Or, are managers write confidently to influence investors? Note that the success of the latter depends on whether investors are capable of identifying the information embedded in writing styles. In this section, we examine whether investors respond to textual confidence in funds' shareholder letters.

We measure investors' reaction to shareholder letters by fund flows, defined as the ratio of total net flows in the six months after the shareholder report release month to fund TNA reported at the end of the report release month. As in the previous two sections, we focus on the confidence rather than the net confidence measures and estimate the following equation:

$$Net\ flow_{i,t+1,t+6} = \alpha + \beta_1 \times Top_{i,t} \times Confidence_{i,t} + \beta_2 \times Mid_{i,t} \times Confidence_{i,t} + \beta_3 \times Btm_{i,t} \times Confidence_{i,t} + \delta_1 \times Top_{i,t} + \delta_2 \times Btm_{i,t} + \gamma \times X_{i,t} + \varepsilon_{i,t+1,t+6}. \quad (5)$$

In addition to the control variables in the previous analyses, we also include funds' flow rates in the last six months, funds' past six-month returns, and the squared past six-month returns as additional controls. The last two variables account for the well-documented convex relationship between flow and past return.

Table 7 reports the results: columns 1-6 report the results for proportional ratios and columns 7-12 report the results for tf.idf scores. We first examine whether investors respond to machine-based textual confidence (columns 1 and 7) and do not detect any significant effects. Columns 2 and 8 examine the human-based confidence measure and show no significant effects either. Given that this confidence measure is human-perception based, these results are somewhat surprising. Next, we partition funds into bottom, medium, and top performers (as before) to examine whether investors respond to textual confidence differently for different fund groups. The results for machine-based measures are in columns 4 and 10 and those for human-based measures are in columns 5 and 11. None of the coefficients on any confidence variable is statistically significant. Overall, investors appear to be insensitive to textual confidence expressed in shareholder letters. This finding could partially explain why unskilled fund managers do not mimic the skilled ones in writing confidently.

We also explore whether investors respond to overstatements in shareholder letters. The average effects are insignificant (columns 3 and 9 of Table 7), whereas there is evidence of negative reactions for overstating bottom performers in columns 6 and 12, as shown by the negative and significant coefficients on $Btm \times Net\ Overstatement$. In other words, when a bottom performer emotionally overstates, it triggers negative reactions from investors, perhaps because emotions are interpreted as a weakness, although as shown before (Table 5), the overstatement itself does not predict changes in future performance for bottom performers. Recall that the results in Table 5 show overstating top performers tend to perform worse than other top performers in the future; investors, however, are not able to capture this information, as shown by the insignificant coefficients on $Top \times Net\ Overstatement$ in columns 6 and 12 of Table 7. Finally, untabulated analyses show investors do not respond to the tone of shareholder letters. We conclude that mutual fund investors are largely unaware of the information embedded in the writing styles of shareholder letters. Since our performance analyses show only machine-based textual confidence can predict future performance and the human-based measures cannot, our finding that investors do not respond to textual confidence is consistent with the notion that investors are unaware of the informational advantage of machine learning vs. human judgments either.

6. Conclusion

We study the information content of the confidence level implied in the text. We use unsupervised machine learning techniques to generate a new lexicon for a writer's level of certainty in expressing opinions. Using it to capture mutual fund managers' textual confidence in letters to shareholders, we show confidence contains important information about the manager's skill: underperforming fund managers who write confidently significantly outperform other underperforming managers in the next six months. Further analyses reveal that our confidence measure is informationally distinct from other textual characteristics such as tone and overstatement, and outperforms human-based confidence measures in predicting performance. Underscoring the informational value of machine learning vs. human judgments, we also show capital flows do not respond to the manager's confidence level.

This paper is the first to examine the relevance and informational role of textual confidence (certainty) in financial disclosures. Compared to other textual characteristics such as readability and tone, textual confidence reflects the writer's self-assessment of how accurately she can evaluate current events and forecast future. By highlighting this new dimension of communication, we contribute to the growing literature on textual analysis and non-descriptive soft information. We are also the first to explore the differences in various machine- and human-based confidence lexicons. Our evidence shows not all forms of the perceived confidence are equally informative. Finally, by examining mutual funds' shareholder letters, we also offer a novel angle to separate "skill" and "luck". By presenting evidence on the relations between skill and a comprehensive set of textual characteristics, we identify writing style as a useful new tool for differentiating skill from luck.

Appendix A. Variable Definitions

Variable	Description
Textual Variables	
Confidence (Proportional)	The ratio of the number of confidence related words to the total number of non-stop words in a shareholder letter. Confidence-M refers to machined based confidence. Confidence-H refers to human-based confidence.
Unconfidence (Proportional)	The ratio of the number of unconfidence words to the total number of non-stop words in a shareholder letter. Unconfidence-M refers to machined based unconfidence. Unconfidence-H refers to human-based unconfidence.
Net Confidence (Proportional)	Confidence score (proportional) minus Unconfidence score (proportional).
Strong Modal (Proportional)	The ratio of the number of strong modal words to the total number of non-stop words in a shareholder letter. Strong modal words are from Laughran and Mcdonald (2011)'s wordlist.
Weak Modal (Proportional)	The ratio of the number of weak modal words to the total number of non-stop words in a shareholder letter. Weak modal words are from Laughran and Mcdonald (2011)'s wordlist.
Net Modal(Proportional)	Strong Modal (proportional) minus Weak Modal (proportional).
Overstatement (Proportional)	The ratio of the number of words related to the emotional overstating to the total number of non-stop words in a shareholder letter. Please refer to Appendix B for the complete wordlist of emotional overstating.
Understatement (Proportional)	The ratio of the number of words related to the emotional understating to the total number of non-stop words in a shareholder letter. Please refer to Appendix C for the complete wordlist of emotional understating.
Net Overstatement (Proportional)	Overstatement (proportional) minus Understatement (proportional).
LM Uncertainty (Proportional)	The ratio of the number of words related to the uncertainty to the total number of non-stop words in a shareholder letter. We use LM's uncertainty wordlist to identify uncertainty words. This variable is used as an alternative measure of unconfidence.
Positive Tone (Proportional)	The ratio of the number of positive tone words to the total number of non-stop words in a shareholder letter. Positive tone related words are based on Laughran and Mcdonald (2011)'s wordlist.
Negative Tone (Proportional)	The ratio of the number of negative tone words to the total number of non-stop words in a shareholder letter. Negative tone related words are based on Laughran and Mcdonald (2011)'s wordlist.
Net Tone (Proportional)	Positive tone score (proportional) minus negative tone score (proportional).

tf.idf Confidence	Confidence tf.idf score based on the machine-based confidence lexicon (tfidf Confidence-M) or human-based lexicon (tf.idf Confidence-H).
tf.idf Unconference	Unconfidence tf.idf score based on the machine-based unconfidence lexicon (tfidf Unconfidence-M) or human-based lexicon (tf.idf Unconfidence-H).
tf.idf Net Confidence	Confidence tf.idf score minus Unconfidence tf.idf score
tf.idf Strong Modal	The strong modal tf.idf score. The strong modal wordlist is from Laughran and Mcdonald (2011).
tf.idf Weak Modal	The weak modal tf.idf score. The weak modal wordlist is from Laughran and Mcdonald (2011).
tf.idf Net Modal	Strong Modal tf.idf score minus Weak Modal tf.idf score
tf.idf Overstatement	Emotional overstating tf.idf score. Please refer to Appendix B for the complete wordlist of emotional overstating.
tf.idf Understatement	Emotional understating tf.idf score. Please refer to Appendix C for the complete wordlist of emotional understating.
tf.idf Net Overstatement	Overstating tf.idf score minus the understating tf.idf score.
tf.idf Uncertainty	Uncertainty tf.idf score. We use LM's uncertainty wordlist to identify uncertainty words. This variable is used as an alternative measure of unconfidence.
tf.idf Positive Tone	The tf.idf positive tone score. The positive tone is based on Laughran and Mcdonald (2011)'s word list.
tf.idf Negative Tone	The tf.idf negative tone score. The negative tone is based on Laughran and Mcdonald (2011)'s word lists.
tf.idf Net Tone	The difference between the tf.idf positive tone score and the tf.idf negative tone score.
Article Length	The number of non-stop words in the shareholder letter.
Readability	Readability is the Flesch–Kincaid Grade Level score defined as follows: $Readability\ Score = 0.39 * \frac{Total\ Words}{Total\ Sentences} + 11.8 * \frac{Total\ Syllables}{Total\ Words} - 15.97.$
Key Dependent Variables	
FF4Alpha (t+1,t+6)	The annualized net alpha in the next six months after the letter release month t, based on the Fama-French three-factor model plus the momentum factor. Betas for the four factors are estimated using returns in the twenty-four months before month t, from CRSP.
Net Flow (t+1,t+6)	The ratio of total net flows over the next six months after the letter release month t to the fund TNA of the release month.
Control Variables	
Top	A dummy equal to one if the fund's *Past Return (t-6,t-1)* is in the top quartile among all sample funds in the specific reporting period, and zero otherwise. t refers to the shareholder letter release month.

