When Words Move Money:

Managerial Rhetorical Expressions and Investor Flows in Mutual Funds

Bing Han [1], Yuanyu Qu[2], Yankai Sheng [3]

Abstract: We study how investors respond to rhetorical expressions in mutual fund managers' communications. We identify rhetorical devices in manager-level disclosures in China using Large Language Models and document that hyperbole and antithesis are the most common forms of rhetoric. Funds with more intensive use of rhetorical expressions attract greater investor inflows after controlling for established textual and fund characteristics. This effect is concentrated among smaller funds and those primarily held by retail investors. Transaction-level evidence further shows that the effect operates primarily through first-time purchases. However, rhetorical expressions do not signal superior managerial skill. Instead, managers who rely more on rhetoric take on greater investment risk and deviate more from their benchmarks. Moreover, rhetoric amplifies flow–performance sensitivity and leads to sharper outflows following underperformance, underscoring the potential costs of promotional narratives.

JEL classification: G11, G23

Keywords: mutual funds, rhetoric devices, narratives, fund flows, LLM

[1] Rotman School of Management, University of Toronto, Toronto, Ontario, Canada M5S3E6, Bing.Han@Rotman.Utoronto.Ca.
[2] China School of Banking and Finance, University of International Business and Economics, Beijing 100029, China, quyuanyu@uibe.edu.cn.
[3] School of Accounting and Finance, Hong Kong Polytechnic University, Hong Kong 999077, yankai.sheng@connect.polyu.hk.

1. Introduction

Does the way information is communicated matter in financial markets? While traditional research has focused primarily on the content of disclosures, a growing body of evidence highlights the importance of narratives in shaping investor beliefs and capital allocation (Kostovetsky & Warner, 2020; Abis & Lines, 2024; Hillert et al., 2024). Holding fixed the economic content of the communications, how the information is framed and articulated can meaningfully influence investors' perceptions and decisions. It suggests that communication style may play an independent role in financial decision-making.

One important yet understudied dimension of managerial communications is rhetorical expression. Unlike general tone or linguistic complexity, rhetorical expression reflects the deliberate use of persuasive language to emphasize, contrast, or frame information without altering its underlying content (McCloskey, 1983; Back, 1989; Hoefer and Green, 2016). In practice, rhetorical expression often takes the form of specific rhetorical devices, such as hyperbole, antithesis, and metaphor, which allow managers to selectively highlight favorable aspects of performance or strategically downplay risk or unfavorable outcomes. These devices can shape how investors interpret disclosed information by directing attention, strengthening emphasis, or creating salient comparisons, even when the underlying facts remain unchanged (Charteris-Balck, 2011 pp.16; Delbaere et al., 2011).

This rhetorical channel of managerial communications is especially salient in mutual fund disclosures. Mutual funds rely heavily on narratives in their reports to communicate with investors whom are typically retail investors with limited ability to process complex quantitative information. This reliance on qualitative communication creates scope for rhetorical expression to influence investor perceptions and capital allocation, above and beyond disclosure of fund performance. Yet, despite the central role of narratives in mutual fund

disclosures, we know little about whether and how rhetorical expression shapes investor behavior and managerial decisions.

In this paper, we examine how rhetorical expression in mutual fund managers' disclosures relates to investor behavior. We focus on narrative sections of mutual fund reports in China, where disclosures are issued at the individual fund level and managerial responsibility is typically concentrated in a single manager, allowing narrative choices to be attributed to individual managers. In this setting, we identify rhetorical expression by detecting specific rhetorical devices in manager-level communications using Large Language Models (LLMs). By construction, this measure captures deliberate, device-based variation in how information is presented while holding constant the underlying information disclosed. LLMs allow us to identify such rhetorical devices directly, which is difficult to achieve using traditional dictionary-based or readability measures. Using both fund-level and transaction-level data, we analyse investor responses to rhetorical expression and examine related patterns in managerial risk-taking and investor discipline.

We first document moderate adoption of rhetorical devices, with fund reports containing 1.08 rhetorical devices per document on average. Hyperbole and antithesis are the most common rhetorical devices, each appearing in about 30% of fund reports. Metaphor, another rhetorical device widely documented in prior linguistic studies, is used in approximately 10% of fund reports. Rhetorical intensity is related to other textual features of fund reports, exhibiting positive associations with sentiment and disclosure length, but a negative association with readability.

Then we examine the relationship between rhetorical expressions and subsequent mutual fund flow. If rhetorical language functions as a persuasion device, its effect should be manifested in investors' capital allocation decisions. Our empirical evidence show that more rhetorical expressions are correlated with higher subsequent fund flows. This conclusion

remains significant after we control for other textual characteristics (Hwang & Kim 2017; Hillert et al. 2025). Consistent with persuasion effect (DeHaan et al., 2026), our heterogeneity analysis illustrates that such effect is more pronounced for small funds and funds dominated by retail investors, which have less established reputation and less sophisticated investor base.

Leveraging unique transaction-level investor data including 200,000 investors, we further explore the micro-level mechanism through which rhetorical language influences investment behavior. Investors who are unfamiliar with a fund are more likely to rely on impressive narratives when forming first-time purchase decision, whereas experienced investors can draw on past information and holding experience, reducing the marginal impact of rhetorical language (Döbeli & Vanini, 2010). We find that rhetorical expressions exert their positive effect at the point of first-time purchase, while their effect reverse once investors have established experience with the fund.

We also run several tests to alleviate endogeneity problems. First, leveraging the regulatory disclosure schedule and transaction-level data, we conduct a timing test and show that rhetorical expressions predict investor flows only in the disclosure month, but not beforehand. Second, our empirical evidence remains robust after controlling for manager or family fixed effects, suggesting that the documented effects are not driven by managerial characteristics or family-level disclosure practices. Third, we exploit the 2019 regulatory policy as a quasi-experiment that strengthened mutual fund disclosure requirements. We find that the persuasive effect of rhetorical language weakens significantly following the policy change, consistent with tighter constraints on discretionary narrative framing. Finally, to alleviate potential misclassification in our rhetorical measures, we convert the continuous variables into dummies and examine the effects of different rhetorical devices. The results are robust with hyperbole and metaphor particularly effective in attracting fund flows.

Then we explore whether investors benefit from chasing rhetorical expressions. We first examine whether rhetorical intensity predicts subsequent fund performance. While the predictive power of mutual fund narratives remains debated (e.g., Chu & Kim, 2019; Hillert et al., 2025), rhetorical expressions—unlike tone, which may convey managerial expectations or information (Li, 2010)—capture a purely stylistic dimension, potentially providing a cleaner measure. Our empirical results reveal that funds employing more rhetorical expressions cannot generate superior performance. This result is robust using raw returns and multiple risk-adjusted alphas, suggesting that rhetorical expressions do not convey information about managers' stock-picking ability or future performance. However, we find that rhetorical expressions are associated with higher future risk-taking. Such managerial behavior pattern is mainly driven by deviating more from their stated benchmarks.

While rhetorical expressions can attract investor attention and shape expectations, they may also entail potential costs when such expectations are not fulfilled. By making narratives more salient, rhetorical language heightens investors' sensitivity to subsequent outcomes. In practice, these expectations are ultimately evaluated against realized returns: when subsequent performance falls short, investors respond more strongly, often through increased redemptions. Consistent with this mechanism, our empirical results show that the effect of rhetorical intensity on fund flows depends critically on subsequent performance. Rhetorical expressions are rewarded when followed by strong returns but penalized when subsequent performance is weak.

Our paper makes several contributions to the literature. First, we contribute to the persuasive effect of marketing in mutual fund market. Previous studies show that mutual funds' marketing and advertising efforts, such as increased visibility through media coverage and promotional expenditure, significantly affect fund flows (Sirri & Tufano, 1998; Jain & Wu, 2000; Gallaher et al., 2015; Roussanov et al., 2021). Investors can be swayed by such marketing efforts even when they provide limited informational content about underlying fund quality

(Cronqvist, 2006; Mullainathan & Shleifer, 2005). In practice, mutual fund advertising takes many forms, including newspaper postings (Jain & Wu, 2000), print and television ads (Aydogdu & Wellman, 2011), as well as interactive formats like tweeting (Gil-Bazo & Imbet, 2025) and livestreaming (DeHaan et al., 2026). In our paper, we illustrate the persuasive power of narratives in fund reports. Specifically, we find that the rhetorics in mutual fund reports can attract investor flow although they do not signal superior performance subsequently.

Secondly, we contribute to a recent line of research on narratives. Prior studies show that strategic use of narratives can influence investor behavior, such as the textual uniqueness of investment strategies (Kostovetsky & Warner, 2020), elaboration on strategy descriptions (Abis et al., 2022), and disclosure of private information (Cao et al., 2024). When mutual funds walk the talk, such as investment strategies (Abis & Lines, 2024) or risk communication (Sheng et al., 2025), investors tend to reward them with inflows. However, literature also reveals the potential cost of strategic information disclosure. Hwang and Kim (2017) find that low readability of shareholder letters leads to more discount for closed-end funds. DeHaan et al. (2021) discover that managers tend to use complex narratives to obfuscate high fees, raising investors' learning costs. Our research illustrates that the strategic adoption of rhetoric by mutual fund managers can not only attract investor inflow but also lead to potential costs when performance is inferior. Hillert et al. (2025) also demonstrate that extremely positive communication accompanied by inconsistent performance can lead to investor outflow.

Third, our research contributes to a rapidly growing literature on the application of LLMs in finance. The recent advent of LLMs has influenced not only the corporates practices but also the financial academia (see a comprehensive review by Eisfeldt & Schubert (2024)). This technique is especially crucial for textual analysis in financial contexts, such as financial news (Chen et al., 2022; Tan et al., 2023; Chen et al., 2025), conference calls (Kim et al., 2023), annual reports (Niu et al., 2025; Breitung & Müller, 2025), loan descriptions (Cong et al., 2024),

analyst reports (Li et al., 2026) and so on. However, relatively little work has applied LLMs to mutual fund disclosures. One notable exception is Gao et al. (2024), who primarily examine the relationship between the sentiment of fund reports and subsequent returns. Rhetoric relies heavily on the contextual information and relationship across sentences, where LLMs can be particularly effective. Our research utilizes LLMs to analyse the rhetorical language in mutual fund reports and reveals how such disclosure influences investors' decisions.

The remainder of this paper is organized as follows. Section 2 presents the institutional background of Chinese mutual fund disclosure and rhetorical expressions. Section 3 displays data, variable construction, and descriptive statistics. Section 4 analyzes the impact of rhetorical expressions on subsequent mutual fund flow. Section 5 uses several methods to address endogeneity concerns. Section 6 explores the relationship between rhetorical expressions and managerial subsequent behavior. Section 7 concludes the paper.

2. Institutional Background

2.1 Managers' Discussions on Chinese Mutual Fund Reports

The CSRC requires mutual fund managers to explain their fund's investment strategy and performance, with reference to prevailing macroeconomic and market conditions in the *Operational Analysis* section of quarterly reports. Such review of performance and macro economy parallels the shareholder letters included in N-CSR/N-CSRS filings submitted to the SEC. Neither disclosure type is strictly limited to specific topics, resulting in substantial variation in content that typically covers fund performance, macroeconomic and market conditions, and investment strategies.

Despite these similarities, several important differences distinguish Chinese *Operational Analysis* from shareholder letters. First, N-CSR/N-CSRS filings in the U.S. are prepared at the

fund family level and often cover multiple funds within a single report[4]. In contrast, Chinese mutual fund reports are issued at the individual fund level, granting fund managers greater flexibility to express their own narrative style. Second, the SEC permits U.S. mutual funds to discuss specific securities in their portfolios, including related operations and analyses. However, Chinese fund managers are prohibited from naming individual securities in their disclosures to prevent the potential transmission of insider information. Finally, regarding disclosure frequency, U.S. N-CSR/N-CSRS filings are submitted semi-annually, while Chinese fund reports are released quarterly, offering a more timely and dynamic perspective. In conclusion, Chinese mutual funds provide us with a unique setting to explore the personal narrative style of mutual fund managers.

2.2 Rhetoric and LLM

Rhetoric is the "disciplined use of language in communication to shape how information is interpreted and weighted, thereby influencing decisions under conditions of limited attention and cognitive resources" (McCloskey, 1983; Back, 1989; Hoefer and Green, 2016). Rhetoric is closely linked to persuasion: by carefully choosing words and structuring narratives, communicators can make certain interpretations more compelling by rhetoric (Young, 2003; Herrick, 2020 pp.3). This persuasive power explains why rhetorical devices are widely used across contexts. For example, politicians sometimes use metaphor as a persuasive instrument to advocate their own policies or oppose the policies of others (Charteris-Black, 2011 pp.16). Additionally, Delbaere et al. (2011) illustrate that personification can lead to more positive emotions and greater brand liking from consumers in advertising field. In the context of mutual funds, managers' deliberate use of rhetorical language shapes how investors perceive and interpret fund disclosures, thereby influencing their beliefs and investment decisions.

[4] As summarized by Hillert et al. (2025), the N-CSR filings contain 8.3 unique funds on average.

Traditional natural language processing (NLP) approaches face inherent limitations in capturing rhetorical expressions in financial communication. Rhetoric is not defined by the frequency of specific words, but by how economic information is structured and emphasized to guide interpretation (McCloskey, 1983; McCloskey, 1998 pp.6). Accordingly, rhetorical expressions are contextual, which makes them difficult to measure using conventional NLP methods. For instance, dictionary-based approaches proxy textual attributes through predefined word lists (Loughran and McDonald, 2016; Frankel et al., 2022). In practice, however, the same lexical item can serve different rhetorical functions depending on context, causing dictionary measures to miss substantial variation in rhetorical expression. Supervised machine-learning methods partially mitigate these concerns by learning patterns from labelled data, but their ability to identify rhetoric is constrained by the nature of the training labels (Bochkay et al., 2023). Such labels typically fail to distinguish among multiple rhetorical devices that may coexist within the same disclosure.

In recent years, LLMs pretrained on enormous texts illustrate effectiveness in capturing semantic and contextual information in long texts (Huang et al. 2023; Tan et al. 2023; Chen et al. 2025). By learning relationships across sentences and paragraphs and integrating information over extended contexts, LLMs are able to recognize narrative structure beyond isolated words. These capabilities allow LLMs to capture how information is emphasized and connected within a disclosure, making them particularly necessary as traditional NLP struggles to capture these rhetorical patterns.

3. Data and Variable Construction

3.1 Measurement of Rhetorical Expression

Prior research suggests that language models are often language-specific, and cross-language application can pose challenges (Du et al., 2022; Chen et al., 2025; Cao et al., 2025).

This is also relevant for rhetorical analysis, as rhetorical devices are closely tied to linguistic conventions (Michael-Luna & Canagarajah, 2015). In this paper, we adopt a Chinese LLM developed by the Qwen team at Alibaba Group, namely *Qwen-3-Max*, released in September 2025. The model is pre-trained on approximately 36 trillion tokens and has over one trillion parameters, enabling it to capture rich contextual and structural features of natural language (Yang et al., 2025).

To measure rhetorical intensity, we adopt a few-shot prompting strategy with an explicit role setting for the LLM[5]. First, assigning the model the role of a fund investor anchors its interpretation to financially relevant reading norms (Bini et al., 2025), ensuring that rhetorical intensity is evaluated from the perspective of investors rather than from the perspective from linguists or psychologists. Second, compared with zero-shot prompt, a few-shot prompt reduces ambiguity in identifying rhetorical devices by explicitly illustrating how different rhetorical forms operate in fund narratives, which is essential given the contextual and interpretive nature of rhetoric (De Kok, 2025).

We instruct the model to identify and quantify eight types of rhetorical devices, namely *Metaphor*, *Personification*, *Hyperbole*, *Antithesis*, *Parallelism*, *Quotation*, *Hypophora*, and *Rhetorical Question*. These devices are widely recognized in rhetoric and communication studies as core tools for persuasion, emphasis, and audience engagement, and are among the most prevalent and salient forms of rhetorical expression in business and financial discourse[6]. By focusing on these well-established rhetorical strategies, we capture the primary mechanisms through which fund managers can shape narratives and influence investor perceptions.

[5] The prompt is Chinese when we import it into the LLM. Appendix Table A1 shows both the English and Chinese version.
[6] Merriam-Webster, one the largest and oldest American reference book publisher, has listed the most commonly used rhetorical devices, which includes our listed devices. See more on https://www.merriam-webster.com/grammar/rhetorical-devices-list-examples. Moreover, the Speaker Lab, a practitioner-oriented platform focused on public speaking and persuasive communication, also listed the power rhetorical devices in communication. See more on https://thespeakerlab.com/blog/rhetorical-devices-examples/.

Our first measure, *REScore*, is the overall rhetorical intensity score directly generated by the LLM. *REScore* ranges from 0 to 100 and reflects the model's assessment of rhetorical intensity from the perspective of an individual investor reading the report. Importantly, this score is not a mechanical aggregation of specific rhetorical devices. Instead, it synthesizes multiple aspects of rhetorical expression into a document-level evaluation, such as the frequency, distribution and salience of rhetorical devices throughout the disclosure.

The second measure, *RERatio*, captures the relative frequency of rhetorical device usage. Specifically, we sum the # of all eight rhetorical devices identified by the LLM and scale this value by the number of sentences in the report as formula (1):

$$RERatio_{i,q} = \frac{\sum_{r=1}^{8} \#\ of\ Rhetorical\ Device_{i,r,q}}{\#\ of\ sentences_{i,q}} \quad (1)$$

where i, r, q denotes fund, rhetorical devices and quarter respectively. $\#\ of\ Rhetorical\ Device_{i,r,q}$ is the frequency of rhetorical device r in report of fund i in quarter q. This normalization accounts for variation in disclosure length and captures the density of rhetorical expressions. Compared with *REScore*, *RERatio* offers a more objective measure by quantifying the actual frequency of rhetorical devices relative to the number of sentences.

These two measures capture rhetorical intensity from complementary perspectives. *REScore* reflects the overall impression of rhetoric in a report, synthesizing frequency, distribution, and salience of rhetorical devices into a holistic document-level evaluation. *RERatio* quantifies the relative frequency of specific rhetorical devices normalized by report length, providing a transparent, density-based view of rhetorical expression. Together, they offer a more comprehensive picture of how rhetorical language is deployed, combining perceptual and quantitative aspects.

We compare the two measures of rhetorical expressions with other textual features including sentiment (Hillert et al., 2025), length and readability (Hwang & Kim, 2017). Following Antweiler & Frank (2005), we generate the sentiment index using positive and

negative words in mutual fund disclosures as identified by the financial sentiment dictionary of Jiang et al. (2019). The length of mutual fund disclosures, denoted by *Length*, is measured by the logarithm of the number of characters in fund reports. Following Loughran and McDonald (2014), we define *Readability* as the negative logarithm of the average sentence length, since longer sentences are more difficult to comprehend.

3.2 Sample and Summary Statistics

We collect quarterly mutual fund reports primarily from iFinD, supplementing missing data with RESSET. Our sample period spans from 2009Q1 to 2024Q4. Since 2009, Chinese mutual funds have been required to file these reports using the eXtensible Business Reporting Language (XBRL) format and follow a standardized template. This regulation ensures a uniform reporting structure, enabling us to consistently identify the Operational Analysis section based on its title. Across the full sample period, we successfully identify Operational Analysis sections for 73,052 out of 74,112 fund–quarter observations, accounting for a proportion of 98.57%.

We obtain data on fund characteristics, fund returns, as well as manager information from the China Stock Market and Accounting Research (CSMAR) database. To ensure comparability across funds, we focus on actively managed equity and balanced funds with at least 60% of assets allocated to stocks[7], excluding international, index, bond, and money market funds. We exclude funds with less than two years of history to address incubation bias (Evans 2010; Brown & Wu 2016). Finally, we merge the mutual fund data with textual disclosures using the unique fund code assigned by the CSRC. We exclude funds whose core variables are missing. Our final sample spans from 2009Q1 to 2024Q4, comprising 2,628

[7] Our fund classification is based on the categories provided by the RESSET database's China Galaxy Securities' Fund Rating. We include funds from the following categories: standard equity funds (Class A), mixed funds with stock allocation limits between 60% and 95% (Class A), mixed funds with an upper stock limit of 95% (Class A), mixed funds with an upper stock limit of 80% (Class A), and flexible funds with stock allocation limits between 60% and 100% (Class A).

unique funds and 53,422 fund-quarter observations. In our analyses, all continuous variables are winsorized at the 1st and 99th percentiles and standardized cross-sectionally.

Summary statistics are presented in Table 1. Panel A presents descriptive statistics for mutual fund disclosures. The average rhetorical intensity is 20.806, indicating that most fund managers rely on relatively conventional language rather than extensive rhetorical expressions. The mean *RERatio* is 0.115, which means fund disclosures contain roughly 1.08 rhetorical devices per report on average, given that each report in our sample comprises 9.391 sentences. Regarding other textual characteristics, the average sentiment score is positive, indicating that fund reports generally convey a favorable tone, with the frequency of positive words exceeding that of negative words. The length of this section varies widely, ranging from 30 to 3,000 Chinese characters, with an average of 432 characters, corresponding to approximately 48 characters per sentence.

Panel A also summarized other fund characteristics. On average, quarterly fund flows amount to 1.0%, aligning with the moderate growth observed in China's actively managed equity fund sector documented in Fang et al. (2024). The distribution of Total Net Assets (TNA) spans a wide range, ranging from 6 million CNY to 10 billion CNY, with an average of approximately 463 million CNY. The sample funds have been in continuous operation for roughly 69 months and are affiliated with fund families managing an average of 17 billion CNY in TNA. The average quarterly raw return in the sample is 1.5%, suggesting that Chinese actively managed equity funds, on average, deliver positive investment performance. In general, our descriptive statistics of fund characteristics is highly consistent with previous literature (Chen et al., 2024; Hong et al., 2025).

We next examine how rhetorical intensity relates to other textual characteristics of fund reports. Panel B of Table 1 shows that increased rhetorical expressions is associated with more

positive sentiment, longer texts and lower readability. These associations are statistically significant.

Lastly, we summarize the importance of different rhetorical devices. Specifically, for each quarter, we calculate the average # of funds employing each type of rhetorical device and report the results in Panel C of Table 1. Our sample includes 853 funds per quarter on average. Hyperbole is the most common rhetorical device in the mutual fund reports, employed by 256 funds per quarter, which represents 30.03% of all funds. The second most frequent device is antithesis, used by 30% of funds. Metaphor, a frequently studied rhetorical device in the prior literature, ranks fourth in our sample, with approximately 89 funds employing it each quarter.

[Insert Table 1 Here]

4. Investors' Response to Rhetorical Expressions

4.1 Rhetorical Expressions and Fund Flow

Prior studies provide evidence that fund disclosures significantly influence investors' decision-making and shape subsequent fund flows (Kostovetsky & Warner 2020; Hillert et al. 2024; Sheng et al. 2025). One unexplored channel through which disclosures may affect investor behavior is the use of rhetorical expressions. Such expressions can attract investors' attention and guide their interpretation of a fund's strategy, market conditions, or performance by making complex information more vivid, or emphasizing key messages in a salient way. In this section, we first examine the relationship between rhetorical expressions and subsequent mutual fund flows.

To examine whether rhetoric expressions can affect mutual fund flows, we adopt the following panel regression model:

$$Flow_{i,q+1} = \alpha + \beta \cdot RE_{i,q} + \gamma \cdot Controls_{i,q} + \lambda_i + \mu_q + \varepsilon_{i,q} \quad (2)$$

where $Flow_{i,q+1}$ represents fund flows in quarter q+1 (one quarter ahead), and $RE_{i,q}$ captures the rhetorical intensity of fund i in quarter q. We use fund flows one quarter ahead since mutual

funds typically release their quarterly reports within 15 business days after the quarter ends, thus investors can only respond to narratives in the subsequent quarter. $Controls_{i,q}$ contains a range of fund-level characteristics, including fund size (*TNA*), age (*Age*), quarterly return (*Return*) and family size (*FTNA*). Since previous literature demonstrates that mutual fund flows are sensitive to various textual characteristics (Hwang & Kim 2017; Hillert et al. 2024), we also take these features in account in other columns, including *Sentiment*, *Length* and *Readability*. We include fund fixed effects (λ_i) to absorb time-invariant fund characteristics, thereby identifying the effect of rhetorical expressions based on within-fund variation over time. We also include time fixed effects (μ_q) to account for quarterly macroeconomic conditions or regulatory changes that may affect fund flows across all funds. Standard errors are clustered at the fund level. Variable definitions are provided in the Appendix Table A2.

The results are shown in Table 2. We use the two rhetoric measures in Section 3.1, namely *REScore*, and *RERatio*. We include mutual fund characteristics in columns (1) and (3), and other textual features in columns (2) and (4). We demonstrate that the adoption of rhetorical language is significantly related to future fund flow. Specifically, funds that use more rhetorical expressions in their reports attract higher investor flows in the subsequent quarter, with the coefficient being significant at the 1% level (t = 3.782). The magnitude of the effect is also economically meaningful. Take column (2) as an example, a one standard deviation increase in rhetorical intensity is associated with a 2.5% standard deviation increase in fund flow. For *RERatio*, the statistical and economical magnitudes are similar. The results remain robust when we control for other textual indicators, which suggests that our measure captures incremental variation beyond other textual dimensions.

[Insert Table 2 Here]

4.2 Heterogeneity Analysis of Fund Characteristics

If rhetorical expressions can persuade investors, their impact should vary with the degree to which investors' attention is drawn to narrative disclosures and the persuasive power of rhetorical language. Such sensitivity is unlikely to be uniform across funds or investor groups, as it depends on factors such as a fund's established reputation and investors' level of sophistication. In this section, we explore potential sources of heterogeneity in the effect of rhetorical expressions on fund flows across different fund and investor characteristics.

Following DeHaan et al. (2026), we first examine heterogeneity in the effect of rhetorical expressions across funds of different *Size*. Fund size captures differences in the extent to which managers need to persuade investors through narrative disclosure. Larger funds typically have more established reputations, longer performance histories, and more stable investor bases, which reduce managers' reliance on rhetorical language as a persuasive instrument. For these funds, investors are more likely to make their decisions on observable and persistent characteristics rather than on rhetorical language in disclosures. As a result, investor responses to rhetorical expressions are expected to be weaker among large funds. We implement this analysis by splitting funds at the cross-sectional median of TNA. We expect the relationship between rhetorical expressions and subsequent fund flows to be stronger among smaller funds, for which rhetorical narratives play a more important role in persuading investors.