Mid	A dummy equal to one if the fund's *Past Return (t-6,t-1)* is in the second or third quartiles among all sample funds in the specific reporting period, and zero otherwise. t refers to the shareholder letter release month.
Btm	A dummy equal to one if the fund's *Past Return (t-6,t-1)* is in the bottom quartile among all sample funds in the specific reporting period, and zero otherwise. t refers to the shareholder letter release month.
Log TNA	The logarithmic value of fund size TNA before the letter release month t, from CRSP.
Expense Ratio	A mutual fund's average expense ratio across all classes in the year of the letter release.
Portfolio Turnover Ratio	A mutual fund's average portfolio turnover ratio across all classes in the year of the letter release.
Fund Age (Years)	The age of the oldest fund class of a mutual fund in the year of the letter release.

Appendix B. Cleaning and Partitioning the IV-4 Overstatement Lexicon Using the K-Means Clustering Method

B1. Words Excluded from this Study (54 Words)
above, account, acknowledgment, after, all, away, basis, billion, board, capital, circle, cohort, commission, course, daily, do, doubt, due, earth, era, etc, few, fundamental, gross, hundred, large, least, mansion, matter, million, millionth, might, once, orphan, over, part, patent, point, principal, privacy, rare, rather, run, say, security, short, simple, thousand, trillion, total, usual, wealth, what, worth

B2. Sub-wordlist 1 (Confidence-related, 119 Words)
absolute, accuracy, accurate, accurateness, admit, appropriate, ascertain, assurance, assure, authentic, authenticity, authoritative, basic, certainty, certification, certify, clarity, coherent, cohesion, complexity, comprehensive, conclusive, confidence, confident, confirm, confirmation, consequence, correct, crucial, decide, decisive, definite, determination, determine, emphasis, emphasize, ensure, essence, essential, establish, eventual, exact, exceed, exceptional, exclusive, exempt, feasible, foremost, furthermore, genuine, ideal, immediate, importance, indispensable, intensive, justification, justify, likelihood, moreover, necessary, necessity, outright, overly, precise, precision, probability, probable, prompt, promptly, proof, prove, purely, purpose, readily, reliability, reliable, rely, robust, scrutiny, secure, significance, sole, speedy, spite, stability, stabilize, standardize, strict, supreme, swift, thereby, therefore, thorough, truth, truthful, ultimate, unanimous, unconditional, unequivocal, universal, unlimited, unrestricted, urgent, utmost, valid, validity, verification, verify, vital, whatsoever, wherever, whoever, *contradictory, custom, enhancement, establishment, practical, privileged, proprietary*,

B3. Sub-wordlist 2 (Intensity-related, 183 Words)
abundant, active, actual, acute, again, almost, alone, altogether, always, amazing, any, anybody, anyone, anything, anywhere, bad, badly, brilliant, bulk, certain, chief, chronic, clear, common, complete, complex, considerable, constant, continuous, countless, critical, crude, dangerous, deep, direct, distinct, dominant, each, either, else, enormous, enough, entire, especially, even, ever, every, everybody, everyone, everything, everywhere, evident, excess, excessive, extensive, extra, extraordinary, extreme, far, fast, feature, final, forever, frequent, full, general, giant, good, grand, grave, great, high, highlight, huge, important, impressive, indeed, inevitable, instant, instantly, intense, intensity, just, last, lead, length, lengthy, likely, literally, mad, magnitude, main, major, majority, many, mark, mass, mean, most, natural, normal, notable, notorious, numerous, obvious, often, ordinary, overall, paramount, particular, perfect, plain, plenty, plus, positive, possibility, possible, present, presumably, primarily, primary, prospect, pure, quick, quite, rapid, real, regular, remarkable, remarkably, right, seldom, severe, sharp, shock, significant, since, sizable, so, soon, sound, spectacular, speed, still, straight, strike, strong, substantial, such, sudden, sure, surprise, terrible, thus, too, tremendous, true , unique, unprecedented, unusual, vast, very, virtual, whatever, whole, widespread, wild, wonder, wonderful, worst, yet, *arrest, celebrity, chaos, danger, effect, fact, fame, heritage, hero, load, star,*

B4. Sub-wordlist 3 (Emotional Overstatement-related, 221 Words)

abominable, absurd, abundance, accentuate, addiction, aghast, alarming, alas, alight, amaze, anarchy, anti-social, appreciable, apt, ascertainment, assuredly, assuredness, astonish, astound, astronomical, atrocious, audacious, audacity, avid, awful, blatant, blurt, brutality, captivation, catastrophe, censor, censorship, chaotic, clique, cogent, cognizance, cognizant, colossal, commodious, commonplace, commotion, completeness, confide, consequent, conspicuous, continual, cynicism, dash, debatable, decadence, decadent, degenerate, delicacy, delirium, deluge, dependability, dependable, desertion, detachment, distracting, distraction, diversion, doubtless, dreadful, eccentric, eccentricity, eminence, emphatic, endurance, endure, engulf, eternal, exaggerate, exaggeration, exclamation, extravagance, extravagant, exuberance, exuberant, fancy, fantastic, fascinate, fiasco, flagrant, flashy, flaunt, focal, frantically, fraught, furious, fury, fussy, gaudy, ghastly, gigantic, glorious, grandeur, gratuitous, grotesque, havoc, hectic, heroic, heroine, hideous, hopeless, hustle, illuminate, illustrious, immaculate, immediacy, immense, immortal, imprecision, incessant, inconceivable, inconsistency, incontestable, incontestability, incredible, incredibility, indeterminable, indeterminate, indisputable, individuality, indulgence, inevitability, infallible, infinite, innumerable, intricate, invariable, invariably, irrefutable, justifiably, likeliness, limitless, lucid, luminous, magnificence, magnificent, magnify, majestic, manageable, marvel, matchless, maximization, mighty, momentous, monstrous, monumental, navigable, necessitate, notoriety, novelty, originality, ostracize, outcast, overwhelm, peerless, perfection, perfectionism, perfectionist, perpetual, perpetuate, plentiful, poignant, predictable, predominant, predominate, priceless, privy, prodigious, prohibitive, prominence, punctual, queer, realistically, rebellious, relevancy, ridiculous, saga, sensational, sheer, shun, simplicity, solace, speechless, speedily, splendid, stark, steadfast, steadfastness, superfluous, superlative, swiftness, torrent, towering, traitor, unavoidable, unbelievable, uncommon, uncontested, undeniable, undisputed, undoubted, undoubtedly, unfailing, unhurried, unmistakable, unmitigated, unquestionable, unquestioned, untold, unwavering, uppermost, utter, utterly, uttermost, vastness, vile, wee

Appendix C. Cleaning and Partitioning the IV-4 Understatement Lexicon Using the K-Means Clustering Method