Previous literature also illustrates that different *Investor Structure* may react heterogeneously to rhetorical expressions. For example, König et al. (2018) show that CEOs' use of metaphors in conference calls can trigger more favourable coverage from journalists, but more unfavourable assessments from analysts, as more sophisticated audiences may interpret rhetorical expressions as distracting noises. In our context, retail investors are more likely to regard rhetorical expressions as promotional or persuasive signals, while institutional investor may perceive such strategic disclosure as noisy or uninformative. Therefore, we expect the relationship between rhetorical expressions and subsequent fund flows to be more

pronounced for funds dominated by retail investors, while weaker for funds dominated by institutional investors. To test this hypothesis, we split funds by investor composition, using a 70% cutoff for high retail investor ownership.

The results are shown in Table 3. Consistent with our hypotheses, we find that the effect of rhetorical expressions on subsequent fund flows is more pronounced among smaller funds. Rhetorical narratives play a more important persuasive role when funds lack established reputations and stable investor bases. We observe similar heterogeneity across investor composition. The relationship between rhetorical expressions and future fund flows is still significant for funds dominated by retail investors but disappears for funds with high institutional ownership. This pattern is consistent with retail investors being more influenced by rhetorical framing, whereas institutional investors exhibit little reaction to such disclosure.

[Insert Table 3 Here]

4.3 Rhetorical Expression and Investor First-time Purchase

If rhetorical expressions can persuade, they are likely to be most effective when investors encounter a fund for the first time. The impact of rhetoric depends on investors' experience and sophistication, with retail investors being particularly sensitive to persuasive language. For these investors, rhetorical expressions play a central role in forming first impressions, which guide their initial investment decisions (Döbeli & Vanini, 2010). In contrast, more experienced or sophisticated investors rely on prior knowledge and accumulated experience, reducing the marginal influence of rhetorical expressions on their behavior.

We have acquired unique and detailed transaction records for 200,000 retail investors on the Ant investment platform to examine the influence of rhetorical expressions on individual investors' decisions. The time interval spans from January 2019 to December 2020. The data set records the timing of each purchase and redemption, the amounts invested or redeemed, and the duration for which investors hold each fund. In addition, it contains a rich set of investor

characteristics, including gender, age, and cash flow classifications[8]. We also construct quarterly fund flows using such data, where validation checks show that the constructed quarterly flows are highly consistent with our baseline fund-level flows, with a correlation of 99%.

We use the following panel regression model, to examine how rhetorical expressions influence retail investor investment decisions, especially their first-time purchase decisions:

$$NetPurc_{i,k,m+1} = \alpha + \beta_1 \cdot RE_{i,m} + \beta_2 \cdot RE_{i,m} \times FirstPurc_{i,k,m} + \gamma \cdot Controls_{i,m} + \lambda_i + \theta_k + \mu_m + \varepsilon_{i,k,m} \quad (3)$$

where k denotes the user and m denotes month. λ_i, θ_k, μ_m are fund, user, month fixed effects respectively. The dependent variable *NetPurc* is monthly amount of investor net purchase, measured as the total amount of investor purchases minus investor redemptions within a month. *FirstPurc* equals 1 if the purchase corresponds to the first time the user has ever purchased the fund, and 0 otherwise[9]. We also control for another two user-level characteristics in this specification, namely the total net purchase across all funds in the previous month (*TNetPurc*) and the total value of the investor's portfolio in the previous month (*TPortSize*). Including these variables allows us to account for investors' recent trading behavior and financial capacity, which is important for isolating the effect of rhetorical expressions on individual investors' decisions. The standard deviation is clustered on user-level.

The results are shown in Table 4. The interaction term between rhetorical intensity and investor first-time purchase dummy is significantly positive, indicating that rhetorical expressions are particularly influential at the first-time purchase stage, when investors are forming their first impressions and are more susceptible to persuasive narratives. Furthermore,

[8] This study was remotely conducted on the Ant Open Research Laboratory in an Ant Group Environment. All data was sampled, desensitized and analyzed on the Ant Open Research Laboratory. The laboratory is a sandbox environment where the authors can only remotely conduct empirical analysis and individual observations are invisible. The main regression variables include basic variables, investment variables, and consumption variables. A detailed documentation can be found on https://www.deor.org.cn/labstore/laboratory.

[9] Although our main analysis sample spans from 2019 to 2020, this variable is not defined based on the first-time purchase within the sample period. Ant Group provides the complete historical record of each user's first-time purchase date for each fund, which allows us to identify whether a purchase in 2019 to 2020 is the user's first-ever investment in that fund, even if the user's first-time fund investment occurred prior to 2019.

the standalone coefficient on rhetorical intensity is significantly negative, demonstrating that rhetorical expressions are associated with subsequent investor outflows among investors who already hold the fund. It suggests that, for these investors, accumulated experience and realized performance reduce the persuasive power of narrative disclosures, and rhetorical language may instead be interpreted as less informative or even distracting. Taken together, these results highlight that the persuasive effect of rhetoric is concentrated at the entry stage, while its influence diminishes, and may even reverse once investors have established longer-term experience with the fund managers.

[Insert Table 4 Here]

5. Addressing Endogeneity Concerns

5.1 Timing Tests of Investor Decisions

A potential concern is whether the observed link between rhetorical expressions and investor flows genuinely reflects responses to disclosed information, rather than being driven by unobserved factors or anticipatory trading. If rhetorical expressions merely capture unobserved fund characteristics or time-varying sentiment, investor flows might respond before the report is formally released. To address this concern, we conduct a timing test that explicitly examines whether investor reactions are concentrated around the disclosure month, rather than occurring in advance.

Our timing test exploits both the regulatory disclosure schedule and monthly account-level investor data of Ant Open Research Laboratory. Since 2004, the CSRC has required mutual funds to complete and disclose their quarterly reports within 15 working days after the end of each quarter[10]. As a result, fund managers typically release their quarterly reports in the

[10] In 2004, CSRC announced the Measures for the Administration of Information Disclosure by Securities Investment Funds No. 4: Content and Format of Quarterly Reports, which standardizes the required timing and content of mutual fund quarterly disclosures. See more on https://fg.amac.org.cn//governmentrules_3854/zcgz_gfxwj/gfxwj_gmjj/gfxwj_gmjj_xxpl/202003/t20200312_24750.html.

first month following the quarter-end, while no new narrative information is released during the quarter-end month. This institutional setting allows us to cleanly separate investor flows before and after disclosure. Empirically, we examine investor responses in the quarter-end month and the subsequent disclosure month, and test whether rhetorical expressions predict fund flows only after the report becomes publicly available.

The results are reported in Table 5. Consistent with the disclosure-based interpretation, we find that rhetorical expressions have a significantly positive effect on investor flows in the report disclosure month, when investors are able to observe the fund managers' narratives. The coefficient of *REScore* and *RERatio* are 1.8% (t = 10.537) and 0.8% (t = 4.953) respectively. In contrast, the corresponding coefficients for the month prior to disclosure are small and statistically insignificant (t = 0.675 for *REScore* and t = 0.341 for *RERatio*). This timing pattern suggests that investors do not react in advance, but instead respond immediately once rhetorical content is released, providing strong evidence that rhetorical expressions influence investor behavior through the information disclosure channel rather than through pre-existing trends or other omitted factors.

[Insert Table 5 Here]

5.2 Controlling for Manager and Family Fixed Effects

It is natural to expect that fund managers change their narrative style following managerial turnover. Prior studies also show that managerial turnover after poor performance can attract investor flows (Kostovetsky & Warner 2015). To account for potential confounding factors associated with both fund flows and managers' narrative choices, we additionally control for fund manager fixed effects. This specification effectively absorbs all time-invariant characteristics of individual funds and managers, as well as quarterly shocks that affect all funds simultaneously. As a result, the identification relies on within-manager, within-fund,

within-quarter variation, isolating the effect of changes in managerial narrative style from unobserved heterogeneity at the fund or manager level.

Mutual fund families often implement standardized templates to constrain the content and format of fund managers' disclosures[11]. Regulatory and industry practices have progressively formalized these templates in recent years[12], and fund families may update or revise templates accordingly. To account for this structured environment, we additionally include fund family fixed effects in our regressions. This specification absorbs family-level shocks and template changes that simultaneously affect multiple funds within the same family in a given quarter. Consequently, the identification comes from within-family, within-fund, within-quarter deviations, allowing us to compare how managers adapt their narrative style under common family-level disclosure templates.

The results are shown in Table 6. In columns (1)-(2) and (5)-(6), we include fund, manager, and quarter fixed effects. In columns (3)-(4) and (7)-(8), we include fund, family, and quarter fixed effects. Across both sets of specifications, our main results remain positive and statistically significant, indicating that the relationship between rhetorical expressions and future fund flows is consistent when controlling for manager-level heterogeneity and family-level disclosure templates, thereby demonstrating the robustness of our results.

[Insert Table 6 Here]

5.3 Evidence from a Quasi-experiment

To further examine how rhetorical expressions influence investors' decisions, we exploit a quasi-experimental setting based on regulatory changes. Stricter disclosure requirements are likely to limit managers' use of rhetorical devices: for instance, metaphors are expected to

[11] In American market, such misbehavior has been detected and charged by SEC. The Northern Lights Fund Trust submitted to shareholders "boilerplate" statements about fund advisers, and paid $50,000 in penalties. See more on https://www.investmentnews.com/mutual-funds/sec-charges-mutual-fund-firm-with-misleading-shareholders-on-advisers/51220. In Chinese market, some mutual funds make mistakes in information disclosure which may be caused by simply pasting the templates in the fund family. See more on https://www.stcn.com/article/detail/1171890.html.

[12] The policy shock we utilize in Section 5.3 may serve as a representative example.

remain closely tied to underlying investment activities, and hyperbolic expressions are constrained to avoid excessive exaggeration. These limitations reduce managers' flexibility in emphasizing favorable narratives or clarifying complex concepts beyond fundamental information, thereby weakening the persuasive power of rhetorical language and its ability to shape investor perceptions and attract flows.

On July 26, 2019, the CSRC issued the *Measures for the Administration of Information Disclosure of Publicly Offered Securities Investment Funds* (MAIDF), which was officially implemented on September 1, 2019. Compared to the 2004 regulation, the MAIDF introduces clearer and more rigorous standards aimed at improving the quality and uniformity of mutual fund disclosures. Under the new framework, fund management companies are mandated to "strengthen the management and control of undisclosed fund information". The CSRC is empowered to "regularly evaluate the quality of information disclosure by fund managers and incorporate it into the fund management company's classified supervision and evaluation index system". In general, these measures significantly increase regulatory scrutiny and promote greater standardization in mutual fund.

To examine the impact of the regulatory change implemented in 2019Q3, we define a dummy variable *Post*, which equals 1 after 2019Q3 and 0 otherwise. To maintain sample consistency, our analysis focuses on the mutual funds that were active as of 2019Q2. Given that the regulation applies universally to all mutual funds in China, we capture its differential effect through the interaction term of rhetorical expressions and the dummy variable for the policy. We use the time window from 2017Q3 (-8) to 2021Q3 (+8) and estimate the following model:

$$Flow_{i,q+1} = \alpha + \beta_1 \cdot RE_{i,q} \times Post_q + \beta_2 \cdot RE_{i,q} + \gamma \cdot Controls_{i,q} + \lambda_i + \mu_q + \varepsilon_{i,q} \quad (4)$$

where *Post* is a dummy variable which equals 1 (0) for the post-event (pre-event) window.

Table 7 reports the regression results. The coefficient on the interaction term between rhetorical expression and the post-regulation indicator is negative and statistically significant, indicating that the association between rhetorical expressions and subsequent fund inflows differs before and after the implementation of the MAIDF. Specifically, funds with higher levels of rhetorical expression exhibit a weaker persuasive power for investors after the policy. This pattern suggests that under a more standardized disclosure framework, rhetorical language plays a less prominent role in shaping investors' capital allocation decisions.

[Insert Table 7 Here]

5.4 Presence and Frequency of Rhetorical Devices

Although our rhetorical measures are derived from LLMs, they may still be affected by potential measurement error due to the contextual and interpretive nature of rhetorical classification. To mitigate the impact of such misclassification on our results, we transform the continuous measures into dummy variables, focusing on whether the presence and frequency of rhetorical devices can affect investor response. This approach focuses on more reliable distinctions which reduce the influence of potential noise in the continuous measures. In addition, we extend the analysis to examine the effects of different categories of rhetorical devices on subsequent fund flows.

In Table 8 Panel A, we first define a dummy variable *# REDevice >0*, which equals to 1 if the report uses at least one rhetorical device. Columns (1) and (2) show that the presence of rhetorical devices significantly increases future fund flows. Then, we additionally define two dummy variables *# REDevice =1* (*# REDevice >1*), which equals to 1 if the report has one (more than one) rhetorical device. The results are shown in columns (3) and (4). The coefficient of *# REDevice >1* is significantly positive while that of *# REDevice =1* is insignificant, suggesting that using multiple rhetorical devices is more effective in attracting investor attention and increasing fund flows.

We next examine how different rhetorical devices contribute to subsequent fund flows. Consistent with the construction of *RERatio*, we measure the dense of each rhetorical device as its frequency scaled by the # of sentences in the report. Table 8 Panel B reports the regression results. We find that the adoption of hyperbole and metaphor is positively and significantly associated with future fund flows, whereas other rhetorical devices show no significant effects. These findings suggest that not all rhetorical expressions are equally effective in influencing investor behavior. Hyperbolic language may attract investors by emphasizing the perceived importance of information (Bochkay et al., 2020) and increasing the salience of fund narratives under limited attention. Meanwhile, metaphors help managers convey abstract market conditions or investment strategies in more concrete and intuitive terms, thereby enhancing the persuasive power of disclosures (Charteris-Black, 2011, pp.15). Overall, these results indicate that rhetorical devices that simplify complexity or heighten emphasis are particularly effective in shaping investor responses.