C1. Words removed from the original IV-4 Understatement Wordlist (31 Words)
all, doubt, event, everyday, few, hand, large, laugh, least, matter, moment, no, number, part, permanent, rare, rather, reserve, roundabout, rule, save, short, simple, snack, speak, time, touch, usual, way, word, wondrous

C2. Sub-wordlist 1 (Unconfidence-related, 57 Words)
about, accident, anyway, apart, appear, approximate, approximately, barely, bit, but, care, chance, despite, fair, handful, hard, impossible, kind, known, less, light, little, luck, maybe, minor, near, never, nobody, nothing, nowhere, only, perhaps, pretty, question, rough, seem, several, single, slow, small, some, somebody, somehow, someone, something, sometime, somewhat, somewhere, sort, tend, tiny, uncertain, uncertainty, unlikely, unsure, unsureness, well

C3. Sub-wordlist 2 (Lack of Intensity-related, 71 Words)
accidental, adequate, ambiguity, ambiguous, apparent, arbitrary, aside, autonomous, brief, careful, casual, caution, cautious, comparative, comparatively, confusion, contingent, controversial, customary, dubious, gamble, gradual, incorrect, indefinite, indirect, interim, limit, meager, mention, mere, minimal, minimize, minimum, moderate, moderation, negligible, nominal, obscure, occasion, occasional, opinion, outline, partial, particle, qualification, qualify, questionable, regardless, relative, reservation, scant, scarce, secondary, selective, slight, speculate, speculation, speculative, suggest, suggestion, suspicion, suspicious, tendency, unclear, undetermined, unfavorable, unnecessary, unspecified, vague, weak, weakness

C4. Sub-wordlist 3 (Emotional Understatement-related, 116 Words)
accustom, allege, aloof, anomaly, antipathy, anyhow, apathetic, apathy, awhile, baffle, bafflement, beware, bewilder, bewilderment, blur, brusque, bungle, calamity, calmness, changeable, coincidence, conceivable, confound, confuse, cursory, daze, dilemma, dim, disbelief, disconcerted, dismay, disputable, doubtful, equivocal, evidently, faint, falter, feeble, fortunate, gingerly, haziness, heed, hesitant, hesitate, hesitation, impossibility, improbability, improbable, incalculable, indecision, indecisive, indecisiveness, indistinct, indistinguishable, inexact, infrequent, insecure, insecurity, insignificant, ironic, irony, luckily, misinform, misinformed, misrepresent, mistaken, mistrust, misunderstanding, modesty, momentary, muddy, mumble, murky, nebulous, oblique, paranoid, perplex, pointless, precarious, puny, puzzle, puzzlement, quandary, scarcely, shady, sketchy, superficial, suppose, temperance, temperate, trifle, trivial, undecided, undefined, undependable, undependability, unfortunate, unimportant, unlikelihood, unlucky, unmoved, unpleasant, unpredictable, unreliability, unreliable, unsatisfactory, unsound, unsoundness, untrue, vacillate, vagueness, vexation, vexing, weakly, whisper, worthless

Appendix D. Confidence Lexicons

D1. Machine-based (119 Words)

absolute, accuracy, accurate, accurateness, admit, appropriate, ascertain, assurance, assure, authentic, authenticity, authoritative, basic, certainty, certification, certify, clarity, coherent, cohesion, complexity, comprehensive, conclusive, confidence, confident, confirm, confirmation, consequence, correct, crucial, decide, decisive, definite, determination, determine, emphasis, emphasize, ensure, essence, essential, establish, eventual, exact, exceed, exceptional, exclusive, exempt, feasible, foremost, furthermore, genuine, ideal, immediate, importance, indispensable, intensive, justification, justify, likelihood, moreover, necessary, necessity, outright, overly, precise, precision, probability, probable, prompt, promptly, proof, prove, purely, purpose, readily, reliability, reliable, rely, robust, scrutiny, secure, significance, sole, speedy, spite, stability, stabilize, standardize, strict, supreme, swift, thereby, therefore, thorough, truth, truthful, ultimate, unanimous, unconditional, unequivocal, universal, unlimited, unrestricted, urgent, utmost, valid, validity, verification, verify, vital, whatsoever, wherever, whoever, *contradictory, custom, enhancement, establishment, practical, privileged, proprietary*,

D2. Human-based (110 Words)

absolute, abundance, abundant, accuracy, accurate, accurateness, appropriate, apt, assurance, assuredness, authentic, authenticity, basic, bulk, certain, certainty, certification, clarity, clear, cognizant, complete, comprehensive, conclusive, confidence, confident, confirm, confirmation, constant, correct, decide, decisive, definite, determination, determine, distinct, emphasis, emphasize, enhancement, ensure, entire, essential, every, evident, exact, exceed, exclusive, extensive, fact, focal, foremost, full, genuine, huge, immediate, importance, important, incontestability, indeed, indeterminable, indisputable, instantly, just, justification, lead, major, majority, maximization, natural, necessary, necessitate, necessity, normal, outright, perfection, perpetuate, plentiful, plus, precise, precision, predictable, primary, prominence, prompt, promptly, prove, quite, rapid, readily, relevancy, reliable, robust, significant, sole, sound, stability, straight, sure, thereby, thorough, uncontested, undeniable, unequivocal, universal, unlimited, unmistakable, unquestionable, unquestioned, valid, validity, verify

Overlapping between Machine-based and Human-based Confidence Wordlists (58 Words)

absolute, accuracy, accurate, accurateness, appropriate, assurance, authentic, authenticity, basic, certainty, certification, clarity, comprehensive, conclusive, confirm, confirmation, confidence, confident, correct, decide, decisive, definite, determination, determine, emphasis, emphasize, enhancement, ensure, essential, exact, exceed, exclusive, foremost, genuine, immediate, importance, justification, necessary, necessity, outright, precise, precision, prompt, promptly, prove, readily, reliable, sole, stability, thereby, thorough, unlimited, unequivocal, universal, valid, validity, verify, robust,

Appendix E. Unconfidence Lexicons

E1. Machine-based (57 Words)

about, accident, anyway, apart, appear, approximate, approximately, barely, bit, but, care, chance, despite, fair, handful, hard, impossible, kind, known, less, light, little, luck, maybe, minor, near, never, nobody, nothing, nowhere, only, perhaps, pretty, question, rough, seem, several, single, slow, small, some, somebody, somehow, someone, something, sometime, somewhat, somewhere, sort, tend, tiny, uncertain, uncertainty, unlikely, unsure, unsureness, well

E2. Human-based (49 Words)

ambiguity, ambiguous, appear, approximately, caution, cautious, comparative, confuse, contingent, dim, evidently, improbable, incalculable, indefinite, inexact, infrequent, insecurity, less, limit, maybe, misinform, moderate, murky, negligible, never, oblique, occasion, perhaps, question, relative, reservation, seem, sometime, somewhere, speculation, speculative, suggestion, suppose, uncertain, uncertainty, undependability, unlikelihood, unlikely, unreliability, unsoundness, unspecified, unsure, unsureness, vacillate

Overlapping between Machine-based and Human-based Unconfidence Wordlists (15 Words)

appear, approximately, less, maybe, perhaps, question, sometime, somewhere, never, seem, uncertain, uncertainty, unlikely, unsure, unsureness

References

Barbalet, J., 1993. Confidence: Time and emotion in the sociology of Action. Journal for the Theory of Social Behaviour 23, 229-47.

Blei, D., Ng, A., and Jordan, M., 2003. Latent dirichlet allocation. Journal of Machine Learning Research 3, 993-1022.

Bochkay, K., Chava, S., Hales, J., 2018. Hyperbole or reality? Investor response to extreme language in earnings conference calls. Working Paper, University of Miami.

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

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

Chu, Y., Kim, H., 2018. Pessimistic fund managers. Working paper, University of South Carolina.

Cremers, M., Petajisto, A., 2009. How active is your fund manager? A new measure that predicts performance. Review of Financial Studies 22, 3329-3365.

Das, S., 2014. Text and context: Language analytics in finance. Foundations and Trends in Finance 8, 145-261.

Dyer, T., Lang, M., and Stice-Lawrence, L., The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. Journal of Accounting Economics 64, 221-245.

Engelberg, J., 2008. Costly information processing: evidence from earnings announcements. Working paper, University of California, San Diego.

Evans, R., 2010. Mutual fund incubation. Journal of Finance 65, 1581-1611.

Fama, E., French, K., 2010. Luck versus skill in the cross-section of mutual fund returns. Journal of Finance 65, 1915-1947.

Hillert, A., Niessen-Ruenzi, A., Ruenzi, S., 2016. Mutual fund shareholder letters: Flows, performance, and managerial behavior. Working paper, University of Mannheim.

Hwang, B., Kim, H., 2017. It pays to write well. Journal of Financial Economics 124, 373-394.

Jurafsky, D., Martin, J. H., 2009. Speech and language processing. Prentice-Hall, Upper Saddle River, NJ.

Kacperczyk, M., Sialm, C., Zheng, L., 2005. On the industry concentration of actively managed equity mutual funds. Journal of Finance 60, 1983-2011.

Kacperczyk, M., Sialm, C., Zheng, L., 2008. Unobserved actions of mutual funds. Review of Financial Studies 21, 2379-2416.

Kearney, C., Liu, S., 2014. Textual sentiment in finance: A survey of methods and models. International Review of Financial Analysis 33, 171-85.

Lehavy, R., Li, F., Merkley, K., 2011. The effect of annual report readability on annual following and the properties of their earnings forecasts. The Accounting Review 86, 1087-1115.

Li, F., 2008. Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics 45, 221-247.

Li, F., 2010a. Textual analysis of corporate disclosures: A survey of the literature. Journal of Accounting Literature 29, 143-65.

Li, F., 2010b. The information content of forward-looking statements in corporate filings - A naive Bayesian machine learning algorithm approach. Journal of Accounting Research 48, 1049-1102.

Liberti, J., Petersen, M., 2019. Information: hard and soft. Review of Corporate Finance Studies 8, 1-41.

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

Loughran, T., McDonald, B., 2014a. Measuring readability in financial disclosures. Journal of Finance 69, 1643-1671.

Loughran, T., McDonald, B., 2014b. Regulation and financial disclosure: The impact of plain English. Journal of Regulatory Economics 45, 94-113

Loughran, T., McDonald, B., 2016. Textual analysis in accounting and finance: A survey. Journal of Accounting Research 54, 1187-1230.

Lundholm, R., Rogo, R., Zhang, J., 2014. Restoring the tower of Babel: How foreign firms communicate with US investors. The Accounting Review 89, 1453-1485.

Pennington, J., Socher, R., and Manning, C.D., 2014. GloVe: Global Vectors for Word Representation. Working Paper, Stanford University

Rogers, J., Van Buskirk, A., Zechman, S., 2011. Disclosure and shareholder litigation. The Accounting Review 86, 2155-2183.

Stajkovic, A., 2006. Development of a core confidence – Higher order construct. Journal of Applied Psychology 91, 1208-1224.

Tetlock, P., 2007. Giving content to investor sentiment: The role of media in the stock market. Journal of Finance 62, 1139-1168.

Tetlock, P., Saar-Tsechansky, M., Macskassy, S., 2008. More than words: Quantifying language to measure firms' fundamentals. Journal of Finance 63, 1437-1467.

Figure 1
Sample Questionnaire at MTurk

Do you agree: if the word below is used in the letter by a fund manager, it indicates that the manager is confident, but not emotional.

	Strongly Agree	Somewhat Agree	Disagree	Not Sure
significance	○	○	○	○
each	○	○	○	○
manageable	○	○	○	○
notable	○	○	○	○
degenerate	○	○	○	○

	Strongly Agree	Somewhat Agree	Disagree	Not Sure
gigantic	○	○	○	○
perfectionist	○	○	○	○
excessive	○	○	○	○
colossal	○	○	○	○
speed	○	○	○	○

Figure 1 - Continued

Do you agree: if the word below is used in the letter by a fund manager, it indicates that the manager is ***lack of confidence*, but not sounding negatively emotional**.

	Strongly Agree	Somewhat Agree	Disagree	Not Sure
unlikelihood	○	○	○	○
apathy	○	○	○	○
indecisiveness	○	○	○	○
rough	○	○	○	○
obscure	○	○	○	○

	Strongly Agree	Somewhat Agree	Disagree	Not Sure
suggestion	○	○	○	○
paranoid	○	○	○	○
perhaps	○	○	○	○
approximately	○	○	○	○
confusion	○	○	○	○

Table 1
Confidence/Unconfidence Lexicons:
Top 30 Most Used Confidence/Unconfidence Words in Mutual Funds' Shareholder Letters

Panel A highlights the misclassification problem in the Harvard IV-4 overstatement/understatement wordlists for applications in business and finance. It reports the top 30 most used IV-4 overstatement/understatement words in our sample mutual funds' shareholder letters. Highlights in grey indicate that the word is removed. Panel B (C) reports the top 30 most used confidence (unconfidence) words in sample letters. The machined-based lexicons are generated through the K-Means Clustering method; the human-based lexicons are generated by human judgments through MTurk; LM strong (weak) modal and uncertainty lexicons are compiled by Loughran and McDonald (2011). The inflectional forms of nouns and verbs are reduced to their base forms for LM lexicons. Stop words are excluded.

Panel A. Top 30 IV-4 Over/Understatement Words in Shareholder Letters

Top 30 IV-4 Overstatement Words in Sample Letters	% of Total IV-4 Overstatement Word Count	Cumulative %	Top 30 IV-4 Understatement Words in Sample Letters	% of Total IV-4 Understatement Word Count	Cumulative %
high	5.85%	5.85%	time	9.38%	9.38%
large	4.27%	10.13%	large	9.10%	18.48%
total	3.98%	14.10%	well	7.52%	26.00%
strong	3.52%	17.62%	small	7.16%	33.15%
capital	3.35%	20.97%	short	5.48%	38.63%
security	3.35%	24.32%	relative	4.42%	43.04%
short	2.57%	26.89%	less	3.66%	46.70%
many	2.39%	29.28%	care	2.98%	49.69%
since	2.33%	31.61%	number	2.73%	52.42%
good	2.13%	33.74%	reserve	2.68%	55.10%
positive	1.77%	35.52%	several	2.56%	57.66%
last	1.77%	37.29%	slow	2.28%	59.94%
account	1.68%	38.97%	part	2.16%	62.11%
lead	1.64%	40.60%	despite	2.10%	64.21%
due	1.48%	42.08%	weak	2.03%	66.24%
significant	1.41%	43.49%	appear	2.02%	68.26%
great	1.28%	44.77%	near	1.73%	69.99%
even	1.18%	45.95%	way	1.53%	71.52%
important	1.17%	47.12%	uncertainty	1.37%	72.89%
point	1.10%	48.22%	seem	1.25%	74.15%
still	1.06%	49.28%	approximately	1.23%	75.38%
overall	1.05%	50.33%	event	1.06%	76.44%
confidence	1.03%	51.36%	limit	1.04%	77.49%
part	1.02%	52.38%	weakness	1.03%	78.51%
likely	1.01%	53.39%	question	1.01%	79.53%
fundamental	0.98%	54.37%	single	0.88%	80.41%
always	0.97%	55.34%	opinion	0.77%	81.17%
far	0.91%	56.25%	least	0.74%	81.92%
major	0.88%	57.13%	little	0.74%	82.66%
basis	0.87%	58.00%	rather	0.73%	83.39%