[Insert Table 8 Here]

6. Further Analyses

6.1 Rhetorical Expression and Fund Performance

Then we explore whether investors benefit from chasing rhetorical expressions. While a substantial literature has studied various predictors of mutual fund returns (e.g. Chen et al., 2004; Kacperczyk et al., 2005, 2008; Hoberg et al., 2018; Kaniel et al., 2023; Gantchev et al., 2024), the predictive power of mutual fund narratives remains debated. For instance, Chu and Kim (2019) find that negative tone in narratives is associated with higher subsequent performance, whereas Hillert et al. (2025) report no significant relationship. Unlike tone, which may reflect managerial expectation or substantive information (Li, 2010), rhetorical

expressions capture a purely stylistic dimension of fund disclosures, potentially yielding a cleaner result.

We estimate fund alphas by first calculating each fund's factor loadings over the prior 24 months and then computing the average difference between the realized monthly return and the return predicted by the factor model (Kaniel et al., 2023). We consider three benchmark models, namely CAPM, Fama-French three-factor model (Fama & French, 1993) and Fama-French-Carhart four-factor model (Carhart, 1997). We regress fund performance on rhetoric expression in the previous quarter using panel regression model:

$$Performance_{i,m+1} = \alpha + \beta \cdot RE_{i,m} + \gamma \cdot Controls_{i,m} + \lambda_i + \mu_m + \varepsilon_{i,m} \quad (5)$$

where dependent variables are performance of fund i in month $m+1$, including raw return (*Ret*), alpha estimated by CAPM (*Alpha CAPM*), Fama-French three-factor model (*Alpha 3-factor*), and Fama-French-Carhart four-factor model (*Alpha 4-factor*). The regressor of interest is *RE*, which denotes rhetorical expression in the previous quarter[13]. *Controls* represent a set of textual characteristics, namely sentiment (*Sentiment*), length (*Length*) and readability (*Readability*), and fund characteristics such as fund size (*TNA*), age (*Age*), quarterly return (*Return*) and fund family size (*FTNA*). We include both fund fixed effects (λ_i) and time fixed effects (μ_m). Standard errors are clustered at the fund level. Variable definitions are provided in the Appendix Table A2.

Table 9 presents the results. In all specifications, we find no significant relationship between usage of rhetorical expressions in fund reports and future fund performance, whether measured by raw returns, or various risk-adjusted alphas. Our results are also consistent with Kostovetsky and Warner (2020), Hillert et al. (2024) and DeHaan et al. (2025), which indicate that mutual fund strategic information disclosure can hardly reflect stock-picking skills. In general, this finding indicates that rhetoric devices in fund disclosures, while potentially

[13] For example, the performance measure from January to March 2019 is matched with the rhetoric intensity in 2018Q4.

effective in attracting investor attention and capital, cannot generate better investment outcomes.

[Insert Table 9 here]

6.2 Rhetorical Expressions and Managerial Risk-taking

Our earlier results show that stronger rhetorical expressions do not translate into better fund performance. This section examines potential implication of managerial usage of rhetorics for their risk taking. Rhetorical language in fund disclosures may steer managers toward more ambitious or unconventional strategies, exposing investors to higher risks. The narrative factor represents a novel channel beyond previously documented determinants of risk-taking, including past performance (Chevalier & Ellison, 1997; Huang et al., 2011), managerial characteristics (Kempf et al., 2009; Hu et al., 2011; Ma & Tang, 2019). To explore this potential channel, we examine whether funds that employ more intensive rhetorical expressions are associated with higher subsequent risk-taking.

We adopt three risk measures following prior research (Ang et al. 2006; Ma & Tang 2019; Sheng et al. 2025), including total risk (*Total Risk*), systematic risk (*Sys Risk*), and idiosyncratic risk (*Idio Risk*), as independent variables. We then estimate the following panel regression:

$$Risk_{i,m+1} = \alpha + \beta \cdot RE_{i,m} + \gamma \cdot Controls_{i,m} + \lambda_i + \mu_m + \varepsilon_{i,m} \quad (6)$$

where the dependent variables are risk-taking proxies. The variable of interest is *RE*, which is rhetorical expression in the previous quarter. *Controls* represent a set of fund fundamentals. The standard errors are adjusted for Newey-West standard errors with 4 lags.

The results are shown in Table 10. We find that more rhetorical expressions are associated with increased risk-taking behavior. In column (1), the coefficient of *REScore* is 0.060 and statistically significant at the 1% level (t = 7.269), indicating that a one standard deviation increase in rhetorical expression corresponds to a 6.0% standard deviation increase in total risk. Column (2) shows a similar positive relationship with systematic risk, with a coefficient of

0.030 (t = 4.172). In column (3), the correlation is more salient for idiosyncratic risk, with a coefficient of 0.057 (t = 7.101), suggesting that a one standard deviation increase in narrative modification corresponds to an 5.7% standard deviation increase in idiosyncratic risk. All the results are similar for *RERatio*. In conclusion, funds with more rhetorical expressions are associated with more aggressive risk-taking behavior.

[Insert Table 10 here]

We further examine how fund managers take risks. Mutual funds typically disclose their investment objectives in prospectuses, but managers do not always adhere strictly to these guidelines in practice (Wermers, 2012; Cao et al., 2017). Such deviations, commonly referred to as style drift, reflect managerial discretion in adjusting portfolio positions relative to benchmark allocations. Our previous empirical results indicate that stronger rhetorical expressions are associated with greater subsequent risk-taking, suggesting that rhetoric may coincide with shifts in managerial behavior beyond what is disclosed. Accordingly, we expect that fund managers who employ more rhetorical languages are more likely to deviate from their self-stated investment styles.

We use the deviation of fund return to benchmark index return to proxy for style drift. Specifically, we define the style drift measure as the absolute difference between the fund's return and the benchmark index return in each month, denoted as *Abs(ReturnBench)*. The benchmark index could be an index that follows the same style as the product's portfolio (Cremers et al., 2022). This measure captures the magnitude of divergence between actual performance and the expected style benchmark return, with larger absolute deviations indicating greater style drift. If a fund consistently adheres to its stated investment style, its return should closely track the return of the corresponding style benchmark, resulting in a small return difference. In contrast, when a fund deviates from its intended style, its realized return

may diverge from the benchmark regardless of whether the deviation leads to outperformance or underperformance.

We use the similar panel regression model in formula (6), with dependent variable replaced by *Abs(ReturnBench)*. The results are shown in Table 11[14]. We find a significantly positive relationship between rhetorical intensity and future fund style drift. A one standard deviation increase in rhetorical intensity leads to 2.7% (t = 4.721) standard deviation increase in in style drift. This finding highlights a tangible cost to investors: when investors chase rhetorical expressions, they bear greater deviations from the fund's stated investment style, resulting in additional unintended risk exposure.

[Insert Table 11 here]

6.3 Potential Cost of Rhetorical Expression for Fund Managers

Our findings show a positive association between rhetorical expressions and subsequent fund inflows, which raises the question of why managers do not continuously use rhetorical devices to attract inflow. Rhetorical language makes narratives more salient, drawing investor attention and amplifying the perceived importance of disclosed information. At the same time, it heightens investors' sensitivity to subsequent outcomes: when realized returns fail to validate rhetorically framed expectations, investors update their beliefs and reassess the credibility of managerial communication (Hillert et al., 2025), potentially reducing the effectiveness of rhetoric in driving fund inflows.

Building on this view, we argue that the impact of rhetorical expressions on subsequent flow is conditional on subsequent performance. When strategically rhetorical expressions are followed by poor performance, investors are likely to interpret such language as promotional talk with limited informational value about managerial ability, leading to investor redemptions. In contrast, when strong subsequent performance accompanies rhetorical disclosures, investors

[14] The observation is slightly smaller in this regression compared to results in Table 9 and Table 10, as some of the funds do not state their benchmark indexes.

are more inclined to view them as credible signals of managerial effort and transparency, which reinforces confidence and attracts additional inflows.

Referring to previous literature (Huang et al. 2007; Franzoni & Schmalz 2017; Li et al. 2017; Huang et al. 2022), we adopt a flow-performance sensitivity framework and introduce an interaction term between rhetorical intensity and fund return in the subsequent quarter. Specifically, we estimate the following panel regression:

$$Flow_{i,q+1} = \alpha + \beta_1 RE_{i,q} \times Return_{i,q+1} + \beta_2 RE_{i,q} + \gamma \cdot Controls_{i,q} + \lambda_i + \mu_q + \varepsilon_{i,q} \quad (7)$$

where the dependent variable includes both mutual fund flow. We also construct two dummy variables, *High* and *Low*, indicating whether a fund's return falls into the top or bottom quintile, respectively, to distinguish the effect of rhetorical expressions for outperforming versus underperforming funds.

The results are shown in Table 12. The interaction term between rhetorical intensity and subsequent returns is significantly positive, indicating that rhetorical expressions attract greater inflows when followed by stronger performance. More importantly, the interaction between rhetorical intensity and *High* is significantly positive, whereas the corresponding interaction with *Low* is significantly negative. These findings indicate that the effect of rhetorical language on investor behavior depends critically on following performance: they reward rhetorically rich disclosures when managers deliver outstanding performance, but penalize such rhetoric when performance is inferior. Overall, rhetorical language amplifies investors' responses to future outcomes, benefiting managers when results are favorable but imposing greater costs when outcomes are poor.

[Insert Table 12 Here]

7. Conclusion

This paper examines whether rhetorical expressions in mutual fund reports influence investor behavior in the Chinese mutual fund market. Leveraging advanced LLMs to quantify rhetorical intensity, we provide evidence that more rhetorically expressive disclosures are associated with higher subsequent fund flows. This effect is particularly pronounced among smaller funds and those dominated by retail investors. Using account-level transaction data, we further show that this effect is concentrated in investors' first-time purchase decisions, indicating that rhetorical language is particularly influential when investors form first impressions. Our findings are robust across a range of robustness tests, including disclosure timing tests, manager- and family-level fixed effects, a regulatory policy shock, and alternative dummy measures of rhetorical usage. Importantly, while rhetorical language attracts investor capital, it does not predict superior future performance and is instead associated with higher subsequent risk-taking and more style drift. Moreover, investors condition their responses on realized outcomes, rewarding rhetorically expressive managers following strong performance but penalizing them when subsequent performance is poor.

Our study contributes to the literature on strategic disclosure and persuasions in financial markets. Whereas prior research primarily focuses on textual attributes such as tone, sentiment, or readability, we highlight rhetorical expressions as a distinct and economically meaningful dimension of financial communication. Rhetorical language illustrates a fundamental tension in financial disclosure: while narratives can enhance communication and attract attention, they may also function as persuasive instruments that do not convey incremental information about managerial skill. In this sense, rhetoric occupies a grey area between transparency and persuasion, with implications for both investor learning and agency conflicts.

Our findings have important implications for investors, fund managers, and regulators. For investors, rhetorical expressions in fund reports should be interpreted with caution, as they

may serve primarily as attention-grabbing devices rather than reliable signals of future performance. For fund managers, the strategic use of rhetorical language entails potential costs: when not supported by subsequent performance, such disclosures may be perceived as impression management, leading to reputational damage and capital outflows. For regulators, our results suggest that while rhetorical language can facilitate communication, it may also amplify the risk of misleading disclosures, underscoring the importance of standardized reporting frameworks and effective supervision to protect investor interest.

Reference

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

Abis, S., & Lines, A. (2024). Broken promises, competition, and capital allocation in the mutual fund industry. *Journal of Financial Economics*, *162*, 103948.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, *61*(1), 259-299.

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance*, *59*(3), 1259-1294.

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

Back, K. W. (1989). Rhetoric as communication and performance. *Communication Research*, *16*(1), 130-148.

Bini, P., Cong, L. W., Huang, X., & Jin, L. J. (2025). Behavioral Economics of AI: LLM Biases and Corrections. *Available at SSRN:* 5213130.

Bochkay, K., Hales, J., & Chava, S. (2020). Hyperbole or reality? Investor response to extreme language in earnings conference calls. *The Accounting Review*, *95*(2), 31-60.

Bochkay, K., Brown, S. V., Leone, A. J., & Tucker, J. W. (2023). Textual analysis in accounting: What's next?. *Contemporary Accounting Research*, *40*(2), 765-805.

Breitung, C., & Müller, S. (2025). Global business networks. *Journal of Financial Economics*, *166*, 104007.

Brown, D. P., & Wu, Y. (2016). Mutual fund flows and cross-fund learning within families. *The Journal of Finance*, *71*(1), 383-424.

Cao, C., Iliev, P., & Velthuis, R. (2017). Style drift: Evidence from small-cap mutual funds. *Journal of Banking & Finance*, *78*, 42-57.

Cao, S., Wang, C. C., & Xiang, Y. (2025). When LLMs go abroad: Foreign bias in AI financial predictions. *Available at SSRN:* 5440116.

Cao, S., Yang, B., & Zhang, A. L. (2024). Beyond the Lines: Uncovering Private Information from Fund Managers' Disclosures. *Available at SSRN:* 3713966.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57-82.

Charteris-Black, J. (2011). *Politicians and rhetoric: The persuasive power of metaphor*. Springer.

Chen, H., Qu, Y., Shen, T., Wang, Q., & Xu, D. X. (2022). The geography of information acquisition. *Journal of Financial and Quantitative Analysis*, *57*(6), 2251-2285.