Panel B. Top 30 Confidence Words in Sample Letters

Machined-Based Lexicon		Human-Based Lexicon		LM Strong Modal Lexicon	
Conf. Word	% of Total Confidence Words in Letters	Conf. Word	% of Total Confidence Words in Letters	Conf. Word	% of Total Confidence Words in Letters
confidence	10.46%	lead	9.04%	best	47.39%
absolute	4.23%	significant	7.78%	always	28.19%
determine	4.12%	important	6.45%	must	7.40%
prove	3.68%	confidence	5.71%	strongly	6.19%
exceed	3.64%	major	4.89%	clearly	4.54%
robust	3.31%	certain	3.44%	never	4.43%
appropriate	3.30%	natural	2.52%	undoubtedly	0.82%
basic	2.96%	full	2.51%	definitely	0.52%
ensure	2.94%	fact	2.41%	unparalleled	0.24%
establish	2.82%	complete	2.35%	unequivocally	0.12%
emphasis	2.71%	absolute	2.31%	definitively	0.05%
therefore	2.66%	determine	2.24%	lowest	0.05%
purpose	2.28%	primary	2.07%	unambiguously	0.03%
essential	1.99%	prove	2.01%	unequivocal	0.02%
confident	1.96%	exceed	1.99%	undisputed	0.01%
exempt	1.59%	every	1.91%	uncompromising	0.00%
stabilize	1.57%	robust	1.81%	highest	0.00%
emphasize	1.53%	appropriate	1.80%	unsurpassed	0.00%
stability	1.42%	entire	1.70%		
necessary	1.39%	basic	1.61%		
rely	1.36%	ensure	1.61%		
importance	1.10%	quite	1.60%		
proprietary	1.09%	emphasis	1.48%		
decide	1.08%	clear	1.33%		
furthermore	1.06%	sound	1.19%		
immediate	0.96%	normal	1.18%		
moreover	0.96%	plus	1.10%		
consequence	0.89%	essential	1.08%		
prompt	0.85%	confident	1.07%		
assure	0.82%	indeed	1.05%		
Total	70.74%		79.23%		100%

Panel C. Top 30 Unconfidence Words in Sample Letters

Machined-Based Lexicon		Human-Based Lexicon		LM Weak Modal Lexicon		LM Uncertainty Lexicon	
Word	% of Total Unconfidence Words in Letters	Word	% of Total Unconfidence Words in Letters	Word	% of Total Unconfidence Words in Letters	Word	% of Total Unconfidence Words in Letters
well	16.81%	relative	22.30%	may	51.45%	may	21.26%
small	16.00%	less	18.49%	could	11.36%	risk	15.70%
less	8.18%	appear	10.22%	appear	9.39%	believe	15.25%
care	6.67%	uncertainty	6.93%	nearly	4.91%	volatility	6.72%
several	5.73%	seem	6.34%	almost	3.63%	exposure	5.37%
slow	5.10%	approximately	6.23%	possible	3.60%	could	4.69%
despite	4.69%	limit	5.28%	suggest	3.35%	volatile	2.64%
appear	4.52%	question	5.11%	might	3.25%	uncertainty	2.63%
near	3.87%	moderate	3.65%	somewhat	2.81%	fluctuate	2.37%
uncertainty	3.07%	cautious	1.81%	perhaps	1.65%	nearly	2.03%
seem	2.80%	perhaps	1.79%	uncertain	1.42%	assume	1.94%
approximately	2.76%	never	1.64%	depend	1.09%	almost	1.50%
question	2.26%	uncertain	1.55%	sometimes	0.91%	possible	1.49%
single	1.97%	comparative	1.42%	possibly	0.49%	suggest	1.38%
little	1.66%	unlikely	1.27%	occasionally	0.24%	might	1.34%
tend	1.57%	contingent	1.22%	maybe	0.21%	anticipate	1.23%
hard	1.51%	speculative	1.07%	apparently	0.17%	somewhat	1.16%
somewhat	1.35%	speculation	0.99%	seldom	0.06%	vary	1.14%
fair	1.35%	caution	0.86%	conceivable	0.02%	roughly	0.90%
light	0.87%	occasion	0.23%	uncertainly	0.00%	differ	0.88%
perhaps	0.79%	maybe	0.23%	seldomly	0.00%	predict	0.76%
bit	0.75%	sometime	0.22%			unpredictable	0.76%
never	0.73%	negligible	0.14%			cautious	0.69%
uncertain	0.69%	suggestion	0.14%			perhaps	0.68%
unlikely	0.56%	dim	0.12%			uncertain	0.59%
chance	0.46%	confuse	0.12%			contingent	0.46%
kind	0.38%	suppose	0.11%			turbulence	0.32%
something	0.36%	somewhere	0.10%			doubt	0.31%
nothing	0.29%	reservation	0.08%			instability	0.31%
impossible	0.26%	vacillate	0.07%			apparent	0.28%
Total	98.01%		99.71%		100%		96.79%

Table 2
Sample Shareholder Letter Characteristics and Fund Characteristics

Panels A reports the summary statistics of textual variables. Panel B (C) reports the pairwise correlations of textual confidence (unconfidence) variables. Panel D reports the pairwise correlations between net confidence variables and other textual variables. Panel E reports the summary statistics of fund characteristics. All variables are defined in Appendix A.

Panel A. Summary Statistics of Letter Characteristics

Variable	# of Obs.	Mean	SD	25th Percentile	Median	75th Percentile
Letter Characteristics: Proportional Ratios						
Confidence-M	10,813	0.0068	0.0056	0.0032	0.0055	0.0091
Confidence-H	10,813	0.0157	0.0085	0.0099	0.0143	0.0201
LM Strong Modal	10,813	0.0033	0.0036	0.0000	0.0023	0.0047
Unconfidence-M	10,813	0.0062	0.0048	0.0028	0.0056	0.0085
Unconfidence-H	10,813	0.0071	0.0053	0.0037	0.0066	0.0097
LM Weak Modal	10,813	0.0066	0.0053	0.0030	0.0059	0.0094
LM Uncertainty	10,813	0.0212	0.0113	0.0140	0.0205	0.0281
Net Confidence-M	10,813	0.0007	0.0071	-0.0031	0.0000	0.0041
Net Confidence-H	10,813	0.0086	0.0104	0.0021	0.0071	0.0138
Net LM Modal	10,813	-0.0033	0.0063	-0.0069	-0.0031	0.0000
Overstatement	10,813	0.0005	0.0010	0.0000	0.0000	0.0005
Understatement	10,813	0.0003	0.0008	0.0000	0.0000	0.0000
Net Overstatement	10,813	0.0002	0.0012	0.0000	0.0000	0.0000
Positive Tone	10,813	0.0304	0.0132	0.0215	0.0299	0.0391
Negative Tone	10,813	0.0307	0.0171	0.0191	0.0296	0.0418
Net Tone	10,813	-0.0002	0.0201	-0.0110	0.0000	0.0112
Letter Characteristics: tf.idf Scores						
tf.idf Confidence-M	10,813	0.1333	0.1036	0.0579	0.1142	0.1868
tf.idf Confidence-H	10,813	0.2278	0.1459	0.1196	0.2101	0.3054
tf.idf LM Strong Modal	10,813	0.0408	0.0336	0.0141	0.0346	0.0598
tf.idf Unconfidence-M	10,813	0.0998	0.0808	0.0347	0.0900	0.1484
tf.idf Unconfidence-H	10,813	0.0858	0.0694	0.0330	0.0747	0.1275
tf.idf LM Weak Modal	10,813	0.0568	0.0536	0.0126	0.0444	0.0856
tf.idf LM Uncertainty	10,813	0.2214	0.1482	0.1080	0.2066	0.3188
tf.idf Net Confidence-M	10,813	0.0339	0.1141	-0.0343	0.0230	0.0951
tf.idf Net Confidence-H	10,813	0.1420	0.1471	0.0431	0.1160	0.2139
tf.idf LM Net Modal	10,813	-0.0160	0.0540	-0.0449	-0.0095	0.0157
tf.idf Overstatement	10,813	0.0206	0.0408	0.0000	0.0000	0.0250
tf.idf Understatement	10,813	0.0115	0.0302	0.0000	0.0000	0.0000

tf.idf Net Overstatement	10,813	0.0090	0.0475	0.0000	0.0000	0.0074
tf.idf Positive Tone	10,813	0.3474	0.2034	0.1994	0.3292	0.4599
tf.idf Negative Tone	10,813	0.5697	0.3782	0.2771	0.5424	0.8041
tf.idf Net Tone	10,813	-0.2215	0.3288	-0.4050	-0.1948	-0.0013
Other Letter Characteristics						
Article Length	10,813	1248.21	1139.62	431.00	858.00	1683.00
Readability	10,813	13.72	2.35	12.21	13.49	14.85

Panel B. Correlations Matrix of Confidence-Related Variables

	(1)	(2)	(3)	(4)	(5)	(6)
(1) Confidence-M	1.0000					
(2) Confidence-H	0.6077***	1.0000				
(3) LM Strong Modal	0.3839***	0.3392***	1.0000			
(4) tf.idf Confidence-M	0.8343***	0.5096***	0.3165***	1.0000		
(5) tf.idf Confidence-H	0.5429***	0.7423***	0.318***	0.7358***	1.0000	
(6) tf.idf LM Strong Modal	0.1893***	0.2286***	0.6367***	0.3414***	0.4544***	1.0000

Panel C. Correlations Matrix of Unconfidence-Related Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Unconfidence-M	1.0000							
(2) Unconfidence-H	0.5545***	1.0000						
(3) LM Weak Modal	0.0654***	0.1692***	1.0000					
(4) LM Uncertainty	0.2902***	0.3093***	0.6018***	1.0000				
(5) tf.idf Unconfidence-M	0.8176***	0.4223***	0.1116***	0.3049***	1.0000			
(6) tf.idf Unconfidence-H	0.5074***	0.7879***	0.2038***	0.3234***	0.6341***	1.0000		
(7) tf.idf LM Weak Modal	0.1829***	0.2419***	0.6695***	0.4247***	0.3771***	0.4778***	1.0000	
(8) tf.idf LM Uncertainty	0.3140***	0.3084***	0.4025***	0.6753***	0.5540***	0.5831***	0.6737***	1.0000

Panel D. Correlation Matrix of Net Confidence Variables and other Text Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Net Confidence-M	1.0000							
(2) Net Confidence-H	0.6056***	1.0000						
(3) Overstatement	0.0307**	0.0165*	1.0000					
(4) Understatement	-0.0107	-0.0244**	0.0785***	1.0000				
(5) Net Overstatement	0.0315***	0.0301***	0.7846***	-0.5546***	1.0000			
(6) Positive Tone	0.0334***	0.1856***	0.0254***	0.0889***	-0.0613***	1.0000		
(7) Negative Tone	-0.1756***	0.0023	0.0313***	0.1146***	-0.079***	0.0944***	1.0000	
(8) Net Tone	0.1728***	0.1167***	0.0029	0.0010	0.0006	0.5530***	-0.7662***	1.0000

Panel E. Fund Characteristics

Fund Characteristics	# of Obs.	Mean	SD	25th Percentile	Median	75th Percentile
TNA ($ million)	10,813	2100.5	6519.12	63.80	219.90	992.30
Expense Ratio (%)	10,813	1.17	0.40	0.93	1.17	1.39
Portfolio Turnover Ratio	10,813	0.81	0.95	0.27	0.52	0.97
Fund Age (Years)	10,813	16.83	14.93	7.00	13.00	21.00
Annualized 6-month Net Ret (%)	10,813	5.24	27.7	-4.52	8.46	20.16
S.D. of Monthly Net Ret (%)	10,813	3.94	2.05	2.5	3.42	4.88
Annualized 6-month Net FF4 Alpha (%)	10,813	-1.20	7.78	-5.23	-1.11	2.97
6-month Net Flow Rate	10,813	0.03	0.25	-0.08	-0.02	0.06

Table 3
Confidence in Funds' Shareholder Letters and Future Performance

This table examines whether net confidence can predict funds' future performance. All variables are defined in Appendix A. *t*-statistics are in parentheses. Standard errors are clustered by fund. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable = FF4Alpha (t+1,t+6)									
	Proportional					tf.idf				
	Machine-Based	Machine-Based	Machine-Based (Annual Reports)	Human-Based	Modal-Based	Machine-Based	Machine-Based	Machine-Based (Annual Reports)	Human-Based	Modal-Based
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Net Confidence$_t$	-0.0362					-0.0014				
	(-0.24)					(-0.15)				
Top*Net Confidence$_t$		-0.1881	-0.9097**	-0.2529	-0.3571		-0.0079	-0.0498*	-0.0206	-0.0114
		(-0.57)	(-2.00)	(-1.11)	(-0.94)		(-0.39)	(-1.84)	(-1.34)	(-0.27)
Mid* Net Confidence$_t$		-0.3038*	-0.3063	-0.2638*	0.0272		-0.0170	-0.0190	-0.0135	0.0082
		(-1.67)	(-1.15)	(-1.89)	(0.13)		(-1.59)	(-1.27)	(-1.39)	(0.36)
Btm* Net Confidence$_t$		0.6397**	1.3342***	0.1915	0.1271		0.0477**	0.1003***	0.0165	0.0157
		(2.29)	(3.04)	(0.96)	(0.40)		(2.54)	(3.72)	(1.12)	(0.43)
Top		-0.0100***	-0.0095***	-0.0099***	-0.0111***		-0.0102***	-0.0086***	-0.0089***	-0.0102***
		(-4.74)	(-3.11)	(-3.58)	(-4.57)		(-4.71)	(-2.75)	(-2.96)	(-4.59)
Btm		0.0102***	0.0045***	0.0068**	0.0112***		0.0086***	0.0016	0.0065**	0.0110***
		(4.90)	(3.47)	(2.45)	(4.73)		(3.98)	(0.51)	(2.13)	(4.96)
Log TNA$_t$	-0.0075***	-0.0073***	-0.0078***	-0.0073***	-0.0072***	-0.0075***	-0.0073***	-0.0077***	-0.0073***	-0.0072***
	(-3.97)	(-3.80)	(-3.40)	(-3.78)	(-3.74)	(-3.97)	(-3.76)	(-3.36)	(-3.78)	(-3.73)
Expense Ratio$_t$	1.0764	1.0946	1.2261	1.1592	1.1244	1.0747	1.1105	1.3550	1.1238	1.1303
	(1.02)	(1.04)	(0.83)	(1.09)	(1.06)	(1.02)	(1.05)	(0.91)	(1.06)	(1.06)
Turnover$_t$	0.0028	0.0029	0.0022	0.0029	0.0028	0.0028	0.0029	0.0025	0.0029	0.0028
	(1.03)	(1.05)	(0.66)	(1.07)	(1.04)	(1.02)	(1.07)	(0.77)	(1.08)	(1.04)
Fund Age$_t$	-0.0002	-0.0001	0.0004	-0.0001	-0.0001	-0.0002	-0.0001	0.0004	-0.0001	-0.0001
	(-0.33)	(-0.23)	(0.78)	(-0.15)	(-0.17)	(-0.32)	(-0.24)	(0.70)	(-0.17)	(-0.18)
Readability$_t$	0.0005	0.0004	0.0001	0.0004	0.0004	0.0005	0.0004	0.0001	0.0004	0.0004
	(0.98)	(0.93)	(0.17)	(0.91)	(0.89)	(0.98)	(0.93)	(0.16)	(0.90)	(0.90)
Article Length$_t$	0.0000**	0.0000**	0.0000	0.0000**	0.0000**	0.0000**	0.0000**	0.0000	0.0000**	0.0000**
	(2.23)	(2.22)	(0.82)	(2.09)	(2.24)	(2.23)	(2.22)	(0.81)	(2.23)	(2.27)
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,813	10,813	5,622	10,813	10,813	10,813	10,813	5,622	10,813	10,813
R-squared	0.15	0.15	0.23	0.15	0.15	0.15	0.15	0.23	0.15	0.15

Table 4
The Information Contents of Confidence and Unconfidence

This table examines whether textual confidence and unconfidence can predict funds' future performance. All variables are defined in Appendix A. *t*-statistics are in parentheses. Standard errors are clustered by fund. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A. The Information Content of Confidence