Chen, J., Hong, H., Huang, M., & Kubik, J. D. (2004). Does fund size erode mutual fund performance? The role of liquidity and organization. *American Economic Review*, *94*(5), 1276-1302.

Chen, J., Tang, G., Zhou, G., & Zhu, W. (2025). ChatGPT and DeepSeek: Can they predict the stock market and macroeconomy? *arXiv preprint arXiv:*2502.10008.

Chen, Y., Kelly, B. T., & Xiu, D. (2022). Expected returns and large language models. *Available at SSRN: 4416687*.

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

Chu, Y., & Kim, H. H. (2019). Pessimistic fund managers. *Available at SSRN:* 2909208.

Clifford, C. P., Fulkerson, J. A., Jame, R., & Jordan, B. D. (2021). Salience and mutual fund investor demand for idiosyncratic volatility. *Management Science*, *67*(8), 5234-5254.

Comprix, J., Lopatta, K., & Tideman, S. A. (2022). The role of gender in the aggressive questioning of CEOs during earnings conference calls. *The Accounting Review*, *97*(7), 79-107.

Cong, L. W., Guo, Y., Zhao, X., & Zhou, W. (2024). Writing Quality and Soft Information in the GenAI Age: Evidence from Online Credit Markets. *Available at SSRN*: *4959535*.

Cremers, K. M., Fulkerson, J. A., & Riley, T. B. (2022). Benchmark discrepancies and mutual fund performance evaluation. *Journal of Financial and Quantitative Analysis*, *57*(2), 543-571.

Cronqvist, H. (2006). Advertising and portfolio choice. *Available at SSRN 920693*.

De Kok, T. (2025). ChatGPT for textual analysis? How to use generative LLMs in accounting research. *Management Science*, *71*(9), 7223-8095.

DeHaan, E., Huang, A. H., Kannan, S., & Qiu, L. (2026). Social Media Livestreaming: Investor Information or Persuasion?. *Journal of Accounting and Economics*, 101861.

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

Delbaere, M., McQuarrie, E. F., & Phillips, B. J. (2011). Personification in advertising. *Journal of Advertising*, *40*(1), 121-130.

Döbeli, B., & Vanini, P. (2010). Stated and revealed investment decisions concerning retail structured products. *Journal of Banking & Finance*, *34*(6), 1400-1411.

Du, Z., Huang, A. G., Wermers, R., & Wu, W. (2022). Language and domain specificity: A Chinese financial sentiment dictionary. *Review of Finance*, *26*(3), 673-719.

Eisfeldt, A. L., & Schubert, G. (2024). Generative AI and Finance. *Annual Review of Financial Economics*, *17*.

Evans, R. B. (2010). Mutual fund incubation. *The Journal of Finance*, *65*(4), 1581-1611.

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

Fang, Y., Luo, D., & Yao, Z. (2024). Belief dispersion in the Chinese stock market and fund flows. *Journal of Banking & Finance*, *166*, 107252.

Ferretti, R., Pancotto, F., & Rubaltelli, E. (2017). Persuasion in financial advertising: Behavioral or rational?. *Journal of Behavioral and Experimental Economics*, *71*, 26-30.

Frankel, R., Jennings, J., & Lee, J. (2022). Disclosure sentiment: Machine learning vs. dictionary methods. *Management Science*, *68*(7), 5514-5532.

Franzoni, F., & Schmalz, M. C. (2017). Fund flows and market states. *The Review of Financial Studies*, *30*(8), 2621-2673.

Gallaher, S. T., Kaniel, R., & Starks, L. T. (2015). Advertising and mutual funds: From families to individual funds. *Available at SSRN:* 2554403.

Gao, Z., Xiong, W., & Yuan, J. (2024). Structured Beliefs and Fund Performance: An LLM-Based Approach. *Available at SSRN*.

Gil-Bazo, J., & Imbet, J. F. (2025). Tweeting for money: Social media and mutual fund flows. *Management Science*.

Herrick, J. A. (2020). *The history and theory of rhetoric: An introduction*. Routledge.

Hillert, A., Niessen-Ruenzi, A., & Ruenzi, S. (2025). Mutual fund shareholder letters: flows, performance, and managerial behavior. *Management Science*, *71*(5), 4453-4473.

Hoberg, G., Kumar, N., & Prabhala, N. (2018). Mutual fund competition, managerial skill, and alpha persistence. *The Review of Financial Studies*, *31*(5), 1896-1929.

Hoefer, R. L., & Green Jr, S. E. (2016). A rhetorical model of institutional decision making: The role of rhetoric in the formation and change of legitimacy judgments. *Academy of Management Review*, *41*(1), 130-150.

Hong, C. Y., Lu, X., & Pan, J. (2025). FinTech platforms and mutual fund distribution. *Management Science*, *71*(1), 488-517.

Hu, P., Kale, J. R., Pagani, M., & Subramanian, A. (2011). Fund flows, performance, managerial career concerns, and risk taking. *Management Science*, *57*(4), 628-646.

Huang, A. H., Wang, H., & Yang, Y. (2023). FinBERT: A large language model for extracting information from financial text. *Contemporary Accounting Research*, *40*(2), 806-841.

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

Huang, J., Sialm, C., & Zhang, H. (2011). Risk shifting and mutual fund performance. *The Review of Financial Studies*, *24*(8), 2575-2616.

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

Huang, J., Wei, K. D., & Yan, H. (2022). Investor learning and mutual fund flows. *Financial Management*, *51*(3), 739-765.

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

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

Jiang, F., Lee, J., Martin, X., & Zhou, G. (2019). Manager sentiment and stock returns. *Journal of Financial Economics*, *132*(1), 126-149.

Kaniel, R., Lin, Z., Pelger, M., & Van Nieuwerburgh, S. (2023). Machine-learning the skill of mutual fund managers. *Journal of Financial Economics*, *150*(1), 94-138.

Kacperczyk, M., Sialm, C., & Zheng, L. (2005). On the industry concentration of actively managed equity mutual funds. *The Journal of Finance*, *60*(4), 1983-2011.

Kacperczyk, M., Sialm, C., & Zheng, L. (2008). Unobserved actions of mutual funds. *The Review of Financial Studies*, *21*(6), 2379-2416.

Kempf, A., Ruenzi, S., & Thiele, T. (2009). Employment risk, compensation incentives, and managerial risk taking: Evidence from the mutual fund industry. *Journal of Financial Economics*, *92*(1), 92-108.

Kim, A., Muhn, M., & Nikolaev, V. (2023). Bloated disclosures: can ChatGPT help investors process information?. *arXiv preprint arXiv:2306.10224*.

König, A., Mammen, J., Luger, J., Fehn, A., & Enders, A. (2018). Silver bullet or ricochet? CEOs' use of metaphorical communication and infomediaries' evaluations. *Academy of Management Journal*, *61*(4), 1196-1230.

Kostovetsky, L., & Warner, J. B. (2015). You're fired! New evidence on portfolio manager turnover and performance. *Journal of Financial and Quantitative Analysis*, *50*(4), 729-755.

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

Li, C. W., Tiwari, A., & Tong, L. (2017). Investment decisions under ambiguity: Evidence from mutual fund investor behavior. *Management Science*, *63*(8), 2509-2528.

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

Li, K., Mai, F., Shen, R., Yang, C., & Zhang, T. (2026). Dissecting corporate culture using generative AI. *The Review of Financial Studies*, *39*(1), 253-296.

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *The Journal of Finance*, *69*(4), 1643-1671.

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, *54*(4), 1187-1230.

Ma, L., & Tang, Y. (2019). Portfolio manager ownership and mutual fund risk taking. *Management Science*, *65*(12), 5518-5534.

McCloskey, D. N. (1983). The rhetoric of economics. *Journal of Economic Literature*, *21*(2), 481-517.

McCloskey, D. N. (1998). *The rhetoric of economics*. Univ of Wisconsin Press.

Michael-Luna, S., & Canagarajah, S. (2015). Multilingual academic literacies. *Journal of Applied Linguistics*, *4*(1), 55-77.

Mullainathan, S., & Shleifer, A. (2005). Persuasion in finance. *NBER Working Paper*: 11838.

Niu, Y., Wu, J., Jiang, S., & Jiang, Z. (2025). The bullwhip effect in servitized manufacturers. *Management Science*, *71*(1), 1-20.

Patel, S., & Sarkissian, S. (2017). To group or not to group? Evidence from mutual fund databases. *Journal of Financial and Quantitative Analysis*, *52*(5), 1989-2021.

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

Sheng, J., Xu, N., & Zheng, L. (2025). Do Mutual Funds Walk the Talk? Evidence from Fund Risk Disclosure. *The Review of Financial Studies, forthcoming*.

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

Tan, L., Wu, H., & Zhang, X. (2023). Large language models and return prediction in China. *Available at SSRN:* 4712248.

Taylor, S. E., & Thompson, S. C. (1982). Stalking the elusive" vividness" effect. *Psychological Review*, *89*(2), 155.

Wermers, R. (2012). Matter of Style: The Causes and Consequences of Style Drift in Institutional Portfolios. *Available at SSRN:* 2024259.

Yang, A., Li, A., Yang, B., Zhang, B., Hui, B., Zheng, B., ... & Qiu, Z. (2025). Qwen3 technical report. *arXiv preprint arXiv:*2505.09388.

Young, J. J. (2003). Constructing, persuading and silencing: the rhetoric of accounting standards. *Accounting, Organizations and Society*, *28*(6), 621-638.

Table A1 Prompt

This table presents the detailed instructions given to Qwen-3-Max to capture the rhetorical intensity in mutual fund reports. We provide both the Chinese and English version of our prompt. We start with the basic information of our input texts and main task the LLM should finish. Then, we provide definition and examples of 8 different rhetorical devices. We also impose restrictions on the rating criteria. Finally, we ask LLM to standardize its response.

Vairable	English Prompt	Chinese Prompt
REScore *RERatio*	Forget all previous instructions. You are now a fund investor. Below is a fund's review of macroeconomic conditions, capital markets, and fund operations over the past quarter. From the perspective of how an investor would feel when reading this text, please evaluate the overall intensity of rhetorical expression in the text and report the frequency of each type of rhetorical device. You may refer to the following types of rhetorical devices and their typical examples when making your assessment: (1) Metaphor: A rhetorical device that explains an object or idea by referring to another object or idea with shared characteristics. For example: • "For this fund, the adopted strategy is based on the long-term certainty of equity assets. Like navigating at sea at night, a distant lighthouse guides the direction toward shore." • "Mr. Market behaves like a pendulum, constantly swinging between overheating and overcooling, spending only brief moments at a rational equilibrium." (2) Personification: Attributing human actions or characteristics to inanimate objects. For example: • "With multiple strong economic support policies introduced domestically, the market instantly shook off its drowsiness and sprang to life." • "From a medium- to long-term perspective, the government is in fact strengthening its body and honing its internal capabilities while safeguarding the bottom line of security and stability." (3) Hyperbole: Deliberately exaggerating the characteristics of people, events, or objects to create a strong impression, causing a clear departure from literal facts. You should judge, based on context, whether the fund manager is intentionally exaggerating or merely stating facts. For example: • "The subsequent bursting of the bubble caused these so-called beautiful dreams to vanish into thin air, with grand narratives collapsing entirely and speculators paying a painful price."	忘记之前给你的提示。你现在是一名基金投资者。下面是一段基金对于过去一个季度宏观经济、资本市场、基金运作情况的回顾。请从投资者阅读这段文本时的感受出发，对这段文本整体使用修辞手法(rhetorical device)的强度进行打分，并且给出每种修辞手法出现频率的统计。 你可以参考下面不同类型的修辞与典型示例进行评分： （1） 比喻(Metaphor)：是在描写事物或说明道理时，将一件事物或道理指成另一件事物或道理的修辞手法，这两件事物或道理之中具有一些共同点，比如： • "对于本基金而言，采取的应对策略是基于股票资产的长期确定性进行投资布局，在海上夜行时，远处的灯塔指引着前进靠岸的方向" • "市场先生就是这样，像个钟摆，总游离于过热和过冷之间，在理性位置停留的时间总是很短" （2） 拟人(Personification)：将无生命的物品赋予具体的行为，使其人性化，比如： • "国内出台多项经济托底政策，力度较大，市场睡意全无，瞬间雄起" • "政府基于中长期的考虑，实际上是在守住安全和稳定的底线下，强身健体，苦练内功" （3）夸张(Hyperbole)：将客观之人、事或物的特点，透过主观情意，故意用夸大铺张地渲染与铺饰描述的手法，使它与真正的事实相差很远，以加深读者的印象。此处你需要结合背景判断基金经理是否故意夸大还是陈述事实。比如： • "随后的泡沫破灭，让这些所谓美好的梦境灰飞烟灭，宏大叙事也荡然无存，投机的人付出了惨痛的教训"

- "In the long run, the application space far exceeds computing power; truly disruptive applications could generate tenfold or even hundredfold returns."

(4) Antithesis: Highlighting the main subject by contrasting it with opposing, differing, or contradictory elements. For example:

- "Pessimists win the present, optimists win the future."
- "The road is tortuous, but the future is bright."

(5) Parallelism: Using three or more words, phrases, or sentences with similar structure and length, and related or identical meanings, to enhance rhetorical force. For example:

- "In the parallel world of the stock market, participants view it from different perspectives depending on their circumstances: it has a surreal side, an irrational side, an uncertain side, a dangerous side, as well as a greedy side and a fearful side."
- "We firmly believe that China's long-term trend of reform, opening-up, and innovative transformation will not change; we firmly believe that a group of outstanding firms will seize this historical opportunity and achieve long-term profit growth; we firmly believe that excessively pessimistic market sentiment will eventually normalize, and the A-share market will ultimately assign reasonable valuations to these future champions, delivering substantial returns to investors."