	Dependent Variable = FF4Alpha (t+1,t+6)							
	Proportional				tf.idf			
	Machine-Based	Machine-Based (Annual Reports Only)	Human-Based	Modal-Based	Machine-Based	Machine-Based (Annual Reports Only)	Human-Based	Modal-Based
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Top* Confidence$_t$	-0.1453	-1.1589**	-0.5925**	-0.6385	-0.0032	-0.0557**	-0.0215	-0.0476
	(-0.36)	(-2.14)	(-2.17)	(-1.07)	(-0.15)	(-2.06)	(-1.43)	(-0.76)
Mid* Confidence$_t$	-0.4704**	-0.7007**	-0.4307***	-0.2894	-0.0006	-0.0065	-0.0085	0.0560
	(-2.00)	(-2.07)	(-2.63)	(-0.71)	(-0.05)	(-0.37)	(-0.85)	(1.36)
Btm* Confidence$_t$	0.9215***	1.6431***	0.0333	0.5367	0.0594***	0.0969***	0.0117	0.0547
	(2.59)	(3.13)	(0.14)	(0.97)	(2.97)	(3.57)	(0.80)	(0.80)
Top	-0.0121***	-0.0064**	-0.0075***	-0.0088***	-0.0095***	-0.0029**	-0.0069****	-0.0056***
	(-3.50)	(-2.28)	(-2.59)	(-2.87)	(-2.72)	(-2.57)	(-2.72)	(-2.60)
Btm	0.0014***	-0.0100**	0.0035***	0.0081***	0.0030***	-0.0078	0.0063**	0.0110***
	(3.43)	(-2.10)	(2.76)	(2.82)	(2.88)	(-1.62)	(2.58)	(3.19)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund FE.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,813	5,622	10,813	10,813	10,813	10,813	10,813	10,813
R-squared	0.15	0.23	0.15	0.15	0.15	0.16	0.15	0.16

Table 4 - Continued

Panel B. The Information Content of Unconfidence

	Dependent Variable = FF4Alpha (t+1,t+6)									
	Proportional					tf.idf				
	Machine-Based	Machine-Based (Annual Reports)	Human-Based	Modal-Based	Uncertainty-Based	Machine-Based	Machine-Based (Annual Reports)	Human-Based	Modal-Based	Uncertainty-Based
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Top* Unconfidence $_t$	-0.0085	0.1173	-0.5212	0.1897	-0.0347	0.0161	0.0055	0.0020	-0.0171	-0.0037
	(-0.02)	(0.21)	(-1.34)	(0.47)	(-0.19)	(0.66)	(0.16)	(0.08)	(-0.46)	(-0.27)
Mid* Unconfidence $_t$	0.0113	-0.1805	-0.0795	-0.1494	0.0402	0.0319**	0.0345	0.0247	0.0180	0.0135
	(0.05)	(-0.47)	(-0.33)	(-0.62)	(0.34)	(2.15)	(1.54)	(1.31)	(0.74)	(1.40)
Btm* Unconfidence $_t$	-0.0919	-0.4119	-0.4864	0.0917	0.1251	0.0117	-0.0172	-0.0165	0.0077	0.0054
	(-0.24)	(-0.75)	(-1.27)	(0.26)	(0.77)	(0.51)	(-0.51)	(-0.54)	(0.20)	(0.41)
Top	-0.0098***	-0.0115**	-0.0069*	-0.0121***	-0.0083*	-0.0082**	-0.0067	-0.0079**	-0.0079**	-0.0061
	(-2.77)	(-2.13)	(-1.81)	(-3.41)	(-1.78)	(-2.36)	(-1.28)	(-2.25)	(-2.46)	(-1.51)
Btm	0.0115***	0.0069	0.0137***	0.0093***	0.0091**	0.0130***	0.0107**	0.0144***	0.0115***	0.0127***
	(3.43)	(1.34)	(3.80)	(2.81)	(2.15)	(3.93)	(2.15)	(4.18)	(3.69)	(3.41)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund FE.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,813	5,622	10,813	10,813	10,813	10,813	5,622	10,813	10,813	10,813
R-squared	0.15	0.23	0.15	0.15	0.15	0.15	0.23	0.15	0.15	0.15

Table 5
Confidence vs. Overstatement

This table examines whether textual over- or under-statement can predict funds' future performance. All variables are defined in Appendix A. *t*-statistics are in parentheses. Standard errors are clustered by fund. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable = FF4Alpha (t+1,t+6)							
	Proportional				tf.idf			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Net Overstatement$_t$	-1.7754**				-0.0360*			
	(-2.42)				(-1.88)			
Top*Net Overstatement$_t$		-7.5130***				-0.1713***		
		(-3.57)				(-3.77)		
Mid*Net Overstatement$_t$		0.5429				0.0194		
		(0.49)				(0.76)		
Btm*Net Overstatement$_t$		-2.7240				-0.0432		
		(-1.41)				(-0.88)		
Top*Overstatement$_t$			-6.8733***	-6.8345***			-0.1464***	-0.1422***
			(-3.22)	(-3.20)			(-3.10)	(-3.01)
Mid*Overstatement$_t$			0.9982	1.0961			0.0509	0.0544*
			(0.76)	(0.83)			(1.61)	(1.71)
Btm*Overstatement$_t$			0.3210	0.3865			0.0624	0.0520
			(0.16)	(0.19)			(1.16)	(0.97)
Top*Understatement$_t$			7.4349**	7.4590**			0.1829**	0.1870**
			(1.98)	(1.99)			(2.48)	(2.55)
Mid* Understatement$_t$			0.6912	0.7549			0.0293	0.0289
			(0.38)	(0.42)			(0.76)	(0.75)
Btm*Understatement$_t$			5.4693*	5.7195*			0.1130	0.1052
			(1.78)	(1.86)			(1.44)	(1.35)
Top*Confidence-M$_t$				-0.1742				-0.0036
				(-0.44)				(-0.18)
Mid*Confidence-M$_t$				-0.4993**				-0.0187*
				(-2.12)				(-1.76)
Btm*Confidence-M$_t$				0.9099**				0.0519***
				(2.55)				(2.69)
Top		-0.0081***	-0.0078***	-0.0100***		-0.0080***	-0.0073***	-0.0075**
		(-3.75)	(-3.17)	(-2.65)		(-3.66)	(-2.95)	(-2.15)
Btm		0.0116***	0.0099***	0.0003***		0.0115***	0.0095***	0.0024***
		(5.36)	(4.03)	(3.08)		(5.29)	(3.75)	(2.67)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,813	10,813	10,813	10,813	10,813	10,813	10,813	10,813
R-squared	0.15	0.15	0.15	0.16	0.15	0.15	0.15	0.16

Table 6
Confidence vs. Tone

This table examines whether textual tone can predict funds' future performance. All variables are defined in Appendix A. *t*-statistics are in parentheses. Standard errors are clustered by fund. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable = FF4Alpha (t+1,t+6)							
	Proportional				tf.idf			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Net Tone$_t$	-0.1826***				-0.0121***			
	(-3.24)				(-3.26)			
Top*Net Tone$_t$		0.0359				0.0032		
		(0.33)				(0.48)		
Mid*Net Tone$_t$		-0.0562				-0.0051		
		(-0.85)				(-1.26)		
Btm*Net Tone$_t$		-0.4026***				-0.0244***		
		(-3.88)				(-3.77)		
Top*Positive Tone$_t$			0.0548	0.0717			0.0149	0.0158
			(0.32)	(0.41)			(1.29)	(1.32)
Mid* Positive Tone$_t$			-0.1300	-0.0929			0.0005	0.0014
			(-1.33)	(-0.95)			(0.07)	(0.20)
Btm* Positive Tone$_t$			-0.3165**	-0.3796**			-0.0304**	-0.0369***
			(-2.06)	(-2.48)			(-2.55)	(-3.06)
Top* Negative Tone$_t$			-0.1265	-0.1207			-0.0049	-0.0041
			(-0.96)	(-0.92)			(-0.73)	(-0.61)
Mid* Negative Tone$_t$			-0.0165	-0.0125			0.0059	0.0063
			(-0.20)	(-0.15)			(1.38)	(1.47)
Btm* Negative Tone$_t$			0.3850***	0.3800***			0.0232***	0.0201***
			(3.35)	(3.31)			(3.63)	(3.14)
Top*Confidence-M$_t$				-0.1970				-0.0076
				(-0.48)				(-0.33)
Mid*Confidence-M$_t$				-0.4609*				-0.0069
				(-1.95)				(-0.53)
Btm*Confidence-M$_t$				0.9693***				0.0580***
				(2.74)				(2.81)
Top		-0.0101***	-0.0125*	-0.0138**		-0.0081***	-0.0091*	-0.0092*
		(-4.70)	(-1.84)	(-1.96)		(-3.34)	(-1.95)	(-1.92)
Btm		0.0089***	0.0026***	0.0040***		0.0050**	0.0100**	0.0059**
		(4.39)	(3.44)	(3.63)		(2.05)	(2.33)	(2.28)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,813	10,813	10,813	10,813	10,813	10,813	10,813	10,813
R-squared	0.15	0.15	0.15	0.16	0.15	0.15	0.15	0.16

Table 7
Investors' Reactions to Textual Confidence

This table examines whether investors respond to the writing styles of funds' shareholder letters. All textual and baseline control variables are defined in Appendix A. Additional control variables include funds' flow rates in the last six months, funds' past six-month returns, and the squared past six-month returns. *t*-statistics are in parentheses. Standard errors are clustered by fund. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Net flow (t+1, t+6)											
	Proportional						tf.idf					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Confidence-M$_t$	0.8838						0.0242					
	(1.02)						(1.04)					
Confidence-H$_t$		0.2665						-0.0060				
		(0.96)						(-0.31)				
Net Overstatement$_t$			-2.0314						-0.0580			
			(-1.37)						(-1.64)			
Top*Confidence-M$_t$				1.1642						0.0416		
				(1.55)						(1.07)		
Mid*Confidence-M$_t$				0.7720						0.0238		
				(1.49)						(0.93)		
Btm*Confidence-M$_t$				0.6934						-0.0099		
				(0.97)						(-0.25)		
Top*Confidence-H$_t$					0.2393						0.0015	
					(0.48)						(0.05)	
Mid*Confidence-H$_t$					0.1802						-0.0056	
					(0.58)						(-0.27)	
Btm*Confidence-H$_t$					0.2333						-0.0286	
					(0.49)						(-0.94)	
Top*Net Overstatement$_t$						-4.3914						-0.0764
						(-1.04)						(-0.86)
Mid*Net Overstatement$_t$						-1.1118						-0.0318
						(-0.50)						(-0.63)
Btm*Net Overstatement$_t$						-6.4522*						-0.1645**
						(-1.96)						(-2.09)
Top				0.0227***	0.0243**	0.0260***				0.0229***	0.0235***	0.0257***
				(3.08)	(2.47)	(4.93)				(3.01)	(2.74)	(4.87)
Btm				-0.0323***	-0.0337***	-0.0318***				-0.0286***	-0.0280***	-0.0318***
				(-4.60)	(-3.79)	(-6.17)				(-4.02)	(-3.53)	(-6.17)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,813	10,813	10,813	10,813	10,813	10,813	10,813	10,813	10,813	10,813	10,813	10,813
R-squared	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35

Internet Appendix

A. Top 30 Over- and Under-statement Words in Funds' Shareholder Letters

Top 30 Overstatement Words in Sample Letters	% of Total Overstatement Word Count	Cumulative %	Top 30 Understatement Words in Sample Letters	% of Total Understatement Word Count	Cumulative %
endure	10.37%	10.37%	unpredictable	37.68%	37.68%
predictable	6.12%	16.49%	hesitate	7.33%	45.01%
undoubtedly	4.20%	20.69%	falter	7.16%	52.18%
unwavering	2.96%	23.65%	fortunate	4.25%	56.43%
overwhelm	2.91%	26.56%	anomaly	3.58%	60.01%
steadfast	2.77%	29.33%	dim	2.21%	62.22%
uncommon	2.67%	32.00%	confuse	2.18%	64.39%
manageable	2.64%	34.64%	suppose	2.04%	66.43%
shun	2.42%	37.06%	hesitant	1.61%	68.04%
incredible	2.12%	39.18%	unpleasant	1.54%	69.58%
stark	2.00%	41.18%	unreliable	1.54%	71.12%
magnify	1.90%	43.08%	dilemma	1.41%	72.52%
abundance	1.78%	44.86%	unfortunate	1.41%	73.93%
dependable	1.58%	46.44%	vacillate	1.37%	75.30%
perfection	1.48%	47.92%	puzzle	1.27%	76.57%
exuberance	1.43%	49.35%	heed	1.17%	77.74%
perpetual	1.43%	50.78%	calamity	1.04%	78.78%
catastrophe	1.21%	51.99%	accustom	0.94%	79.72%
plentiful	1.21%	53.20%	worthless	0.94%	80.66%
unavoidable	1.09%	54.29%	confound	0.87%	81.53%
predominant	1.06%	55.35%	ironic	0.80%	82.33%
exuberant	1.01%	56.36%	precarious	0.77%	83.10%
solace	0.99%	57.35%	insignificant	0.74%	83.84%
commonplace	0.96%	58.31%	murky	0.70%	84.54%
avid	0.94%	59.25%	insecurity	0.67%	85.21%
havoc	0.91%	60.16%	coincidence	0.60%	85.81%
necessitate	0.89%	61.05%	doubtful	0.60%	86.41%
dash	0.86%	61.91%	faint	0.60%	87.01%
endurance	0.84%	62.75%	weakly	0.60%	87.62%

B. Top 30 Positive/Negative Tone Words in Funds' Shareholder Letters

Top 30 Positive Tone Words in Sample Letters	% of Total Positive Tone Word Count	Cumulative %	Top 30 Negative Tone Words in Sample Letters	% of Total Negative Tone Word Count	Cumulative %
strong	9.81%	9.81%	decline	6.13%	6.13%
gain	8.80%	18.61%	volatility	4.31%	10.43%
good	5.94%	24.55%	concern	3.60%	14.03%
opportunity	5.82%	30.37%	loss	3.59%	17.62%
benefit	5.19%	35.56%	slow	2.81%	20.42%
positive	4.94%	40.50%	crisis	2.69%	23.11%
best	4.55%	45.05%	negative	2.61%	25.73%
great	3.56%	48.61%	weak	2.49%	28.22%
outperform	3.30%	51.91%	recession	2.03%	30.25%
attractive	3.14%	55.05%	poor	1.92%	32.17%
improve	2.91%	57.95%	bad	1.85%	34.02%
despite	2.75%	60.70%	late	1.80%	35.82%
advantage	1.99%	62.68%	challenge	1.74%	37.55%
strength	1.78%	64.46%	volatile	1.69%	39.25%
achieve	1.38%	65.84%	unemployment	1.54%	40.79%
rebound	1.32%	67.16%	fear	1.33%	42.12%
pleased	1.31%	68.47%	difficult	1.32%	43.44%
success	1.30%	69.78%	weakness	1.26%	44.70%
improvement	1.27%	71.05%	question	1.24%	45.94%
favorable	1.22%	72.27%	detract	1.17%	47.11%
reward	1.19%	73.46%	cut	1.16%	48.27%
able	1.05%	74.51%	underperform	1.16%	49.43%
effective	1.03%	75.54%	lag	1.16%	50.59%
sable	1.02%	76.56%	lose	1.15%	51.74%
boost	0.97%	77.53%	problem	1.09%	52.83%
progress	0.80%	78.32%	correction	1.06%	53.89%
optimistic	0.76%	79.08%	slowdown	1.00%	54.88%
superior	0.70%	79.77%	hurt	0.91%	55.80%
excellent	0.69%	80.46%	sharply	0.84%	56.63%
successful	0.68%	81.14%	downturn	0.81%	57.44%