(6) Quotation: Citing famous individuals, events, literary works, classical allusions, idioms, proverbs, maxims, or poems to support the author's viewpoint and strengthen the argument. For example:

- "In the second quarter, the market moved beyond the 'mountains and rivers with no road ahead' of early Q1 and entered a phase of 'another village after the dark willows and bright flowers.'"
- "'Once one recognizes the face of the spring wind, a thousand blossoms reveal that it is spring.' We hope that, as our investment framework continues to improve and mature, we can deliver more stable, superior, and sustainable returns to investors."

(7) Hypophora: Asking a question with an obvious answer to emphasize a point and enhance persuasiveness. For example:

- "For globally priced commodities, the missed opportunity traces back to the second half of 2022, when I began to wonder: if the world's major currencies are, to some extent, artificial, then to what extent can commodity prices be genuine?"

- "长期看，应用的空间远大于算力，若存在颠覆式的应用，则会创造十倍甚至百倍的回报"

（4）反衬(Antithesis)：用相反、有差别或对立的事物做衬体，从反面陪衬主体事物。反衬中，被衬托的主要事物和陪衬的次要事物相反或相对，或是具有不同的方向，比如：

- "悲观者赢得当下，乐观者赢得未来"
- "道路是曲折的，前途是光明的"

（5）排比(Parallelism)：利用三句或以上结构和长度均类似、意义相关或相同的词，短语或句子排列起来，达到一种加强语势的效果，比如：

- "在股市这个平行世界中，参与者在不同处境当中看待它有不同的角度和不同的看法，它有超现实的一面、非理性的一面、不确定的一面、危险的一面，也有贪婪的一面和恐惧的一面。"
- "我们坚信中国经济改革开放、转型创新的大趋势不会改变；我们坚信终究会有一批优秀的企业能够把握住这个历史机遇并实现利润的长期增长；我们坚信过冷的市场情绪终将恢复正常，A 股市场终将会给这些未来的优秀企业以合理的估值，从而给投资者带来 丰厚的回报"

（6）引用(Quotation)：援用名人的话，或名人的事、物、诗文、典故、寓言、成语、俗语、格言、谚语等，来支持作者的立场，以达到证明和加强自己所说的理论，让文章的内容更为充实，比如：

- "二季度，市场摆脱了一季度前后的山重水复疑无路，走出了一段独立的柳暗花明又一村的好光景"
- "等闲识得东风面，万紫千红总是春。希望随着我们投资框架 的不断完善和丰富，为投资人带来更加稳健优异且持续的投资回报"

（7）设问(Hypophora)：用问语提示下文，突出语意的重点，以吸引读者注意，以"自问自答"的方式在文章中呈现，比如：

- "对于全球定价的商品，错过的原因则要回溯至 2022 年下半年，当时我有些怀疑： 如果全球最主要的货币在一

- "However, as the disclosure season for annual and quarterly reports approached, market characteristics reversed. With many emerging growth stocks missing earnings expectations, investors began to question whether future prospects had already been excessively priced in."

(8) Rhetorical Question: Posing a question and then answering it within the text to emphasize key points and attract the reader's attention. For example:

- "What is risk? Setting aside strict semantic or economic definitions, we believe the core of risk lies in the possibility of actual loss."
- "What is the primary contradiction in this market? The A-share market has transitioned from a complex game involving institutions and numerous individual investors to one dominated by institutional investors, implying that it has become increasingly difficult for institutions to profit from less sophisticated retail investors. As a result, value reversion driven by fundamentals is strengthening while technical arbitrage opportunities are weakening."

Additional requirements:

1. Your score must range from 0 to 100, where a higher score indicates greater rhetorical intensity. A score of 0 means no rhetorical devices are used, while 100 indicates that the text is entirely rhetorical. You should comprehensively consider the frequency, distribution, and salience of rhetorical devices throughout the text, rather than assigning a high score based on isolated instances.
2. The scoring guidelines are as follows:

- 0–20: Only very occasional, mildly decorative wording
- 20–40: Everyday, restrained writing with one or two mildly rhetorical sentences
- 40–70: Frequent rhetoric, but still primarily focused on standard informational disclosure
- 70–100: Extremely dense rhetoric with extensive use of multiple rhetorical devices

3. For each text, you must provide a numerical score and a count of each rhetorical device. Your response must be in JSON format only, without any additional text. For example:

{
"REScore": <overall rhetorical intensity score>,
"Metaphor": <count>,

定程度上是假的，那么商品价格在多大程度上能是真的？"

- "但进入年报和季报披露密集区后，市场特征出现逆转，大量新兴成长股的业绩不达预期，投资者开始怀疑是否已经为未来过度的透支？"

（8）反问(Rhetorical Question)：无疑而问、明知故问，把答案寓于问话中并加以强调，对方从问话中思考出答案的句式，并增强语气和说服力，比如：

- "风险是什么？摒弃语义或经济学界的严格定义，我们认为风险的核心是实际亏损的可能性"
- "这个市场的主要矛盾是什么？A 股市场已经从机构和众多个人投资者组成的复杂博弈的市场转向以机构投资者为主的市场，这意味着机构投资者实际上已经很难从技术和专业欠缺的个人投资者中赚得收益，A 股市场的价值回归规律在减弱，基本面驱动的价值差异化规律在增强"

我还有如下要求:

（1）你的评分需要控制在 0 到 100 之间，分数越高则文本使用的修辞手法的强度越高，0 表示没有修辞的使用，100 表示全部都是修辞的内容。请综合考虑修辞手法在文本种的频率、分布范围和显眼程度，对整体修辞强度进行评分，而不是出现个别修辞就给出较高的评分。

（2）我有如下的评分标准：

只有极个别词语稍显修饰性文本的，评分在 0 到 20 分

日常、较克制、偶尔出现一两句修饰性文本的，评分在 20 到 40 分

修辞较多但仍以正常信息披露为主的文本，评分在 40 到 70 分

修辞极为密集、大量出现多种修辞手法的文本，评分在 70 到 100 分

（3） 每段文本你需要给出数字形式的评分，和每种修辞出现频率的统计，回复需要以 JSON 格式输出，不要回复其他内容，比如：

"Personification": <count>, "Hyperbole": <count>, "Antithesis": <count>, "Parallelism": <count>, "Quotation": <count>, "Hypophora": <count>, "RhetoricalQuestion": <count> } Below is the fund's operational analysis text.	{“ RhetoricIntensity”：整体修辞强度的评分，“比喻”：比喻出现的次数，“拟人”：拟人出现的次数，“夸张”：夸张出现的次数，“反衬”：反衬出现的次数，“排比”：排比出现的次数，“引用”：引用出现的次数，“设问”：设问出现的次数，“反问”：反问出现的次数} 下面是基金的运作回顾正文。

Table A2 Variable Definition

Variable Name	Definition
Panel A: Textual Characteristics	
REScore	Rhetorical intensity scores generated by qwen-3-max using prompt in Table A1.
RERatio	$RERatio_{i,q} = \frac{\sum_{r=1}^{8} \#\ of\ Rhetorical\ Device_{i,r,q}}{\#\ of\ sentences_{i,q}}$ The total number of all eight rhetorical devices in a fund report divided by the number of sentences in the report.
#REDevice	Number of rhetorical devices.
Sentiment	$Sentiment_{i,q} = \ln(\frac{pos_{i,q}+1}{neg_{i,q}+1})$ where *pos* and *neg* are positive and negative word numbers in mutual fund disclosures identified by the financial sentiment dictionary proposed by Jiang *et al.* (2019).
Length	The logarithm of number of characters in the fund report, plus one.
Readability	The logarithm of the sentence lengths in the fund report, plus one.
Panel B: Fund Characteristics	
TNA	The logarithm of Total Net Assets.
Age	The logarithm of the number of months since the fund's inception.
Return	Accumulative raw return of the fund in the quarter.
Flow	$Flow_{i,q} = \frac{TNA_{i,q} - TNA_{i,q-1}}{TNA_{i,q-1}} - Return_{i,q}$ where *TNA* is total net assets，*Return* is quarterly raw return.
FTNA	The logarithm of the total net assets of all equity funds within the fund family.
Panel C: Individual Investor Characteristics	
NetPurc	Monthly net purchase of a fund by an investor. Net purchase is the total purchase value minus total redemption value.
TNetPurc	Total monthly net purchase across all funds by an investor.
TPortSize	Total value of portfolio across all funds by an investor.
Panel D: Manager Characteristics	
Gender	Equals 1 if the manager is male, and equals 0 if the manager is female.
PhD	Equals 1 if the manager has PhD degree, and 0 otherwise.
Analyst	Equals 1 if the manager used to work as an analyst, and 0 otherwise.
Panel E: Performance and Risk-taking	
Ret	Monthly raw return before fee of the fund.
Alpha CAPM	Alpha generated from CAPM. Beta is estimated using monthly fund returns in the previous 24 months.
Alpha 3-factor	Alpha generated from the Fama and French (1993) three-factor model. Beta is estimated using monthly fund returns in the previous 24 months.
Alpha 4-factor	Alpha generated from the Fama and French (1993) three-factor and momentum model. Beta is estimated using monthly fund returns in the previous 24 months.
Abs(RetBench)	The absolute value of monthly raw return minus monthly return of benchmark index.
Total Risk	Standard deviation of daily raw return in a month.
Sys Risk	Estimated coefficient of excess return on the MKT of Fama and French (1993) three-factor model. Betas are estimated based on the daily fund returns and factor return.
Idio Risk	Standard deviation of residuals of Fama and French (1993) three-factor model. The estimation is based on the month level.

Table 1 Summary Statistics

This table reports summary statistics of report textual and fund characteristics (Panel A), correlation between rhetorical intensity and other textual characteristics (Panel B), and average number of funds using rhetorical devices per quarter (Panel C). For each variable, the table displays the number of available observations (Obs), the mean (Mean), the standard deviation (SD), the 1st (P1), 25th (P25), 50th (P50), 75th (P75), and 99th (P99) percentiles. The sample includes all actively managed equity funds from 2009Q1 to 2024Q4. All variables are defined in detail in Appendix Table A2.

Panel A: Textual and Fund Characteristics

VarName	Obs	Mean	SD	P1	P25	Median	P75	P99
REScore	53784	20.806	10.096	10.000	15.000	15.000	25.000	55.000
RERatio	53784	0.115	0.166	0.000	0.000	0.000	0.188	0.750
Sentiment	53784	0.712	0.776	-1.099	0.194	0.693	1.204	2.565
Length	53784	6.072	0.655	4.454	5.642	6.089	6.515	7.656
Readability	53784	-3.897	0.305	-4.677	-4.092	-3.892	-3.699	-3.183
Flow	53784	0.010	0.462	-0.517	-0.082	-0.030	-0.000	1.582
TNA	53784	19.950	1.630	15.632	18.839	20.075	21.156	23.055
Age	53784	4.242	0.607	3.178	3.738	4.234	4.718	5.429
Return	53784	0.015	0.113	-0.225	-0.056	0.002	0.075	0.352
FTNA	53784	23.559	1.510	18.435	22.783	23.770	24.638	25.901

Panel B: Correlation of Textual Characteristics

	REScore	*RERatio*	*Sentiment*	*Length*	*Readability*
REScore	1				
RERatio	0.737***	1			
Sentiment	0.147***	0.114***	1		
Length	0.413***	0.134***	0.061***	1	
Readability	-0.196***	-0.276***	-0.024***	-0.490***	1

Panel C: Average Number of Rhetorical Devices

VarName	# of Funds Using Rhetorical Devices	Percentage of Funds Using Rhetorical Devices
Hyperbole	256.381	30.031%
Antithesis	256.095	29.998%
Parallelism	129.968	15.224%
Metaphor	89.397	10.472%
Personification	36.381	4.261%
Quotation	25.762	3.018%
Hypophora	9.111	1.067%
RhetoricalQuestion	0.508	0.060%

Table 2 Rhetoric Expression and Fund Flow

This table presents coefficient estimates from OLS regressions of fund flows on rhetorical intensity scores. The sample includes all actively managed equity funds from 2009Q1 to 2024Q4, with observations constructed at the fund-quarter level. All independent variables are lagged by one quarter. *REScore* is the rhetorical intensity scores generated by Qwen-3-Max based on the prompt in Appendix Table A1. *RERatio* is the number of rhetorical devices divided by the number of sentences in the report. Textual characteristics (*Sentiment, Length* and *Readability*) and funds fundamentals (*TNA, Age, Return* and *FTNA*) are also controlled. The detailed variable definitions are provided in Appendix Table A2. Time (quarter-level) and fund fixed effects are included. Standard errors are clustered at the fund level. *t*-values are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

	(1)	(2)	(3)	(4)
		Flow		
REScore	**0.021*** **	**0.025*** **		
	(3.432)	**(3.782)**		
RERatio			**0.018*** **	**0.018*** **
			(3.274)	**(3.153)**
Sentiment		0.011**		0.012**
		(2.140)		(2.295)
Length		-0.018**		-0.008
		(-2.189)		(-1.024)
Readability		-0.005		-0.000
		(-0.742)		(-0.020)
TNA	-0.448***	-0.448***	-0.448***	-0.448***
	(-26.619)	(-26.656)	(-26.636)	(-26.633)
Age	0.158***	0.157***	0.157***	0.156***
	(5.374)	(5.323)	(5.340)	(5.291)
Return	0.138***	0.137***	0.138***	0.137***
	(24.293)	(24.253)	(24.311)	(24.281)
FTNA	0.059**	0.061**	0.060**	0.061**
	(2.468)	(2.555)	(2.489)	(2.544)
Constant	-0.006	-0.006	-0.006	-0.006
	(-0.534)	(-0.485)	(-0.517)	(-0.470)
N	53721	53721	53721	53721
R-squared	0.114	0.114	0.114	0.114
Fund FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Table 3 Heterogeneity Analysis Based on Fund Characteristics

This table presents the heterogeneity analysis using fund characteristics of the correlation between rhetorical intensity and subsequent fund flow. The sample includes all actively managed equity funds from 2009Q1 to 2024Q4, with observations constructed at the fund-quarter level. In columns (1) and (2), *Small Fund* refers to funds with size lower than the cross-sectional median while *Large Fund* refers the opposite. In column (3) and (4), *High Retail* (*Low Retail*) refers to holding proportion of retail investors is higher (lower) than 70%. All independent variables are lagged by one quarter. The detailed variable definitions are provided in Appendix Table A2. Time (quarter-level) and fund fixed effects are included. Standard errors are clustered at the fund level. *t*-values are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					Flow			
	Small Fund	*Large Fund*	*High Retail*	*Low Retail*	*Small Fund*	*Large Fund*	*High Retail*	*Low Retail*
REScore	0.050***	0.004	0.024***	0.030				
	(3.930)	(0.582)	(3.737)	(1.618)				
RERatio					0.044***	-0.003	0.019***	0.016
					(3.988)	(-0.601)	(3.184)	(1.164)
Sentiment	0.013	0.006	0.005	0.006	0.014	0.007	0.006	0.007
	(1.417)	(1.276)	(1.008)	(0.514)	(1.504)	(1.403)	(1.146)	(0.589)
Length	-0.028*	-0.001	-0.015*	0.012	-0.010	0.001	-0.006	0.025
	(-1.826)	(-0.064)	(-1.816)	(0.549)	(-0.667)	(0.159)	(-0.734)	(1.157)
Readability	0.009	-0.013**	0.001	-0.001	0.021	-0.014**	0.006	0.003
	(0.757)	(-1.999)	(0.102)	(-0.044)	(1.590)	(-2.124)	(0.861)	(0.155)
TNA	-0.612***	-0.477***	-0.383***	-0.578***	-0.611***	-0.476***	-0.382***	-0.577***
	(-17.075)	(-16.678)	(-16.790)	(-15.452)	(-17.107)	(-16.679)	(-16.776)	(-15.423)
Age	0.241***	0.014	0.183***	-0.027	0.238***	0.013	0.183***	-0.028
	(3.885)	(0.374)	(5.528)	(-0.301)	(3.834)	(0.356)	(5.512)	(-0.308)
Return	0.145***	0.130***	0.109***	0.202***	0.145***	0.130***	0.109***	0.202***
	(16.139)	(21.919)	(19.189)	(15.353)	(16.189)	(21.925)	(19.218)	(15.347)
FTNA	0.100***	0.053*	0.093***	-0.043	0.100***	0.053*	0.093***	-0.043
	(2.814)	(1.775)	(3.639)	(-0.738)	(2.836)	(1.773)	(3.627)	(-0.752)
Constant	-0.403***	0.325***	0.002	0.018	-0.402***	0.325***	0.002	0.018
	(-12.995)	(10.553)	(0.111)	(0.486)	(-12.996)	(10.561)	(0.118)	(0.491)
Observations	24995	28637	37632	14065	24995	28637	37632	14065
R-squared	0.133	0.170	0.163	0.189	0.133	0.170	0.163	0.188
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 4 Rhetoric Expression and Investor Initial Purchase Decisions

This table presents coefficient estimates from OLS regressions for the influence of rhetorical intensity on investor initial purchase decision. The sample includes 200,000 retail investors provided by Ant Open Research Laboratory and covers all actively managed equity funds they held from January 2019 to December 2020. Observations are constructed at the investor–fund–month level based on their transaction records. We exclude inactive investors with monthly portfolio size less than 100 RMB. The dependent variable is monthly investor net purchase value, measured as the total amount of investor purchases minus investor redemptions within a month. The independent variable is the interaction term between rhetoric intensity and *FirstPurc* (equals to 1 if it is the first time the user buys the fund). The detailed variable definitions are provided in Appendix Table A2. Time (month-level), fund and user fixed effects are included. The standard errors are clustered at the user level. t-values are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

	(1)	(2)
	NetPurc	
REScore × *FirstPurc*	**0.146*** **	
	(65.511)	
REScore	**-0.023*** **	
	(-22.400)	
RERatio × *FirstPurc*		**0.110*** **
		(40.440)
RERatio		**-0.016*** **
		(-18.302)
Sentiment	-0.004***	-0.001
	(-4.702)	(-1.424)
Length	0.007***	0.007***
	(4.270)	(4.627)
Readability	-0.007***	-0.006***
	(-5.991)	(-5.278)
TNA	-0.157***	-0.158***
	(-44.089)	(-44.381)
Age	-0.028***	-0.014
	(-2.893)	(-1.501)
Return	0.017***	0.018***
	(22.509)	(24.566)
FTNA	-0.065***	-0.076***
	(-10.855)	(-12.652)
TNetPurc	-0.061***	-0.061***
	(-35.936)	(-35.960)
TPortSize	0.074***	0.075***
	(21.129)	(21.280)
N	2649923	2649923
R-sq	0.088	0.085
User FE	Yes	Yes
Fund FE	Yes	Yes
Time FE	Yes	Yes

Table 5 Timing Test Based on Individual Decisions

This table presents coefficient estimates from OLS regressions for the timing effect of rhetorical expressions. The sample includes 200,000 retail investors provided by Ant Open Research Laboratory and covers all actively managed equity funds they held from January 2019 to December 2020. Observations are constructed at the investor–fund–month level based on their transaction records. We exclude inactive investors with monthly portfolio size less than 100 RMB. The dependent variable is monthly investor net purchase value, measured as the total amount of investor purchases minus investor redemptions within a month. The key independent variable is the measure for rhetorical expressions. Columns (1) and (3) use only observations from the report release month, whereas columns (2) and (4) use only observations from the month prior to the report release. The detailed variable definitions are provided in Appendix Table A2. Time (month-level), fund and user fixed effects are included. The standard errors are clustered at the user level. t-values are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

	(1)	(2)	(3)	(4)
			NetPurc	
	Report Month	*Pre-Report Month*	*Report Month*	*Pre-Report Month*
REScore	**0.019*** **	**0.001**		
	(10.537)	**(0.675)**		
RERatio			**0.008*** **	**0.000**
			(4.953)	**(0.341)**
Sentiment	-0.011***	0.000	-0.010***	0.000
	(-7.458)	(0.238)	(-6.826)	(0.276)
Length	0.007**	-0.003*	0.014***	-0.002*
	(2.492)	(-1.817)	(5.215)	(-1.763)
Readability	0.003	-0.002*	0.004*	-0.002*
	(1.351)	(-1.922)	(1.783)	(-1.877)
TNA	-0.098***	0.005	-0.098***	0.005
	(-16.003)	(1.540)	(-15.999)	(1.531)
Age	-0.048**	0.002	-0.037**	0.002
	(-2.560)	(0.162)	(-1.994)	(0.209)
Return	0.033***	-0.001*	0.031***	-0.001*
	(24.880)	(-1.705)	(23.602)	(-1.718)
FTNA	0.037***	-0.001	0.045***	-0.000
	(3.180)	(-0.122)	(3.896)	(-0.064)
TNetPurc	-0.047***	0.118***	-0.047***	0.118***
	(-18.696)	(27.001)	(-18.694)	(27.001)
TPortSize	0.088***	0.011**	0.088***	0.011**
	(16.698)	(2.019)	(16.705)	(2.019)
N	836113	836113	836113	836113
R-sq	0.205	0.277	0.205	0.277
User FE	Yes	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Table 6 Controlling for Manager and Company Fixed Effects

This table presents coefficient estimates from OLS regressions of fund flows on rhetorical scores. The sample includes all actively managed equity funds from 2009Q1 to 2024Q4, with observations constructed at the fund-quarter level. All independent variables are lagged by one quarter. In columns (1)-(2) and (5)-(6), we include manager, fund and quarter fixed effects. In columns (3)-(4) and (7)-(8), we include family, fund and quarter fixed effects. The detailed variable definitions are provided in Appendix Table A2. Standard errors are clustered at the fund level. *t*-values are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					Flow			
REScore	**0.013****	**0.015****	**0.021*****	**0.025*****				
	(1.987)	**(2.172)**	**(3.427)**	**(3.777)**				
RERatio					**0.013****	**0.013****	**0.018*****	**0.018*****
					(2.203)	**(2.184)**	**(3.270)**	**(3.148)**
Sentiment		0.014***		0.011**		0.014***		0.012**
		(2.614)		(2.137)		(2.669)		(2.291)
Length		-0.013		-0.018**		-0.007		-0.008
		(-1.371)		(-2.186)		(-0.778)		(-1.022)
Readability		-0.001		-0.005		0.002		-0.000
		(-0.163)		(-0.741)		(0.322)		(-0.020)
TNA	-0.638***	-0.638***	-0.448***	-0.448***	-0.638***	-0.638***	-0.448***	-0.448***
	(-28.371)	(-28.348)	(-26.581)	(-26.618)	(-28.378)	(-28.345)	(-26.598)	(-26.595)
Age	0.221***	0.219***	0.158***	0.157***	0.221***	0.219***	0.157***	0.156***
	(5.351)	(5.315)	(5.367)	(5.316)	(5.352)	(5.320)	(5.332)	(5.284)
Return	0.128***	0.127***	0.138***	0.137***	0.128***	0.127***	0.138***	0.137***
	(22.894)	(22.840)	(24.258)	(24.218)	(22.903)	(22.854)	(24.276)	(24.246)
FTNA	0.029	0.030	0.059**	0.061**	0.029	0.030	0.060**	0.061**
	(0.821)	(0.867)	(2.465)	(2.551)	(0.829)	(0.868)	(2.485)	(2.540)
Constant	-0.019	-0.018	-0.006	-0.006	-0.019	-0.018	-0.006	-0.006
	(-1.124)	(-1.084)	(-0.533)	(-0.485)	(-1.129)	(-1.092)	(-0.516)	(-0.470)
Observations	53621	53621	53721	53721	53621	53621	53721	53721
R-squared	0.184	0.184	0.114	0.114	0.184	0.184	0.114	0.114
Manager FE	Yes	Yes	No	No	Yes	Yes	No	No
Company FE	No	No	Yes	Yes	No	No	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7 Evidence from a Quasi-experiment

This table presents coefficient estimates from OLS regressions of fund flows on rhetorical intensity scores. We adopt the implementation of *Measures for the Administration of Information Disclosure of Publicly Offered Securities Investment Funds* (MAIDF) as a quasi-experiment, which brings a more standardized and strict information disclosure framework for mutual funds. The dummy variable Post equals 1 after 2019Q3 when MAIDF starts to implement. The time interval spans from 2017Q3 to 2021Q3, with observations constructed at the fund-quarter level. The detailed variable definitions are provided in Appendix Table A2. Time (quarter-level) and fund fixed effects are included. Standard errors are clustered at the fund level. *t*-values are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

	(1)	(2)	(3)	(4)
			Flow	
REScore × *Post*	**-0.055*** **	**-0.054*** **		
	(-3.115)	**(-3.030)**		
RERatio × *Post*			**-0.037** **	**-0.037** **
			(-2.174)	**(-2.133)**
REScore	0.064***	0.061***		
	(4.235)	(3.912)		
RERatio			0.044***	0.040***
			(3.086)	(2.810)
Sentiment		0.012		0.014
		(1.353)		(1.573)
Length		-0.006		0.005
		(-0.347)		(0.315)
Readability		-0.012		-0.008
		(-0.942)		(-0.652)
TNA	-0.558***	-0.558***	-0.558***	-0.558***
	(-12.208)	(-12.258)	(-12.171)	(-12.203)
Age	0.223***	0.222***	0.230***	0.229***
	(3.097)	(3.084)	(3.191)	(3.179)
Return	0.114***	0.113***	0.114***	0.113***
	(11.427)	(11.319)	(11.431)	(11.332)
FTNA	0.028	0.031	0.028	0.030
	(0.528)	(0.577)	(0.526)	(0.562)
Constant	-0.051	-0.051	-0.055	-0.054
	(-1.399)	(-1.384)	(-1.493)	(-1.478)
Observations	15100	15100	15100	15100
R-squared	0.144	0.144	0.143	0.143
Fund FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Table 8 Presence and Frequency of Rhetorical Devices

This table presents coefficient estimates from OLS regressions of fund flows on presence and frequency of rhetorical devices. The sample includes all actively managed equity funds from 2009Q1 to 2024Q4. All independent variables are lagged by one quarter. In Panel A, we use several dummy variables as our independent variable. *# REDevice >0* (*# REDevice =1, # REDevice >1*) is a dummy variable equals to 1 if the report contains more than 0 (equals to 1, more than 1) rhetorical devices. Panel B reports the results of different rhetorical devices. The independent variable is the number of rhetorical devices divided by the number of sentences. For instance, *HyperboleRatio* means the number of hyperboles in the fund report divided by the number of sentences. The detailed variable definitions are provided in Appendix Table A.2. Time (quarter-level) and fund fixed effects are included. Standard errors are clustered at the fund-level. t-values are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Panel A: Total Frequency of Rhetorical Devices

	(1)	(2)	(3)	(4)
		Flow		
# REDevice >0	**0.026****	**0.030****		
	(2.314)	**(2.554)**		
# REDevice =1			**0.015**	**0.017**
			(1.138)	**(1.209)**
# REDevice >1			**0.037*****	**0.046*****
			(2.635)	**(3.142)**
Sentiment		0.012**		0.012**
		(2.259)		(2.260)
Length		-0.015*		-0.019**
		(-1.821)		(-2.175)
Readability		-0.005		-0.005
		(-0.695)		(-0.739)
TNA	-0.455***	-0.455***	-0.456***	-0.456***
	(-25.785)	(-25.806)	(-25.795)	(-25.832)
Age	0.158***	0.157***	0.158***	0.157***
	(4.979)	(4.924)	(4.983)	(4.927)
Return	0.138***	0.136***	0.138***	0.137***
	(23.689)	(23.627)	(23.692)	(23.634)
FTNA	0.050**	0.052**	0.050**	0.052**
	(2.078)	(2.154)	(2.073)	(2.160)
Constant	-0.016	-0.018	-0.017	-0.020
	(-1.052)	(-1.125)	(-1.117)	(-1.261)
Observations	51040	51040	51040	51040
R-squared	0.116	0.116	0.116	0.116
Fund FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Panel B: Frequency of Specific Rhetorical Devices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					Flow			
HyperboleRatio	**0.016*** **							
	(2.918)							
AntithesisRatio		**0.007**						
		(1.381)						
ParallelismRatio			**0.003**					
			(0.628)					
MetaphorRatio				**0.010** **				
				(1.981)				
PersonificationRatio					**0.004**			
					(0.989)			
QuotationRatio						**-0.006**		
						(-1.243)		
HypophoraRatio							**0.004**	
							(0.793)	
RhetoricalQuestionRatio								**-0.004**
								(-0.914)
Sentiment	0.012**	0.013**	0.013**	0.013**	0.013**	0.013**	0.013**	0.013**
	(2.407)	(2.549)	(2.458)	(2.533)	(2.528)	(2.548)	(2.557)	(2.543)
Length	-0.008	-0.008	-0.008	-0.009	-0.008	-0.007	-0.008	-0.008
	(-0.972)	(-1.013)	(-0.976)	(-1.114)	(-1.036)	(-0.900)	(-1.035)	(-0.969)
Readability	-0.002	-0.004	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005
	(-0.278)	(-0.568)	(-0.676)	(-0.676)	(-0.742)	(-0.774)	(-0.763)	(-0.774)
TNA	-0.447***	-0.447***	-0.446***	-0.447***	-0.446***	-0.446***	-0.446***	-0.446***
	(-26.591)	(-26.614)	(-26.597)	(-26.650)	(-26.595)	(-26.606)	(-26.608)	(-26.598)
Age	0.155***	0.155***	0.155***	0.155***	0.155***	0.154***	0.154***	0.154***
	(5.259)	(5.253)	(5.245)	(5.251)	(5.245)	(5.240)	(5.237)	(5.238)
Return	0.137***	0.137***	0.137***	0.137***	0.137***	0.137***	0.137***	0.137***
	(24.257)	(24.295)	(24.290)	(24.287)	(24.293)	(24.307)	(24.302)	(24.297)
FTNA	0.061**	0.061**	0.062**	0.062**	0.062**	0.061**	0.061**	0.062**
	(2.538)	(2.549)	(2.554)	(2.556)	(2.554)	(2.544)	(2.551)	(2.558)
Constant	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005
	(-0.455)	(-0.446)	(-0.436)	(-0.437)	(-0.438)	(-0.434)	(-0.426)	(-0.431)
Observations	53721	53721	53721	53721	53721	53721	53721	53721
R-squared	0.114	0.114	0.114	0.114	0.114	0.114	0.114	0.114
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 9 Rhetorical Expression and Fund Performance

This table reports the panel regressions of fund performance on rhetorical expression and various fund characteristics, where independent variables are lagged by one quarter. The sample includes all actively managed equity funds from 2009Jan to 2024Dec. The observations are at the fund-month level. Independent variables are various fund performance proxies. *Ret* is the monthly raw return. *Alpha CAPM*, *Alpha 3-factor* and *Alpha 4-factor* are abnormal return-adjusted by CAPM, Fama-French 3 factor model, and Carhart 4 factor model respectively. The detailed variable definitions are provided in Appendix Table A2. Time (month-level) and fund fixed effects are included. The standard errors are clustered on fund level. *t*-values are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret	*Alpha CAPM*	*Alpha 3-factor*	*Alpha 4-factor*	*Ret*	*Alpha CAPM*	*Alpha 3-factor*	*Alpha 4-factor*
REScore	**0.006**	**0.004**	**0.005**	**0.004**				
	(1.461)	**(1.076)**	**(1.130)**	**(0.998)**				
RERatio					**0.003**	**0.002**	**0.004**	**0.005**
					(0.902)	**(0.547)**	**(1.204)**	**(1.424)**
Sentiment	0.000	-0.000	-0.001	-0.000	0.001	-0.000	-0.001	-0.000
	(0.154)	(-0.079)	(-0.328)	(-0.018)	(0.234)	(-0.010)	(-0.306)	(-0.028)
Length	-0.003	-0.004	-0.003	-0.004	-0.001	-0.003	-0.001	-0.002
	(-0.581)	(-0.869)	(-0.651)	(-0.773)	(-0.138)	(-0.551)	(-0.310)	(-0.484)
Readability	-0.002	-0.004	-0.004	-0.002	-0.001	-0.003	-0.003	-0.001
	(-0.538)	(-0.888)	(-0.978)	(-0.553)	(-0.332)	(-0.750)	(-0.704)	(-0.232)
TNA	-0.135***	-0.125***	-0.108***	-0.097***	-0.134***	-0.125***	-0.108***	-0.097***
	(-16.251)	(-15.702)	(-14.227)	(-12.312)	(-16.204)	(-15.670)	(-14.233)	(-12.338)
Age	0.027***	0.018*	0.027**	0.027**	0.027***	0.018*	0.027**	0.027**
	(2.612)	(1.750)	(2.557)	(2.423)	(2.586)	(1.729)	(2.547)	(2.419)
Return	-0.018***	-0.004	-0.005	-0.014***	-0.018***	-0.004	-0.005	-0.014***
	(-6.181)	(-1.402)	(-1.556)	(-4.713)	(-6.166)	(-1.390)	(-1.543)	(-4.701)
FTNA	-0.037***	-0.039***	-0.035***	-0.023*	-0.037***	-0.039***	-0.035***	-0.023*
	(-2.863)	(-3.110)	(-2.891)	(-1.845)	(-2.864)	(-3.110)	(-2.896)	(-1.851)
Constant	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(-5.286)	(-4.895)	(-8.640)	(-4.107)	(-5.164)	(-4.806)	(-8.577)	(-4.021)
Observations	152850	152850	152850	152850	152850	152850	152850	152850
R-squared	0.023	0.023	0.026	0.027	0.023	0.023	0.026	0.027
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 10 Rhetorical Expression and Subsequent Risk-taking

This table presents the results of panel regressions examining the relationship between rhetorical expression and various measures of risk-taking behavior, controlling for a set of fund characteristics. The sample includes all actively managed equity funds from 2009Jan to 2024Dec, with observations constructed at the fund-month level. In columns (1) and (4), the dependent variables are *Total Risk*, measured by the standard deviation of daily return within a month. In columns (2) and (5), the dependent variables are *Sys Risk*, the market beta estimated by the Fama-French three-factor model using daily return within a month. In columns (3) and (6), the dependent variables are *Idio Risk*, defined as the standard deviation of residuals estimated by the Fama-French three-factor model. The detailed variable definitions are provided in Appendix Table A2. Time (month-level) and fund fixed effects are included. The standard errors are clustered on fund level. *t*-values are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Risk	*Sys Risk*	*Idio Risk*	*Total Risk*	*Sys Risk*	*Idio Risk*
REScore	**0.060*** **	**0.030*** **	**0.057*** **			
	(7.269)	**(4.172)**	**(7.101)**			
RERatio				**0.029*** **	**0.010***	**0.029*** **
				(4.374)	**(1.764)**	**(4.697)**
Sentiment	0.035***	0.022***	0.027***	0.038***	0.024***	0.030***
	(6.839)	(5.055)	(5.255)	(7.304)	(5.415)	(5.713)
Length	-0.025**	-0.015	-0.034***	-0.001	-0.003	-0.012
	(-2.190)	(-1.503)	(-3.025)	(-0.123)	(-0.323)	(-1.078)
Readability	-0.007	-0.019***	0.001	0.001	-0.016**	0.008
	(-0.814)	(-2.778)	(0.070)	(0.136)	(-2.360)	(0.968)
TNA	-0.017	-0.013	-0.078***	-0.014	-0.011	-0.075***
	(-0.852)	(-0.734)	(-3.872)	(-0.692)	(-0.620)	(-3.711)
Age	0.013	0.048**	-0.002	0.010	0.047**	-0.004
	(0.527)	(2.341)	(-0.063)	(0.419)	(2.266)	(-0.162)
Return	0.036***	0.021***	0.058***	0.037***	0.021***	0.059***
	(9.338)	(5.664)	(14.586)	(9.443)	(5.726)	(14.661)
FTNA	0.035	-0.012	0.011	0.035	-0.012	0.011
	(1.064)	(-0.430)	(0.333)	(1.055)	(-0.430)	(0.326)
Constant	-0.000***	-0.000***	-0.000	-0.000***	-0.000***	0.000
	(-5.601)	(-13.519)	(-0.030)	(-5.031)	(-13.290)	(0.563)
Observations	152850	152850	152850	152850	152850	152850
R-squared	0.408	0.280	0.403	0.407	0.279	0.402
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 11 Rhetorical Expression and Subsequent Style Drift

This table presents the results of panel regressions examining the relationship between rhetorical intensity and style drift, controlling for a set of fund characteristics. The sample includes all actively managed equity funds from 2009Jan to 2024Dec, with observations constructed at the fund-month level. The dependent variables is *Abs(ReturnBench)*, the absolute value of monthly return comparing to the benchmark index. The detailed variable definitions are provided in Appendix Table A2. Time (month-level) and fund fixed effects are included. The standard errors are clustered on fund level. *t*-values are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

	(1)	(2)
	Abs(ReturnBench)	
REScore	**0.027*** **	
	(4.721)	
RERatio		**0.011** **
		(2.466)
Sentiment	0.015***	0.016***
	(3.775)	(4.105)
Length	-0.011	-0.000
	(-1.432)	(-0.008)
Readability	-0.006	-0.003
	(-0.998)	(-0.484)
TNA	-0.046***	-0.045***
	(-3.679)	(-3.548)
Age	0.022	0.021
	(1.296)	(1.219)
Return	0.030***	0.030***
	(8.933)	(9.004)
FTNA	0.023	0.022
	(1.056)	(1.051)
Constant	0.001	0.001
	(1.011)	(0.948)
Observations	139691	139691
R-squared	0.128	0.127
Fund FE	Yes	Yes
Time FE	Yes	Yes

Table 12 Rhetoric Expression and Subsequent Flow-performance Sensitivity

This table presents coefficient estimates from OLS regressions for subsequent flow-performance sensitivity. The sample includes all actively managed equity funds from 2009Q1 to 2024Q4, with observations constructed at the fund-quarter level. The dependent variable is quarterly mutual fund flow. The independent variables include the interaction term between rhetorical intensity and raw return. We also include *High* (*Low*), which equals 1 if the raw return is in the top (bottom) quintile. The detailed variable definitions are provided in Appendix Table A2. Time (quarter-level) and fund fixed effects are included. The standard errors are clustered at the fund level. t-values are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

	(1)	(2)	(3)	(4)
			Flow	
	RE = REScore		*RE = RERatio*	
RE×Return	**0.114*** **		**0.070*** **	
	(18.228)		**(12.134)**	
RE×High		**0.106*** **		**0.078*** **
		(16.811)		**(11.861)**
RE×Low		**-0.031*** **		**-0.025*** **
		(-7.943)		**(-6.322)**
RE	0.022***	0.009	0.015***	-0.003
	(3.348)	(1.341)	(2.655)	(-0.499)
Sentiment	0.010**	0.011**	0.011**	0.012**
	(1.977)	(2.247)	(2.180)	(2.434)
Length	-0.017**	-0.018**	-0.008	-0.008
	(-2.091)	(-2.168)	(-0.999)	(-1.027)
Readability	-0.005	-0.006	-0.001	-0.001
	(-0.703)	(-0.850)	(-0.094)	(-0.095)
TNA	-0.427***	-0.427***	-0.442***	-0.439***
	(-25.608)	(-25.442)	(-26.425)	(-26.261)
Age	0.151***	0.150***	0.154***	0.151***
	(5.224)	(5.226)	(5.242)	(5.206)
Return	0.141***	0.140***	0.139***	0.139***
	(25.050)	(24.968)	(24.606)	(24.738)
FTNA	0.069***	0.068***	0.066***	0.066***
	(2.909)	(2.907)	(2.761)	(2.800)
Constant	-0.005	-0.005	-0.005	-0.005
	(-0.418)	(-0.412)	(-0.433)	(-0.392)
Observations	53721	53721	53721	53721
R-squared	0.128	0.128	0.119	0.121
Fund FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes