Seek the Seeking Alpha: The Role of Social Media Research on the Behaviour and Economic Outcomes of Mutual Fund Investors*

Sha Liu[1], Wenjie Ding[2], Ariel Gu[3], Qingwei Wang[4]

This version: January 2024

Abstract

This paper investigates the impact of social media investment research on mutual fund investor behaviour and economic outcomes. We find that increased social media research coverage of stocks held by mutual funds predicts higher short-term fund flows but not better long-term fund returns, indicating an attention-driven response from investors. Higher bullishness of such coverage, while predicts lower fund returns, is associated with a near-term fund inflow but a longer-term outflow. As such, it hints that more sophisticated investors leverage sentiment for contrarian bets. The impact of social media research on fund flows is more pronounced post-filing of holdings and in funds with higher investor recognition or those holding stocks with greater social media visibility.

Keywords: Social media; Investor attention; Investor sentiment; Fund performance; Fund flows.

JEL Classification: G11; G12; G14; G23.

* This is a preliminary draft. Please consult the authors before citation. All errors are our own.

[1] School of Finance, Southwestern University of Finance and Economics, China
[2] Corresponding Author, Business School, Sun Yat-sen University, China
[3] School of Economics, University of East Anglia
[4] Cardiff Business School, Cardiff University

1. Introduction

The rise of social media in the digital age has transformed how investors access information and interact with others. Crowd-sourced platforms such as Seeking Alpha (*SA*) provide low-cost financial analysis and commentary to over 20 million users every month in 2023,[5] catering particularly to individual investors (see, Gomez et al., 2022; Farrell et al., 2022; Dim, 2023). Amidst this new information landscape, one crucial question remains, how does investment research on social media platforms (referred to as 'social media research' or *SMR*) such as *SA* affect mutual fund investors who hold a lion's share of the US stock market?

This paper aims at addressing this question, arguing that such an effect can be multifaceted. On one hand, wisdom of crowds (WOC) posits that, with a large community of contributors, *SMR* offers a wide range of investment perspectives to allow collective intelligence outperforming individual judgements. Therefore, *SA* has the potential to attract investors to invest in mutual funds with large exposure to *SA*-covered stocks and help them achieve high investment returns. Consistent with this WOC view, the literature has found that social media can play an informational role in the stock market by conveying novel and value-relevant news (Bartov et al., 2018; Gu and Kurov, 2020). Particularly, *SA* research predicts returns (Chen et al., 2014; Dim, 2023), lead to more informed trading (Farrell et al., 2022) and less information asymmetry in earnings announcements (Gomez et al., 2022).

On the other hand, at least some *SMR* may amplify behavioural bias and fuel attention-driven or sentiment-driven trading, the types of irrational behaviour extensively documented in the literature (e.g., Barber and Odean, 2008; Barber et al., 2022; De Long et al., 1990; Baker and Wurgler, 2006). Unlike traditional forms of social interaction, such as conversations among a small group of individuals like neighbours, colleagues, or between advisors and clients,

[5] https://about.seekingalpha.com/?source=footer

platforms like *SA* facilitate the rapid and wide dissemination of information and opinions to a broad investment community. This dynamic can escalate minor biases into significant market impacts through a compounding effect driven by user interactions. Theoretical models and empirical findings suggest the presence of a 'social transmission bias' when investment strategies are communicated and adopted within social networks (Hirshleifer, 2020; Han et al., 2022). Content on investment-focused social media can create or exacerbate disagreements among less sophisticated investors, potentially leading to persistent excessive trading (Hirshleifer et al., 2023). Farrell et al. (2022) find that a small subset of *SA* research reports induces uninformed retail trading that causes price deviations from fundamentals over short horizons.

Before delving into the informational or behavioural role of *SMR* on mutual fund investors, it is essential to test the hypothesis that investors react to *SMR* by adjusting their fund investments, which can be observed through net fund flows. To effectively capture the *SMR* features of mutual funds, we focus on two prominent aspects of *SA* research on individual stocks: coverage and sentiment. Coverage is quantified by the number of *SA* articles about a specific stock, which serves as a measure of the attention *SA* pays to that stock. Sentiment, on the other hand, is gauged by the overall tone of these articles, assessing their bullishness or bearishness towards the stock. The stock-level coverage and sentiment metrics are aggregated to the fund level, resulting in three variables that reflect the *SMR* characteristics of mutual funds. The first two measures reflect the incremental *SA* coverage of a fund's holdings during a particular period. The third variable captures the change in *SA* sentiment regarding the stocks held by a fund. These variables exhibit minimal correlation with the characteristics of the fund or its holdings. This lack of correlation is particularly important for our empirical analysis as it helps mitigate endogeneity concerns. We propose that investors are likely to respond to these variables because these features either capture their attention or provide information perceived as valuable.

Our fund-level *SMR* variable, derived from stock-level data rather than direct *SA* coverage or sentiment about specific funds, is posited to influence investors' decisions in mutual fund investments. For individual investors who show interest in popular and bullish stocks on social media, delving deeply into news and financial statements of each stock can be a daunting and costly task. Instead, these investors might opt for mutual funds that include a variety of stocks covered by *SA*. The rationale behind this choice is that mutual funds are professionally managed and offer well-diversified portfolios, thus providing a more efficient alternative in terms of information and opportunity costs. Additionally, investors who already hold mutual funds and follow *SA* articles are likely to monitor the collective *SA* attention and sentiment of the fund holdings. This aggregated information from *SA* can serve as a key factor in their decision-making process, influencing whether they choose to buy more shares or redeem some of their holdings in a particular fund.

In our analysis of 1,141 actively managed U.S. equity funds from January 2009 to March 2020, we discover that a greater increase in *SA* coverage predicts higher fund flows in the subsequent quarter or month, after accounting for fund characteristics, past fund performance, and fund-level measures of traditional media coverage or tone. Specifically, a one standard deviation increase in the percentage change of the number of stocks covered by *SA* is associated with a 0.60% increase in fund flows. Similarly, a one standard deviation increase in the percentage change of the number of *SA* articles about stocks held by a fund leads to a 4.34% increase in fund flows. This effect is more pronounced in funds with higher investor recognition (such as those that are large, old, or low-cost) or those holding stocks that receive higher *SA* coverage (including large, growth-oriented, high-priced, or low-idiosyncratic-volatility stocks), especially for the incremental number of *SA* articles about stocks held by a fund.

To address concerns that the observed relationship between incremental *SA* coverage and fund flows might be influenced by omitted factors, we follow the methodology used by Agarwal et al. (2022) and implement an event study and a Difference-in-Differences (DID) analysis. The 'event' in this context is defined as the SEC filing of a fund's holdings at the end of each quarter, which we use to examine the reaction in fund flows to these disclosures. In this analysis, funds with high incremental *SA* coverage are designated as the treatment group, while those with low coverage are considered the control group. The objective is to discern whether the treatment group, with greater increase in *SA* attention, exhibits different flow patterns compared to the control group, particularly in response to the event of holdings disclosure. Our findings provide robust evidence that the impact of incremental *SA* attention on fund flows is significantly more pronounced immediately following the full disclosure of fund holdings at the end of a quarter. This is in stark contrast to the months leading up to the filing (pre-filing months).

Regarding future performance, we find that the two-quarter performance of mutual funds positively correlates with both measures of incremental *SA* attention. This short-term outperformance may be attributed to a temporary increase in demand for stocks held by the funds, driven by heightened investor attention. However, this positive relationship reverses in the third and fourth quarters, culminating in no significant impact of these measures on one-year fund performance. These findings support the *behavioural view* that the influence of *SA* coverage is primarily on investor attention and their immediate investment decisions, rather than on providing informational content that would benefit long-term fund performance.

The relationship between the change in fund-level *SA* bullish sentiment and fund flows presents a more complex narrative. A positive correlation exists between the monthly change in *SA* sentiment and subsequent month's fund flows. However, an increase (or decrease) in *SA* sentiment over a quarter is linked with lower (or higher) fund flows in the following quarter. Specifically, a one standard deviation increase (or decrease) in the change of *SA* sentiment is associated with a 2.56% decline (or increase) in fund flows. The event study and DID analysis further reveal that the positive effect of changes in *SA* attention on monthly fund flows is more pronounced in the months following the fund's SEC filing.

These findings might indicate that different investor responses to short-term versus longer-term changes in *SA* sentiment. It is also possible that sophisticated investors initially favour funds holding stocks with an increased bullish sentiment on *SA* to capitalise on the short-lived bullish sentiment, yet they tend to divest from these funds over a longer period before the revelation of underlying fundamentals or a shift in prevailing sentiment. Additionally, investors might opt to invest in funds that have experienced a decrease in bullishness, typically associated with poorer past performance, in anticipation of potential price appreciation or a reversal in trends. This behaviour aligns with a contrarian investment approach, where investors seek to exploit market inefficiencies or temporary mispricing resulting from sentiment-driven movements.

Further analysis reveals that fund investors' negative response to increased bullishness on *SA* seems to be a prudent decision. Specifically, a higher increase in *SA* bullishness correlates with poorer short-term performance of funds. Interestingly, this effect does not reverse in the following two quarters. Consequently, the cumulative one-year performance of the funds is significantly and negatively correlated with the increase in *SA* bullishness during the current quarter. This finding implies that the level of sentiment expressed by *SA* articles about fund holdings is not merely noise; instead, it provides valuable information, especially for contrarian investors, supporting the *information-based view* of the role of *SMR*.

Our paper makes several contributions to the existing literature. The existing body of research on social media and asset pricing has extensively documented the impact of social media coverage and tone, regardless of their connection to specific news events, on stock trading and short-term returns (Sprenger et al., 2014a and 2014b; Renault, 2017; Jia et al., 2020; Jiao et al., 2020; Rakowski et al., 2021; Gu and Kurov, 2020; Liu et al., 2022). Building upon these findings, our study contributes by demonstrating the influence of social media coverage and content on mutual fund investing, a crucial segment of the stock market. By ensuring that the coverage and sentiment measure we create are not merely reflections of intrinsic fund or stock characteristics, we can more confidently attribute any observed effects on fund flows or performance to the influence of social media research itself, rather than to underlying fund attributes. More intriguingly, we demonstrate that the attention and opinions expressed about *individual stocks* on social media platforms can be translated into mutual fund investment decisions.

This research also makes a meaningful contribution to a growing but more focused area of literature on the role of social media research and social media analysts in financial markets (e.g., Chen et al., 2014; Campbell et al., 2019; Drake et al., 2022; Farrell et al., 2022; Gomez et al., 2022; Dim, 2023; Cookson et al., 2023). Our work delves into what social media research, such as articles from crowd-sourced platforms, can offer mutual fund investors. Specifically, we explore whether there are tangible benefits to following investment research on these platforms. Our findings indicate that mutual fund investors who focus on the aggregate research coverage of stocks within a fund do not reap long-term benefits from this approach. However, a contrarian perspective towards the collective sentiment expressed in social media research about fund holdings inform successful investment decisions, a finding consistent with Farrell et al. (2022)' s conclusion that investors actively glean valuable information from *SA* rather than trading directly on article sentiment.

Lastly, this research complements existing studies on the impact of *traditional* media coverage on mutual fund flows, performance, and fund managers' investment decisions. This body of work includes research by Sirri and Tufano (1998), Fang et al. (2014), Solomon et al. (2014), and Kaniel and Parham (2017). While mixed results are observed regarding the relationship between traditional media coverage of fund holdings and mutual fund flows (Fang et al., 2014 versus Solomon et al., 2014), we show that social media coverage and sentiment regarding fund holdings do have a significant effect on fund flows, after controlling for the traditional media coverage or sentiment of fund holdings. Sirri and Tufano (1998) and Kaniel and Parham (2017) base their analyses on the mentions of mutual fund names in major newspapers, while our approach to measuring fund-level attention and sentiment is rooted in the analysis of coverage and sentiment of individual stocks within these funds. By focusing on fund holdings' representation on social media, we reveal that the attention and opinions expressed about individual stocks on social media platforms can lead to mutual fund investment decisions.

The rest of the paper is organised as follows. Section 2 reviews the related literature and develops our hypotheses. Section 3 describes our mutual fund sample, *SA* and traditional media data sets and how *SMR* variables are constructed and provides summary statistics. Section 4 reports the empirical analysis and discusses our findings. Section 5 concludes.

2. Literature Review and Hypotheses Development

2.1 Social media and asset pricing

The scholarly focus on the impact of social media platforms on the financial market is pervasive; however, the implications of social media analyst reports are diverse. On one hand, a strand of literature supports the argument that social media analyst reports are beneficial for the financial market (Chen et al., 2014; Bartov et al., 2018; Campbell et al., 2019; Gu and Kurov, 2020;

Farrell et al., 2022). One supportive argument posits that by revealing timely, valuable or firm-specific information, social media research reports enhance the integration of information into stock prices and improve the market quality. For example, Drake et al. (2021) demonstrate that social media analysts divulge decision-useful information akin to sell-side analyst reports but in a timely manner. Another theory supporting the positive impact of the aggregate opinion (i.e., positivity/negativity) of social media reports is the wisdom-of-the-crowd theory (e.g., Chen et al., 2014; Dim, 2023).

On the other hand, social media can amplify biased investor behaviour by affecting investor sentiment or disseminating outdated information (e.g., Heimer, 2016; Chawla et al., 2021; Chen and Hwang, 2022; Cookson et al., 2023). Chen and Hwang (2022) introduce a novel perspective on the overpricing of stocks covered by *SA*, suggesting that the impression management considerations of *SA* analysts lead to the propagation of noise. Cookson et al. (2023) point out the 'echo chambers' phenomenon in the social media platform, wherein investors selectively expose themselves to information aligning with their pre-existing beliefs or opinions. Their findings indicate that being in an echo chamber exacerbates misinformed sentiment.

2.2 Investor sentiment, attention, and mutual fund flows

Investor sentiment and attention are two different concepts which both lead to mispricing. Investor sentiment is broadly defined as the biased belief of investors which makes price deviating from its fundamental value when there are limits of arbitrage (Baker and Wurgler, 2007). Mutual fund flow has been widely acknowledged as a financial outcome closely tied to investor sentiment. For example, Akbas et al (2015) show that aggregate mutual fund flows represent dumb money and exacerbate the stock market anomalies. Ben-Rephael et al. (2012) differentiate between various types of funds, measuring investor sentiment through shifts in mutual fund flows between bond funds and equity funds. Cooper et al. (2005), Lou (2012), and Kamstra et al. (2017), amongst others, also illustrate that mutual fund flows contain information about investor demands and/or sentiment.

Investor attention is scarce (Kahneman, 1973; Peng and Xiong, 2006). Investors constrained by attention, particularly individual investors, encounter a substantial search challenge when making purchases, but less so when selling. Therefore, they tend to be net purchasers of attention-grabbing stocks (Barber and Odean, 2008). Sirri and Tufano (1998) argue that investors face a costly search problem. Mutual funds with high media attention, high marketing efforts and strong prior performance lower investors' search cost, therefore attracting more fund flows. Mutual funds may enhance advertising efforts to attract investors, whether through traditional media (e.g., Jain et al., 2000) or, more recently, through social media (Gil-Bazo, 2020), given its rapid development. Cooper et al (2005) find that mutual funds changing their names to reflect hot investment styles leads to abnormal fund flows, indicating that investors are susceptible to a cosmetic effect. Solomon et al. (2014) assert that the media coverage of fund holdings, a proxy of attention, affect investors' capital allocations to mutual funds.

2.3 Hypotheses development

Social media sentiment and attention, as potential indicators of investor sentiment and attention, exhibit distinct characteristics and result in different economic impacts. For example, Cookson et al. (2023) discovered a positive correlation between social media sentiment and the next-day stock return, in contrast to a negative relationship between social media attention and the next-day return. Furthermore, even across various social media platforms, sentiment and attention patterns can differ. Cookson et al. (2023) noted a high correlation in investor attention across platforms like Twitter, Stocktwits, and *SA*, but observed that the levels of attention varied among them. Therefore, it is crucial to distinguish between the effects of social media sentiment and attention. While both are important in understanding market movements and investor behaviour, they contribute in different ways and to varying extents.

Before examining whether *SMR* affect mutual fund flows through attention-driven, sentiment-driven, or information-driven trading, our initial hypothesis posits that incremental *SA* attention or bullish sentiment at the fund level positively predicts fund flows in the subsequent period. This hypothesis is grounded in the extensive literature documenting the impact of social media coverage and sentiment on stock trading and short-term returns, as explored in studies by Sprenger et al. (2014a and 2014b), Renault (2017), Jia et al. (2020), Jiao et al. (2020), Rakowski et al. (2021), and Gu and Kurov (2020). Given this background, it is reasonable to hypothesise that *SA* attention and sentiment similarly influence mutual fund investment decisions. This hypothesis also aligns with the assumption that investors, particularly individual investors, are likely to be net purchasers of stocks receiving significant social media attention or bullish sentiment, as demands for these stocks are partly driven by excessive attention [6] and overoptimistic views. While investors have the option to buy individual stocks that are active or bullish on social media, this approach is often more costly and less diversified compared to investing in mutual funds holding a bunch of such stocks.

Hypothesis 1*: The fund-level incremental SA attention or bullish sentiment positively predicts next-period fund flows.*

When testing *Hypothesis 1*, we also operate under the assumption that fund investors can view complete fund holdings after the funds' quarterly filings with the SEC. In addition, to alleviate the concerns that the relationship between incremental *SA* coverage and fund flows might be

[6] Traditional proxies for investor attention include traditional media coverage (Solomon et al., 2014), extreme price movements (Barber and Odean, 2008), and advertising expenses (Lou, 2014). These proxies, however, capture only passive investor attention. Unlike news coverage on traditional media, a large proportion of social media posts are written by their users and are then shared with their fellow users. Therefore, social media coverage can potentially capture both active and passive investor attention on these social media platforms.

influenced by omitted factors, we follow the framework of Agarwal et al. (2022) and examine whether the effect is more pronounced in the post-filing months.

Hypothesis 2*: The reaction of fund flows to incremental SA attention or bullish sentiment is more pronounced in the post-filing months.*

The literature finds that *SA* primarily serves the information needs of individual investors over institutional ones (Farrell et al., 2022; Gomez et al., 2022; Dim, 2023). Individual investors, who often face significant challenges in stock selection due to limited attention resources (Barber and Odean, 2008), encounter similar issues when considering mutual fund investments. This leads to our next hypothesis that the impact of *SA* attention and sentiment on fund flows will be more pronounced in certain types of funds. For instance, larger funds typically enjoy greater investor recognition, so the influence of *SA* attention and sentiment on these funds' flows is expected to be more significant. Additionally, funds with holdings that attract higher investor attention on *SA* are also likely to experience a more substantial impact. These include funds with larger holdings in stocks that are frequently covered by *SA* analysts, such as large-cap, growth-oriented, and high-priced stocks, as summarized in the literature (Dim, 2023).

Hypothesis 3*: The impact of fund-level incremental SA attention or bullishness on fund flows is more pronounced in funds with higher investor recognition or those holding a greater proportion of stocks with higher visibility on SA.*

Hypotheses 4 and *5* are proposed to investigate the effect of *SA* attention and sentiment on mutual fund performance, respectively. They aim to address whether investors' reactions to *SMR* translate into any long-term benefits. Regarding fund-level *SA* attention, we posit that the stock-level *SA* attention primarily influences investor attention towards individual stocks, potentially leading to a temporary increase in demand for those stocks and the funds holding them. As this investment behaviour is driven by attention rather than fundamental analysis, we anticipate possible short-term outperformance but not sustained long-term benefits.

***Hypothesis 4**: The fund-level incremental SA attention is not correlated with fund performance in the long term.*

Regarding *SA* sentiment, we argue that it may contain valuable information that predicts long-term fund performance, reflecting the collective returns of the fund's stock holdings. This is supported by studies indicating that social media platforms, including *SA*, can disseminate novel and value-relevant information (Chen et al., 2014; Bartov et al., 2018; Gu and Kurov, 2020; Farrell et al., 2022; Gomez et al., 2022). However, the direction in which *SA* sentiment affects fund performance remains an open empirical question. For instance, Farrell et al. (2022) find that the incremental information revealed by post-research retail trading is largely orthogonal to the information revealed by report tone, suggesting that investors actively extract valuable information from *SA* research rather than making immediate trades based on report sentiment. Consequently, we hypothesize that fund-level incremental *SA* bullish sentiment could either positively or negatively correlate with long-term fund performance.

***Hypothesis 5a**: The fund-level incremental SA bullish sentiment is positively correlated with fund performance in the long term.*
***Hypothesis 5b:** The fund-level incremental SA bullish sentiment is negatively correlated with fund performance in the long term.*

3. Sample and Data Description

3.1 Mutual fund sample construction

Our dataset synergises mutual fund holdings data with stock-level social media variables. We construct our primary sample of mutual funds by merging the CRSP Survivor-Bias-Free Mutual Fund Database with the Thomson Reuters Mutual Fund Holdings Database. Aligning with the existing literature, our analysis is confined to actively managed, diversified U.S. domestic equity funds. Index funds, international funds, municipal bond funds, bond and preferred stock funds, sector-specific funds, and any funds that cannot be linked to the CRSP database via Wharton Research Data Services' Mutual Fund Links dataset are excluded from our sample. Furthermore, we require that at least 80% of a fund's assets under management must be allocated in common stocks to ensure the funds are genuinely equity focused.

The Thomson Reuters Mutual Fund Holdings Database provides quarterly updates on U.S. common stock holdings for mutual funds, specifically detailing long positions. For funds that do not regularly report quarterly updates of portfolio holdings, we use the most recent report to deduce holdings for those quarters. Due to the lack of intra-quarter trading data, we assume consistency in holdings throughout the report period.

Our study spans from January 2009 to March 2020, and we have instituted the following selection criteria: funds holding fewer than 10 stocks are excluded to affirm diversification; those managing less than $5 million in assets are excluded to concentrate on funds with a significant investment impact; and any fund with a history shorter than one year is excluded to guarantee a considerable track record for analysis. Furthermore, fund-quarters with fewer than 5 stocks identified as *SA* stocks (definition explained in Section 3.2) are also excluded to preserve the integrity of our *SA*-focused analysis. Our final sample consists of 1,141 mutual funds. After removing fund-quarters without *SA* data, we are left with 36,181 fund-quarter observations. The fund-quarter dataset is then merged with the social media variables, which will be further described in the subsequent section.

3.2 Seeking-Alpha and traditional media data

The social media coverage and sentiment data utilized in this study are derived from articles on Seeking Alpha (www.SeekingAlpha.com), a prominent investment-focused crowd-sourced platform. Touted as the world's largest investing community, *SA*'s registered contributors include a diverse mix of individual and institutional investors, fund managers, analysts, college students, retirees, and others who share their investment insights, expertise, and ideas. These contributors, known as social media analysts (SMAs), ensure a dynamic exchange of perspectives. *SA*'s editorial team maintains a minimum quality standard for published articles (Gomez et al., 2022). The literature finds that SMAs primarily cover stocks that are large, growth-oriented, high-priced, liquid, and have low idiosyncratic volatility (Dim, 2023). *SA*'s content is low-cost, initially free but now accessible through a subscription starting at $239 a year, which is more affordable than many other business information sources. Hence, SMAs mostly cater to the information needs of individual investors (see, Gomez et al., 2022; Farrell et al., 2022; Dim, 2023). Given these characteristics, *SA* serves as a particularly apt source for analysing mutual fund flows, as mutual funds are a popular investment choice for individual investors.

We designed a web-scraping algorithm to download all 'Long Ideas' investment articles pertaining to stocks, published on *SA* during our sample period. Each article is tagged with its publication date and the associated stock tickers. We then compile the daily total number of articles for each stock ever held by the mutual funds in our sample throughout the sample period. An '*SA Stock*' is defined as a stock with at least one article on *SA* during a certain period (either quarterly or monthly).

For the textual sentiment analysis, we employ the classic 'bag of words' approach, which treats text as a collection of individual words without considering their order or grammar. We adopt a simple proportional weighting scheme, where the importance or weight of each word is determined by its frequency relative to the total number of words in the text. The finance-specific positive and negative word lists developed by Loughran and McDonald (2011) are utilised for this purpose. The same textual sentiment analysis approach has been widely adopted in the finance literature (e.g., Liu and McConnell, 2013; Garcia, 2013; Huang et al., 2014; Solomon et al., 2014; Jiang et al., 2019).

The daily sentiment score for each *SA* stock is determined as follows:

$$Net\ Bulishness\_Stock = \frac{No.of\ positive\ words - No.of\ negative\ words}{Total\ No.of\ words} \quad (1)$$

The *stock-level* quarterly or monthly sentiment score is determined by averaging the daily scores within the respective period, excluding days when no *SA* articles were published. For the *quarterly* fund-level *Net Bullishness* score, the calculation employs a weighted average method, using the fund's current-quarter holdings as weights for each *SA* stock within the fund. This ensures that the score reflects the proportionate influence of each *SA* stock on the fund based on its holding size. Similarly, the *monthly* fund-level *Net Bullishness* score is computed using the same weighted average method, but with a slight variation. It utilises holding information from the previous quarter in conjunction with the current month's stock-level sentiment score (*Net Bullishness_Stock*). This calculation method ensures that the monthly score reflects the publicly disclosed holding information and timely captures the most recent *SA* coverage and sentiment.

For each fund, we acquire 'raw' *SA*-related variables at the fund level, which include the total number of *SA* stocks (*No. of SA Stocks*), total number of *SA* articles (*No. of SA Articles*), and the *Net Bullishness* score. Figure 1(a) to 1(c) each graphs the average monthly values of *No. of SA Stocks, No. of SA Articles,* and *Net Bullishness* across all funds included in the study. These figures indicate that the raw *SA*-related variables demonstrate distinct temporal trends. Specifically, *No. of SA Stocks* and *No. of SA Articles* both experienced a notable surge, reaching their zenith around 2011 and 2012, followed by a gradual decline thereafter. *Net Bullishness* shows a trend of being relatively negative prior to 2011 and again at the very end of the sample period, which coincides with the onset of the Covid-19 pandemic.

We then create 'adjusted' *SA*-related variables by calculating the percentage first difference of *No. of SA Stocks* and *No. of SA Articles,* along with the simple first difference of *Net Bullishness*. These adjusted variables are labelled as *%Diff_No. of SA Stocks, %Diff_No. of SA Articles*, and *Diff_Net Bullishness*, respectively, and collectively referred to as the fund's *SMR* feature or variables. The use of these adjusted, rather than raw, *SMR* variables in our regression analysis is further explained in Section 3.3.

Traditional media coverage and tone, serving as control variables in our regression analysis, are sourced from RavenPack. Consistent with the approach of Fang and Peress (2009), we focus on articles from the four most widely circulated newspapers in the US: *The New York Times, The Washington Post, USA Today*, and *The Wall Street Journal*. We download firm-specific articles along with their corresponding Composite Sentiment Score (CSS). As per the RavenPack manual, a positive (negative) CSS signifies a 'go long' ('go short') signal. For each fund in our study, we determine the fund-level total number of newspaper articles (*Newspaper_Articles*) and the article tone measure (*Newspaper_Tone*). These are calculated on a quarterly or monthly basis, employing a methodology similar to that used for creating the fund-level raw *SMR* variables.

3.3 Descriptive statistics

Table 1 provides an overview of the summary statistics for our final dataset, comprising 36,181 fund-quarter observations from actively managed U.S. domestic equity funds. We calculate the Total Net Assets (*TNA*s) at the fund level by summing the *TNA*s across all share classes of a fund. Return, common stock percentage, expense ratio, turnover, and flow are then derived as *TNA*-weighted averages across all these share classes. The fund's age is determined by counting the number of years since the inception of the oldest share class within the fund. These funds have an average *TNA* of approximately $4.56 billion, stretching across a broad spectrum from $5 million to nearly $897.62 billion. The average fund age stands at 26.8 years suggesting a predominance of long-established funds, though the actual ages span from as young as 2 years to 60 years. The funds have an average of 94.08% of their portfolios allocated to common stocks, in line with our sample selection criteria. The average expense ratio is 1.047%, albeit with considerable variability that ranges from 0.093% to 2.384%. The average turnover ratio is observed at 60.8% pointing to a moderate trading frequency among the funds.

[Insert Table 1 About Here]

Quarterly alphas are estimated from the Carhart (1997) four-factor model using daily fund returns within each quarter.[7] The *Net Alpha*, which reflects fund performance post-expenses, averages -0.292%. In contrast, the *Gross Alpha*, which does not consider fund expenses, averages at 2.847%, with certain funds reaching heights of nearly 30%. Fund flow is derived from *TNA* and quarterly raw returns,[8] and it is indicative of investor contributions and redemptions. The average fund flow is marginally negative at -0.02%, suggesting slight net capital outflows on average, yet the wide range from -0.923% to 6.541% reflects diverse investor behaviours across our sampled funds.

Panel B of Table 1 provides a detailed descriptive analysis of *SA* and traditional media characteristics for our *fund-quarter* sample. The substantial range in coverage and sentiment indicates a considerable influence of both *SA* and traditional media on these funds. On average, funds hold approximately 72 *SA* S*tocks* within a calendar quarter, with the number ranging from a minimum of 5 to a maximum of 1,581. This reflects the diverse level of *SA* coverage

[7] Converted to quarterly returns.
[8] Fund flow is calculated as $\frac{TNA_{i,t} - TNA_{i,t-1}*(1+ret_{i,t})}{TNA_{i,t-1}*(1+ret_{i,t})}$, where $ret_{i,t}$ is the *TNA*-weighted averages of quarterly raw returns (holding period returns computed from monthly returns) across all fund share classes

among the funds. The fraction of holdings in asset size covered by *SA* stocks averages 64.6%, with some funds featuring as little as 3% of their holdings discussed on *SA* in a quarter, while others achieve full coverage. The sample also shows an average of 782 *SA* articles coverage per fund-quarter, underscoring the active engagement of SMAs with the stocks in these funds. The average *Net Bullishness* is 16.8%, though there is significant variation across different funds. In examining *%Diff_No. of SA Stocks* and *%Diff_No. of SA Articles*, we observe an average of 3.1% and 8.3%, respectively. Such variability highlights the fluctuations in *SA* coverage and content volume across the 11-year period. Regarding traditional media coverage, the funds are covered by an average of approximately 1,540 newspaper articles per fund-quarter. *Newspaper_Tone*, which is the fund-level holding-weighted CSS, is on average 0.014. However, media tone for holdings of different funds varies widely and reflects the broad range of views in traditional media.

To examine the fund characteristics associated with holding *SA* stocks, we classify funds into quintile portfolios based on one of the fund-level raw *SMR* variable from the formation quarter. Quintile 1 (Low) comprises funds with the lowest value of the selected *SMR* variable, whereas Quintile 5 (High) includes funds with the highest value. These quintile portfolios are rebalanced every calendar quarter. Following this classification, we calculate the average values of *TNA*, *Age*, *Expense ratio (%)*, *Turnover*, fund betas[9], and *Net Alpha* for the funds within each quintile during the formation quarter. Next, we compute the mean of the time-series of these average values for each characteristic across each quintile. We also calculate the differences between the High (Quintile 5) and Low (Quintile 1) quintiles, providing the Newey-West (1987) adjusted t-statistics for these differences. The results are presented in Table 2.

[Insert Table 2 About Here]

[9] Fund betas are estimated from 60-month rolling window fund returns.

Panel A categorises mutual funds based on the *No. of SA Stocks* measure. The results indicate a clear trend: *TNA* progressively increases while *Expense ratio* decreases across quintiles as the number of *SA Stocks* grows. This suggests that larger funds with lower expense ratios tend to hold more *SA Stocks*. There is a stark difference in average *TNA* between the High and Low quintiles: funds with the most *SA Stocks* average $14,424.85 million in size, compared to just $1,231.55 million for those with the fewest. Additionally, funds with a higher number of *SA Stocks* generally have a longer history, lower turnover, and greater exposure to market risk and momentum factors, but less exposure to the SMB and HML factors. This profile suggests that funds inclined towards *SA Stocks* typically invest in large, growth-oriented, and winning stocks. This observation aligns with the findings by Dim (2023). Despite these differences in fund characteristics and risk profiles, the performance of funds in terms of *Net Alpha* does not significantly differ between the High and Low quintiles during the formation quarter. This suggests that while the inclination towards *SA Stocks* is associated with certain fund characteristics and risk exposures, it does not necessarily correlate with contemporaneous fund performance.

Panel B of Table 2 sorts the mutual funds by the *No. of SA Articles* measure. This sorting reveals even more consistent monotonic patterns across quintile portfolios compared to the sorting based on *No. of SA Stocks* in Panel A. Overall, the observed trends align closely with those in Panel A, with the exception of the momentum factor beta, where no significant difference is noted between funds experiencing high versus low *SA* article coverage for their holdings. Moreover, some trends are even more pronounced in this panel, particularly regarding *Age, Beta_SMB,* and *Beta_HML*.

Panel C of Table 2 organises funds based on their *Net Bullishness* scores. While there are significant differences in *TNA*, *Age*, and *Expense ratio* between the lowest and highest quintiles, the data do not show a clear monotonic trend across quintiles for these characteristics. This suggests a more complex or less direct relationship between a fund's size, age, expense ratio, and its holdings' overall bullish sentiment as expressed on *SA*, although funds characterised by greater bullish sentiment in their holdings tend to be the smallest, youngest, and have the highest expense ratios and turnover. Comparatively, the patterns for four factor loadings and *Net Alpha* are more consistent. Funds with higher *Net Bullishness* scores generally invest in smaller, growth-oriented, and winning stocks. This aligns with the literature examining the relationship between media tone and stock characteristics, such as the work by Liu and Han (2020). Interestingly, a greater *Net Bullishness* score correlates with better contemporaneous performance.

An insightful takeaway from Table 2 is the evident correlation between the raw *SMR* variables and a variety of fund characteristics, as well as the characteristics of their holdings. Essentially, a fund's raw *SMR* metrics are reflective of its diverse attributes or can be seen as composites of the fund and stock fundamentals. Notably, these raw *SMR* variables demonstrate a high degree of persistence, indicating they consistently reflect certain aspects of the funds over time. This revelation introduces a critical endogeneity concern when investigating the effects of *SMR* variables on fund flows and performance. The raw *SMR* variables may not effectively isolate the new information or sentiment-driven shocks intended to be captured from the count and content of *SA* articles. Consequently, to mitigate these concerns and to focus on the incremental aspects of social media coverage and sentiment, the study utilises the aforementioned 'adjusted' *SMR* variables in all regression analyses. These include the percentage first difference of *No. of SA Stocks* (*%Diff_No. of SA Stocks*), the percentage first difference of *No. of SA A*rticles (*%Diff_No. of SA Articles*), and the simple first difference of *Net Bullishness* (*Diff_Net Bullishness*). Employing these adjusted variables aims to refine the analysis and ensure that the *SMR* variables incorporated into regression models reflect shifts and sentiments more accurately, rather than just mirroring fund or stock characteristics. We repeat the analysis in Table 2 by using the adjusted *SMR* variables. These results are presented in Table A2 in the Appendix. Unlike the raw *SMR* variables, no distinct trends emerge in relation to fund or stock characteristics across the quintiles sorted by *%Diff_No. of SA Stocks*, *%Diff_No. of SA Articles*, or *Diff_Net Bullishness*. The only exception appears in the context of *Net Alpha* across the *Diff_Net Bullishness* quintiles. While the High quintile of funds in terms of *Diff_Net Bullishness* perform significantly better than the Low quintile contemporaneously, findings in the subsequent parts of the paper suggest that *Diff_Net Bullishness* does not simply mirror fund alphas, suggesting that the relationship between a fund's holding-weighted change in *SA* bullishness and its performance is more complex than a direct correlation. [10]

Table 3 presents a correlation matrix detailing the relationships between various mutual fund characteristics, their *SMR* variables, and traditional media variables. Intuitively, there is a relatively high positive correlation (0.648) between *No. of SA Stocks* and *No. of SA Articles*, suggesting that funds which hold more *SA Stocks* also attract more article coverage for its holdings. Meanwhile, the *Net Bullishness* score, indicative of the overall positive sentiment of holdings, exhibits only a minimal correlation with both *No. of SA Stocks* (0.018) and *No. of SA Articles* (0.022). This indicates that the sentiment attached to a fund's holdings is not tied to how often those holdings are mentioned or discussed in *SA*. Moreover, the correlation matrix reveals a strong connection between coverage in *SA* and traditional media, with a correlation

[10] Firstly, Table 4 and 5 corroborate existing literature by demonstrating a strong positive relationship between net flows and past fund performance. However, *Diff_Net Bullishness* is negatively correlated with flows in the subsequent quarter. Secondly, if *Diff_Net Bullishness* were simply a reflection of past performance, we would expect it to positively predict mid-to-long-term fund returns, despite the possibility of a short-term reversal (i.e., negative relationship). This expectation stems from the short-term reversal and mid-to-long-term momentum phenomenon documented in the asset pricing literature (see the results presented by Bali et al., 2016). However, the findings in Table 8 (specifically columns (7) and (9)) contradict this expectation. They reveal that the negative relationship between future fund performance and *Diff_Net Bullishness* persists into the mid-to-long term.

coefficient of 0.759 between article count in *SA* and those in traditional media platforms. This reflects a significant overlap in the attention stocks receive across different media channels.

Interestingly, while raw *SA* coverage variables moderately correlate with fund size and expense ratios (absolute values between 0.25 and 0.45), the adjusted *SMR* variables show almost no correlation with these fund characteristics. This observation lends further support to our decision to use adjusted *SMR* variables in regression analyses, affirming that they are more suited to capturing new information or sentimental shocks rather than merely reflecting inherent fund or stock fundamentals.

We do not find any surprising evidence of intra-fund correlations when compared to existing literature. For example, larger funds are usually associated with lower expense ratios, as indicated by a coefficient of -0.559. Additionally, the turnover ratio exhibits weak correlations with other variables, suggesting that trading activities within funds are relatively independent of factors such as age, size, and expense ratio.

[Insert Table 3 About Here]

4. Results and Discussions

This section presents the results and discussions pertaining to the testing of *Hypothesis 1* through *Hypothesis 5*. These hypotheses collectively examine the effects of social media on mutual fund investors' decision-making and the resultant economic outcomes. This section seeks to provide a comprehensive understanding of how social media coverage and sentiment influence the dynamics of mutual fund investments and the broader implications for market efficiency and investor behaviour.

4.1 SMR and fund flows

We start by examining whether a fund's *SMR* variables affect mutual fund flows. The regression models are specified as follows:

$$Flow_{i,t+1} = \lambda_0 + \lambda_1 \times SMR_{i,t} + \sum_{k=1}^{K} \lambda_k \times CONTROLS_{i,k,t} + \varepsilon_{i,t+1} \quad (2)$$

where the dependent variable is fund *i*'s percentage net flow in quarter *t*+1, and the primary explanatory variable is one of fund *i*'s adjusted *SMR* variables, *%Diff_No. of SA Stocks*, *%Diff_No. of SA Articles*, and *Diff_Net Bullishness* in quarter *t*. Control variables include the natural logarithm of the number of newspaper articles (*Newspaper_Article(ln)*) or *Newspaper_Tone*, the natural logarithm of fund *TNA* (*ln(TNA)*), the natural logarithm of fund age (*ln(Age)*), *Expense ratio*, *Turnover*, and fund betas in quarter *t*. To account for the nonlinear flow-performance relationship as suggested in the literature (e.g., Chevalier and Ellison, 1997; Sirri and Tufano, 1998), the variables $Low_{i,t}$, $Mid_{i,t}$, and $High_{i,t}$ are also employed as controls. Specifically, following Agarwal et al. (2022), for each quarter *t*, we assign all funds fractional ranks ($Rank_{i,t}$), according to their Carhart (1997) four-factor alpha, which are uniformly distributed between 0 (worst performance) and 1 (best performance). The variable $Low_{i,t}$ is defined as MIN (0.2, $Rank_{i,t}$), $Mid_{i,t}$ is defined as MIN (0.6, $Rank_{i,t}$ - $Low_{i,t}$), and $High_{i,t}$ is computed as $Rank_{i,t}$ - $Low_{i,t}$ - $Mid_{i,t}$. We control for all or a subset of the control variables in different model specifications. All regressions include fund and year fixed effects, and standard errors are clustered at the fund level.

[Insert Table 4 About Here]

Columns (1) and (4) of Table 4 indicate that after adjusting for fund characteristics and past performance, both *%Diff_No. of SA Stocks* and *%Diff_No. of SA Articles* are positively correlated with subsequent quarter fund flows. This suggests that increases in either the number of *SA* stocks or articles related to a fund's holdings are associated with significant increases in fund flows, although investors' reaction to changes in the number of *SA Stocks* is much smaller than that to changes in the number of *SA* articles. Specifically, a one standard deviation increase in *%Diff_No. of SA Stocks* corresponds to 0.60% (0.024*0.252) increase in fund flows, and a similar increase in *%Diff_No. of SA Articles* results in a 4.34% (0.086*0.505) increase in fund flows. Results are consistent when controlling for the extent of traditional media coverage or tone related to the stocks held by the fund (columns (2) and (5)) and when adding additional factor loadings of fund returns (columns (3) and (6)). These findings support H1, indicating that investors are attentive to *SA Stocks* and attracted to invest in funds that hold these stocks, especially those with greater total *SA* article count.

In contrast, an increase (decrease) of bullish *SA* sentiment, as measured by *Diff_Net Bullishness*, predicts lower (higher) fund flows in the subsequent quarter. In all three model specifications (columns 7, 8, and 9 in Table 4), the coefficients of *Diff_Net Bullishness* are negative and statistically significant at the 1% level. A one standard deviation increase (decrease) in *Diff_Net Bullishness* corresponds to a 2.56% (0.094*0.272) decline (increase) in fund flows.

Notably, the number of news articles on the traditional media related to stocks held by a fund predicts an increase in its fund flows in the following quarter, while the tone of those news articles predicts a decline in its fund flow. These results are comparable to the effects of corresponding *SMR* variables. Fang et al. (2014), however, find no significant relationship between traditional media coverage of fund holdings and fund flows.

The quarterly regressions might not fully capture the more immediate, short-term effects of *SMR* variables on fund flows. Investors might initially be drawn to funds with a higher *SA* feature due to recent coverage or sentiment but then adjust their positions before the quarter ends. To delve into these potential short-term dynamics, we conduct predictive regressions at a monthly frequency, maintaining the same model specifications as equation (2). This approach allows for a more granular view of how *SA* influences fund flows on a shorter time scale, potentially revealing investment patterns that quarterly data might not discern.

[Insert Table 5 About Here]

Table 5 reports monthly regression results. It shows that, similar to the quarterly regressions, both *%Diff_No. of SA Stocks* and *%Diff_No. of SA Articles* significantly predict higher fund flows in the subsequent month, though the statistical significance of their coefficients is lower compared to those in Table 4. Interestingly, traditional media coverage of stocks held by a fund does not predict its next-month flow, corroborating with the findings in Fang et al. (2014).

On the contrary, *Diff_Net Bullishness* predicts higher fund flows in the subsequent month, a finding that is in sharp contrast to the quarterly regression outcomes. Taken together, the monthly and quarterly regressions results might suggest that investors temporarily favour funds holding stocks associated with greater increase in bullish *SA* sentiment but tend to divest from them over the longer term, such as the next quarter. Alternatively, it could imply that sophisticated investors recognise the transient nature of social media sentiment towards individual stocks. Consequently, they might choose to capitalise on this short-lived sentiment by investing in funds exhibiting higher *SA* sentiment but strategically exit their positions before underlying fundamentals are revealed or before the prevailing sentiment reverses. Moreover, sophisticated investors might be inclined to invest in funds characterised by a decrease in bullishness, which is typically associated with poorer performance (see the last column of Panel C, Table 2). This investment choice could be driven by the expectation of price appreciation or a reversal in performance.

4.2 Difference-in-Differences analysis

In our baseline analysis of quarterly and monthly fund flows, we operate under the assumption that fund investors can view complete fund holdings subsequent to the funds' quarterly filings with the SEC. Adopting the methodology of Agarwal et al. (2022), we analyse the periods before and after the filing month to ascertain if the reactions of fund investors are particularly concentrated in the months immediately following the filing. This approach helps to mitigate endogeneity concerns, particularly the possibility that the predictive power of the *SMR* variables on fund flows could be influenced by omitted factors affecting both *SMR* variables and fund flows. To this end, we employ two methodologies: an event study and a DID analysis.

The *monthly* indicator variable I(TREAT$_{i,m}$) is set to be 1 if fund *i*'s *%Diff_No. of SA Stocks*, *%Diff_No. of SA Articles*, or *Diff_Net Bullishness* measurement ranks in the top 30% in the calendar quarter, and 0 if the measurement is in the bottom 30%. The monthly binary variables I(Lead_Lag $_{i,M+n}$) (where n = -1, 0, 1) are assigned a value of 1 to signify that the specific month $M+n$ is n periods away from the SEC filing month M at the end of any quarter (i.e., March, June, September, and December), and 0 otherwise. Therefore, in months M-1 and M, the fund holdings and the corresponding fund-level *SA* coverage and sentiment for that quarter are not yet fully known to investors. In contrast, by month $M+1$, this information would have been observed by investors. Filing in month M is considered as an 'event'. In Panel A of Table 6, an event study is conducted by employing the interaction terms between I(Lead_Lag $_{i,M-1}$) and the first lead of I(TREAT$_{i,m}$), I(Lead_Lag $_{i,M}$) and I(TREAT$_{i,m}$), and I(Lead_Lag $_{i,M+1}$) and the first lag of I(TREAT$_{i,m}$) as explanatory variables. [11] The dependent variable in this analysis is the fund flows in month $m+1$. The regression models also include a complete list of control variables employed in the baseline analysis, in addition to fund and time fixed effects. Standard errors are clustered at the fund level.

[11] The first lead and first lag of I(TREAT$_{i,m}$) are employed to refer to the treated group on the event of interest.

Columns (1) to (3) in Panel A of Table 6 show that for the pre-filing and filing months, the coefficients of the interaction terms are generally not significant at the 5% level, except for the pre-filing period in column (1). This suggests that in the subsequent month (month *m+1*) of these periods, investors' reactions to fund-level *SA* coverage and sentiment do not significantly differ between the treated and control groups in terms of *%Diff_No. of SA Articles* and *Diff_Net Bullishness*. An exception is noted with the *%Diff_No. of SA Stocks* variable. There are differing flow patterns between the treated and control groups before investors can fully observe the portfolio composition and the *SA* coverage and sentiment of fund holdings for that quarter. This may indicate some endogeneity, suggesting that *%Diff_No. of SA Stocks* might not be an ideal proxy of 'clean' fund-level *SA* attention. It is also possible that some funds disclose part of their holdings before the mandatory filing dates, and investors who actively follow *SA* might have reacted during the pre-filing and filing periods. Nevertheless, it is noteworthy that the reaction difference between the treated and control groups concerning *%Diff_No. of SA Stocks* is considerably more pronounced in the post-filing month compared to the pre-filing month (column (1)). The coefficients of the interaction terms for the post-filing period in columns (2) (for *%Diff_No. of SA Articles*) and (3) (for *Diff_Net Bullishenss*) are also positive and significant at the 1% level, highlighting a distinct shift in investor behaviour following the disclosure of portfolio composition. The positive coefficient for the post-filing months, as presented in column (3) of Panel A, resonates with the monthly baseline regression results shown in columns (7) to (9) of Table 5. Figure 2(a) to 2(c) graphically represent results presented in Panel A, illustrating that the 95% confidence intervals for the coefficients of pre-filing and filing months encompass zero on the vertical axis (i.e., coefficients are not significantly different from zero), except for the pre-filing period in column (1). In contrast, for the post-filing months, these intervals are well above zero.

[Insert Table 6 About Here]

Panel B of Table 6 presents the results from a classic DID analysis. In this analysis, the monthly indicator variable I(POST$_{i,m}$) is designated a value of 1 for the first month following a filing month *M*, and 0 otherwise. The primary variable of interest in this context is the interaction between the first lag of I(TREAT$_{i,m}$) and I(POST$_{i,m}$). All control variables and fixed effects as previously mentioned are incorporated into the regression models. Results show that the coefficients of the interaction terms in all regressions are positively significant at the 1% level. This indicates that fund flows are considerably higher for the treated groups, characterized by high *SA* coverage or bullish sentiment, compared to the control group in month *M*+2. This month follows the period when investors have had the opportunity to observe complete fund holdings in month *M*+1. The evidence presented in Table 6 strongly supports the hypotheses that fund-level *SA* coverage and sentiment (the first difference) exert an impact on the investment decision-making of fund investors, particularly in the period following the disclosure of fund holdings (*Hypothesis 1* and *Hypothesis 2*).

4.3 Investor recognition and attention

The literature finds that *SA* predominantly caters to the information needs of individual investors, rather than institutional ones (Farrell et al., 2022; Gomez et al., 2022; Dim, 2023). Given that attention is a limited resource for individual investors, who face a significant search problem when buying a stock (Barber and Odean, 2008), this issue extends to their decisions regarding mutual fund investments. We posit that funds which are more familiar to investors naturally garner more attention, and higher aggregated *SA* coverage at the fund level also act as an attention catalyst for mutual fund investors. Consequently, as stated in *Hypothesis 3*, we expect that the impact of *SMR* variables on fund flows will be more pronounced in funds that have higher investor recognition or those that hold a larger proportion of stocks with higher visibility on *SA*.

We anticipate that larger funds, those with a longer history and lower expense ratios (which are typically also larger) are likely to have higher investor recognition. Additionally, since *SA* primarily covers large[12], growth-oriented, and high-priced stocks (Dim, 2023), we expect the reaction of fund flows to *SMR* variables to differ among funds with varying exposure to the SMB and HML factors, as well as among funds with different holding-weighted average stock prices. Further, we are interested in examining the influence of funds' holding-weighted MAX measure (*MAX*), which represents the maximum daily returns of a stock within a month and averaged over a quarter. This examination is driven by two motivations. Firstly, we aim to determine if a fund's *SA* stock holding exhibits characteristics similar to its lottery-like stock holdings, as explored by Agawal et al. (2022). Secondly, considering that *SA* typically covers stocks with low idiosyncratic volatility (Dim, 2023), we posit that the impact of *SMR* variables on fund flows will be more pronounced among funds characterised by a lower holding-weighted *MAX*, since lottery stocks (characterised by high *MAX*) exhibit high idiosyncratic volatility (Kumar, 2009; Bali et al., 2021). These examinations will help elucidate the nuances of how different fund characteristics and the nature of stock holdings interact with *SA* coverage and sentiment to influence investor behaviour.

For each quarter, we classify funds into three groups based on the 30th and 70th percentiles of certain fund or fund holding characteristics, ordered from low to high (Group 1 to Group 3): *TNA, Age, Expense Ratio, Beta_SMB, Beta_HML, Price* (holding-weighted), and *MAX* (holding-weighted). The fund portfolios are rebalanced every quarter. We then apply the same regression model used in Table 4 to each group across the entire sample period, incorporating the full set of control variables.

[12] Large stocks have the advantage of more data available on its economic activities and longer history compared to small stocks, making them more cost-effective targets for analysts (Begenau et al 2018; Veldkamp and Chung, forthcoming).

Table 7 presents the coefficients of *%Diff_No. of SA Stocks* (Panel A), *%Diff_No. of SA Articles* (Panel B), and *Diff_Net Bullishness* (Panel C), along with the corresponding t-statistics in the parentheses. By examining the magnitude of these coefficients and their t-statistics, we can discern how the responsiveness of fund flows to *SMR* variables varies across different fund or fund holding characteristics. Additionally, we conduct the Chow (1960) test to assess the equality of the coefficients between Group 1 and Group 3, with the p-values of the F-statistics reported. For the sake of brevity, coefficients and t-statistics for all other independent variables are omitted from the report.

[Insert Table 7 About Here]

Table 7's findings largely support *Hypothesis 3* when considering the reactions of fund flows to *%Diff_No. of SA Articles* and *Diff_Net Bullishness*. Specifically, the coefficients are typically largest in magnitude and more significant (reflected by higher absolute value of the t-statistics) for funds with the largest *TNA*, oldest age, lowest expense ratio, lowest exposure to the SMB factor (indicating a preference for larger stocks), highest holding-weighted price, and lowest holding-weighted *MAX*. This suggests that the reaction to these *SMR* variables is most pronounced among funds with higher investor recognition or those holding stocks more frequently featured on *SA*. The prediction regarding funds holdings in growth-oriented stocks (*Beta_HML*) is confirmed for *Diff_Net Bullishness* but not as much for *%Diff_No. of SA Articles*. The findings with respect to MAX also suggest that the *SMR* feature is inherently different from the lottery-stock characteristics.

Conversely, when considering *%Diff_No. of SA Stocks*, the support for *Hypothesis 3* is more limited, primarily observed through metrics like *TNA*, *Beta_SMB*, and *MAX*. As discussed in Section 4.2, *%Diff_No. of SA Stocks* may not purely reflect direct *SA* attention. Coupled with the observation that the impact of *%Diff_No. of SA Stocks* on fund flows is considerably smaller than that of *%Diff_No. of SA Articles* (as reported in Table 4), it can be inferred that investors are more responsive to changes in the total count of *SA* articles rather than the number of *SA Stocks* held by a fund.

4.4 SA attention and bullishness and future fund performance

Table A2 in the Appendix has established that *%Diff_No. of SA Stocks* and *%Diff_No. of SA Articles* signify incremental *SA* attention at the fund level, distinct and uncorrelated with fund or fund holding characteristics. Further insights from Tables 6 and 7 suggest that *%Diff_No. of SA Articles* serves as a more accurate proxy for 'clean' investor attention that leads to investment reactions. Additionally, it is evident that investors respond to incremental *SA* bullishness at the fund level, as indicated by *Diff_Net Bullishness*. However, the wisdom and value of acting upon this incremental *SA* attention and bullishness for fund investors remain to be assessed.

To address this, we examine future fund performance, measured by cumulative net alphas across different investment horizons, against the *SMR* variable and the same set of control variables used in the fund flow regressions, along with the current-quarter *Net Alpha*. While acknowledging that the *SMR* feature of a fund might undergo significant changes in future periods, the purpose of this analysis is to determine whether investor reactions to *SMR* variables are, on the whole, justifiable and beneficial from a performance perspective. This is crucial for understanding whether the attention and sentiment driven by *SMR* content contribute to tangible investment benefits, or if they simply reflect transient market trends without long-term performance advantages. The analysis covers both short-term and mid-to-long-term fund performance, specifically over the window periods of [t+1, t+2], [t+3, t+4], and [t+1, t+4], where t represents the current quarter. Alphas for all investment periods are annualised and normalised, which ensures that the magnitude of their coefficients is directly comparable across different time spans. All regressions control for the fund and year fixed effects. Standard errors are clustered at the fund level. The regression results are presented in Table 8.

[Insert Table 8 About Here]

Columns (1) and (4) of Table 8 reveal that the future two-quarter performance of mutual funds is positively related to both *%Diff_No. of SA Stocks* and *%Diff_No. of SA Articles*. This positive relationship, however, reverses during the period from quarter *t*+3 to quarter *t*+4 (columns (2) and (5)), leading to an insignificant effect of these two incremental *SA* attention measures on the one-year fund performance (columns (3) and (6)). This pattern suggests that investors who respond to *SMR* by investing in funds with higher incremental *SA* attention may experience positive abnormal returns in the short term, but this outperformance tends to disappear if the investment is held for an additional two quarters. This evidence supports *Hypothesis 4*.

On the other hand, the coefficient of *Diff_Net Bullishness* on the next two-quarter cumulative alphas is significantly negative, as shown in column (7). This indicates that a higher fund-level incremental *SA* bullishness is associated with worse short-term fund performance. No significant reversal of this effect is observed in quarter *t*+3 to quarter *t*+4 (column (8)), resulting in the one-year fund performance being significantly negatively correlated with current-quarter incremental *SA* bullishness (column (9)), a result consistent with *Hypothesis 5b*. This finding aligns with observations from Tables 4 and 5 that while investors may initially respond to higher *SA* bullishness by investing more immediately after the complete fund holding disclosure, on average, they tend to divest from funds with high *Diff_Net Bullishness* in the subsequent quarter. Therefore, fund investors who consult *SA* articles generally adopt a contrarian stance to sentiment changes, and their negative response to *SMR* bullishness appears to be a judicious decision.

5. Conclusions

By analysing 36,181 fund-quarter observations for 1,141 actively managed US equity mutual funds between January 2009 and March 2020, this research delineates that fund investors' reactions to social media research are twofold. Firstly, an increase in incremental *SA* coverage, especially a notable rise in the number of *SA* articles related to fund holdings, captures investor attention and prompts increased investment in such funds. This effect is more pronounced in the post-filing months and in funds that have higher investor recognition or those that hold a larger proportion of stocks with higher visibility on *SA*, emphasising the role of investor attention in this process. However, this strategy does not yield rewards in the mid-to-long term, suggesting that while *SA* coverage initially attracts investors, it may not lead to sustained performance benefits.

Secondly, an increase (or decrease) in holding-weighted *SA* bullishness is linked to lower (or higher) fund flows in the subsequent quarter, and this investment strategy proves to be prudent over the long term. Specifically, a rise in bullishness predicts underperformance of funds, whereas a decline forecasts overperformance. This suggests that heightened bullishness or sentiment expressed in *SA* articles is informative and serves the interests of contrarian investors, who seek to profit from subsequent market adjustments. This finding aligns with conclusions drawn by Farrell et al. (2022), who show evidence that investors have the skills to glean valuable information from *SA* rather than trade directly on article sentiment.

The fact that fund flow reactions to changes in sentiment are more pronounced among funds with higher investor recognition or those holding a larger proportion of stocks frequently featured on *SA* does not contradict the idea that *SA* sentiment is informative. This could be because sophisticated investors recognise the transient nature of sentiment and seek to capitalise on short-term market deviations, which are often amplified by increased investor attention. In other words, while these investors are attentive to the sentiment shifts, they may also be strategically positioning themselves to benefit from the temporary nature of such sentiment, especially when it is magnified in funds that are generally more familiar to investors or hold stocks that are more visible on *SA*.

In summary, while the coverage of fund holdings by social media analysts may not be reliably informative in the long term, it plays a significant role in directing investor attention in the short term. Conversely, the sentiment level expressed by social media analysts regarding fund holdings provides valuable information, particularly for contrarian investment strategies. Therefore, this evidence supports the *investor attention hypothesis* in the context of coverage and the *information hypothesis* concerning sentiment in the role of social media research for mutual fund investment decision making.

References

Agarwal, V., Jiang, L., & Wen, Q. (2022). Why do mutual funds hold lottery stocks?. *Journal of Financial and Quantitative Analysis*, 57(3), 825-856.

Akbas, F., & Genc, E. (2020). Do mutual fund investors overweight the probability of extreme payoffs in the return distribution?. *Journal of Financial and Quantitative Analysis*, 2020, 55(1): 223-261.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance*, 61(4), 1645-1680.

Bali, T. G., Engle, R. F., & Murray, S. (2016). Empirical asset pricing: The cross section of stock returns. John Wiley & Sons.

Bali, T. G., Hirshleifer, D., Peng, L., & Tang, Y. (2021). Attention, social interaction, and investor attraction to lottery stocks (No. w29543). National Bureau of Economic Research.

Barber, B. M., Huang, X., Odean, T., & Schwarz, C. (2022). Attention-induced trading and returns: Evidence from Robinhood users. *Journal of Finance*, 77(6), 3141-3190.

Barber, B. M., & Odean, T. (2008) All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, 21, 785–818.

Bartov, E., Faurel, L., Mohanram, P. (2018). Can Twitter help predict firm-level earnings and stock returns? *Accounting Review*, 93, 25–57.

Begenau, J., Farboodi, M., & Veldkamp, L. (2018). Big data in finance and the growth of large firms. *Journal of Monetary Economics*, 97, 71-87.

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

Campbell, J. L., DeAngelis, M. D., & Moon, J. R. (2019). Skin in the game: Personal stock holdings and investors' response to stock analysis on social media. *Review of Accounting Studies*, 24, 731-779.

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

Chawla, N., Da, Z., Xu, J. & Ye, M. (2021). Information diffusion on social media: does it affect trading, return, and liquidity? Available at SSRN: https://ssrn.com/abstract=2935138

Chen, H., De, P., Hu, J., & Hwang, B. H. (2014). Wisdom of the crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies*, 27, 1367-1403.

Chen, H., & Hwang, B. H. (2022). Listening in on investors' thoughts and conversations. *Journal of Financial Economics*, 145(2), 426-444.

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

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society*, 591-605.

Cookson, J. A., Lu, R., Mullins, W., & Niessner, M. (2023) The social signal. Working paper, available at SSRN: https://ssrn.com/abstract=4241505

Cooper, M., Gulen, J., H., & Rau, P. R. (2005). Changing names with style: Mutual fund name changes and their effects on fund flows. *Journal of Finance*, 60: 2825–58.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, 98(4), 703-738.

Dim, C. (2023), Social media analysts' skills: Insights from text-implied beliefs. Available at SSRN: https://ssrn.com/abstract=3813252

Drake, M. S., Moon Jr, J. R., Twedt, B. J., & Warren, J. D. (2023). Social media analysts and sell-side analyst research. *Review of Accounting Studies*, 28(2), 385-420.

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *Journal of Finance*, 64, 2023–2052.

Fang, L., Peress, J., & Zheng, L. (2014). Does media coverage of stocks affect mutual funds' trading and performance? *Review of Financial Studies*, 27, 3441–3466.

Farrell, M., Green, T. C., Jame, R., & Markov, S. (2022). The democratization of investment research and the informativeness of retail investor trading. *Journal of Financial Economics*, 145(2), 616-641.

Garcia, D. (2013). Sentiment during recessions. *Journal of Finance*, 68, 1267–1300.

Gil-Bazo, J., & Imbet, J. F. (2022) Tweeting for money: social media and mutual fund flows. Université Paris-Dauphine Research Paper No. 3719169, available at SSRN 3719169.

Gomez, E., Heflin, F., Moon, J., & Warren, J. (2022). Financial analysis on social media and disclosure processing costs: Evidence from Seeking Alpha. Georgia Tech Scheller College of Business Research Paper, (18-45).

Gu, C. & A. Kurov (2020). Informational role of social media: Evidence from Twitter sentiment. *Journal of Banking & Finance*, 121, 105969.

Han, B., Hirshleifer, D., & Walden, J. (2022). Social transmission bias and investor behavior. *Journal of Financial and Quantitative Analysis*, 57(1), 390-412.

Hirshleifer, D. (2020). Presidential address: Social transmission bias in economics and finance. *Journal of Finance*, 75(4), 1779-1831.

Hirshleifer, D. A., Peng, L., & Wang, Q. (2023). News diffusion in social networks and stock market reactions. Working paper. Available at SSRN: https://ssrn.com/abstract=3824022.

Huang, A. H., Zang, A. Y., & Zheng, R. (2014). Evidence on the information content of text in analyst reports. *Accounting Review*, 89(6), 2151-2180.

Jia, W., Redigolo, G., Shu, S., & Zhao, J. (2020). Can social media distort price discovery? Evidence from merger rumors. *Journal of Accounting and Economics*, 70(1), 101334.

Jain, P. C., & Wu, J. S. (2000). Truth to mutual fund advertising: Evidence on future performance and fund flows. *Journal of Finance*, 55: 937–58.

Jiang, G. J., & Yüksel, H. Z. (2019). Sentimental mutual fund flows. *Financial Review*, 54(4): 709–738.

Jiao, P., Veiga, A., & Walther, A. (2020). Social media, news media and the stock market. *Journal of Economic Behavior & Organization*, 176, 63-90.

Kahneman, D. (1973). Attention and effort (Vol. 1063, pp. 218-226). Englewood Cliffs, NJ: Prentice-Hall.

Kamstra, M. J., Kramer, L. A., Levi, M. D., & Wermers, R. (2017). Seasonal asset allocation: evidence from mutual fund flows. *Journal of Financial and Quantitative Analysis*, 52, 71–109.

Kaniel, R., & Parham, R. (2017). WSJ Category Kings –The impact of media attention on consumer and mutual fund investment decisions. *Journal of Financial Economics*, 123, 337-356.

Kumar, A. (2009). Who gambles in the stock market?. *Journal of Finance*, 64(4), 1889-1933.

Liu, S., Gaskell, P., & McGroarty, F. (2022). Where and about what? Price relevant narratives depend on topic and media type. *Economics Letters*, 213, 110363.

Liu, S. & Han, J. (2020). Media tone and expected stock returns. *International Review of Financial Analysis*, 70, 101522.

Lou, D. (2012). A flow-based explanation for return predictability. *Review of Financial Studies*, 25, 3457–3489.

Lou, D. (2014). Attracting investor attention through advertising. *Review of Financial Studies*, 27(6), 1797-1829.

Peng, L., & Xiong, W. (2006) Investor attention, overconfidence and category learning. Journal of Financial Economics, 80(3): 563-602.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? *Journal of Finance*, 66, 35–65.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708.

Rakowski, D., Shirley, S., & Stark, J. R. (2021) Twitter activity, investor attention, and the diffusion of information. *Financial Management*, 50: 3– 46.

Renault, T. (2017). Intraday online investor sentiment and return patterns in the US stock market. *Journal of Banking & Finance*, 84, 25-40.

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

Solomon, D. H., Soltes, E. F., & Sosyura, D. (2014). Winners in the spotlight: media coverage of fund holdings as a driver of flows. *Journal of Financial Economics*, 113, 53–72.

Sprenger, T.O., Tumasjan, A., Sandner, P.G. & Welpe, I.M. (2014a). Tweets and trades: the information content of stock microblogs. *European Financial Management*, 20: 926-957.

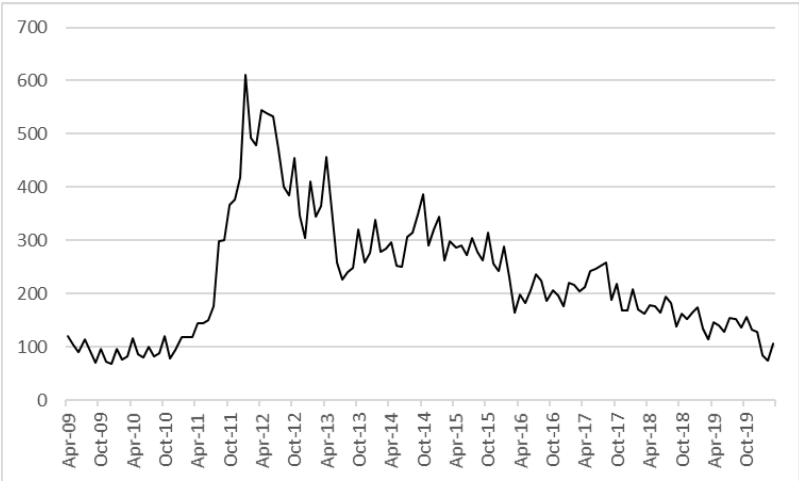
Sprenger, T.O., Sandner, P.G., Tumasjan, A. & Welpe, I.M. (2014b). News or noise? Using Twitter to identify and understand company-specific news flow. *Journal of Business Finance & Accounting*, 41: 791-830.

Veldkamp, L., & Chung, C. (forthcoming). Data and the aggregate economy. *Journal of Economic Literature*.

Figure 1

(a) Average Monthly *No. of SA Stocks*

(b) Average Monthly *No. of SA Articles*



(c) Average Monthly *Net Bullishness*

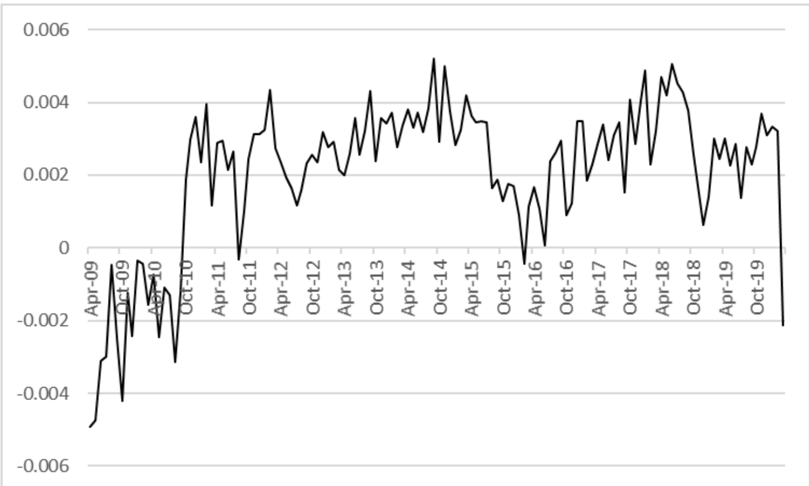
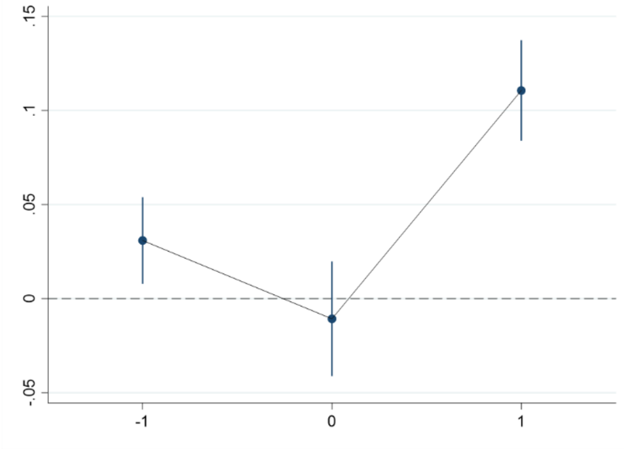
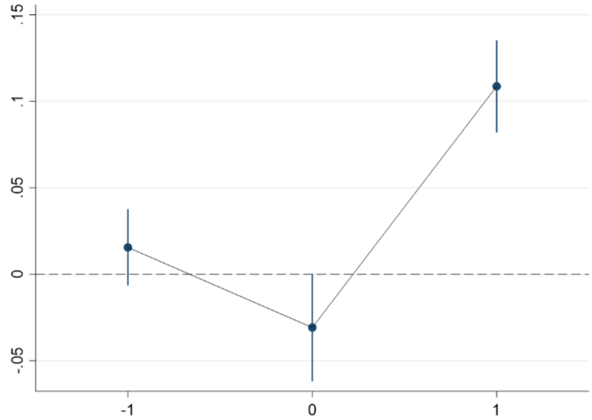


Figure 2

(a) %Diff_No. of SA Stocks



(b) *%Diff_No. of SA Articles*



(c) *Diff_Net Bullishness*

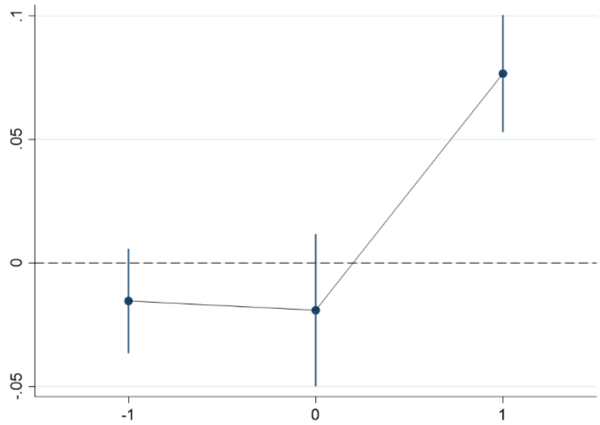


Table 1: Descriptive Summary Statistics

Notes: This table reports summary statistics of the mutual fund characteristics and the fund-level *SA* and traditional media variables. Definitions for all variables are listed in Table A1 in the Appendix.

Variable	Obs	Mean	Std. Dev.	Min	Median	Max
Panel A: Fund Characteristics (by Fund-Quarter)						
TNA ($ Million)	36,181	4563.387	22405.610	5.000	730.600	897614.500
Age (Year)	36,181	26.813	11.270	2.000	24.000	60.000
Common Stock (%)	36,181	94.078	4.474	80.001	95.120	100.030
Expense ratio (%)	36,181	1.047	0.383	0.093	1.041	2.384
Turnover	36,181	0.608	0.534	0.020	0.470	3.080
Net Alpha (%)	36,181	-0.292	2.754	-36.895	-0.221	24.576
Gross Alpha (%)	36,120	2.847	2.941	-32.883	2.734	29.946
Fund Flow (quarterly)	34,811	-0.020	0.145	-0.923	-0.027	6.541
Panel B: Seeking Alpha (SA) and Traditional Media characteristics (by Fund-Quarter)						
No. of SA Stocks	36,181	72.592	100.554	5.000	43.000	1581.000
Fraction of holdings covered by SA	36,181	0.646	0.213	0.030	0.707	1.000
No. of SA Articles	36,181	782.420	1083.908	5.000	388.000	12207.000
Net Bullishness (holding-weighted) (%)	36,181	0.168	0.339	-4.023	0.222	2.228
%Diff_No. of SA Stocks	34,618	0.031	0.252	-0.941	0.000	9.554
%Diff_No. of SA Articles	34,618	0.083	0.505	-0.977	0.003	23.875
Diff_Net Bullishness (%)	34,619	0.018	0.272	-4.312	0.010	3.884
Newspaper Articles	36,181	1540.713	2004.595	0.000	718.000	13960.000
Newspaper_Tone (holding-weighted)	36,181	0.014	0.029	-0.520	0.014	0.560

Table 2: Mutual Fund Quintile Portfolios Sorted on Raw SMR Variables

Notes: Funds are classified into quintile portfolios based on *No. of SA Stocks*, *No. of SA Articles*, or *Net_Bullishness* from the formation quarter. Quintile 1 (Low) comprises funds with the lowest value of the selected variable, whereas Quintile 5 (High) includes funds with the highest value. These quintile portfolios are rebalanced every calendar quarter. Following this classification, we calculate the average values of *TNA, Age, Expense ratio (%), Turnover*, fund betas, and *Net Alpha* for the funds within each quintile during the formation quarter. Next, we compute and report the mean of the time-series of these average values for each characteristic across each quintile. We also calculate the differences between the High (Quintile 5) and Low (Quintile 1) quintiles, providing the Newey-West (1987) adjusted t-statistics for these differences.

				Panel A: ***No. of SA Stocks***					
	TNA ($ Million)	***Age (Year)***	***Expense ratio (%)***	***Turnover***	***Beta_Mkt***	***Beta_SMB***	***Beta_HML***	***Beta_UMD***	***Net Alpha (%)***
Low	1231.550	25.564	1.193	0.580	0.951	0.255	0.083	-0.012	-0.222
2	1525.025	26.176	1.155	0.603	0.973	0.209	-0.006	0.012	-0.307
3	2514.249	27.892	1.090	0.628	0.995	0.136	-0.043	0.024	-0.420
4	4161.566	28.628	1.019	0.662	0.996	0.097	-0.033	0.032	-0.319
High	14424.850	27.510	0.736	0.528	0.990	0.142	0.029	0.013	-0.219
High-Low	13193.300	1.946	-0.457	-0.052	0.038	-0.113	-0.053	0.024	0.004
t-stat	***6.87***	***5.55***	***-43.87***	***-4.44***	***8.76***	***-3.97***	***-3.68***	**2.58**	0.03
				Panel B: ***No. of SA Articles***					
	TNA ($ Million)	***Age (Year)***	***Expense ratio (%)***	***Turnover***	***Beta_Mkt***	***Beta_SMB***	***Beta_HML***	***Beta_UMD***	***Net Alpha (%)***
Low	1230.320	24.174	1.219	0.594	0.958	0.509	0.112	0.015	-0.186
2	2060.773	25.552	1.100	0.686	0.977	0.305	0.016	0.027	-0.284
3	2498.541	27.604	1.087	0.609	0.986	0.106	-0.032	-0.002	-0.372
4	3630.279	28.946	1.019	0.589	0.995	-0.023	-0.055	0.023	-0.345
High	14284.090	29.420	0.775	0.522	0.988	-0.055	-0.008	0.004	-0.299
High-Low	13053.770	5.246	-0.445	-0.073	0.030	-0.564	-0.121	-0.011	-0.113
t-stat	***6.88***	***45.46***	***-101.27***	***-6.34***	***4.58***	***-49.77***	***-6.40***	-0.85	-0.93
				Panel C: ***Net Bullishness***					
	TNA ($ Million)	***Age (Year)***	***Expense ratio (%)***	***Turnover***	***Beta_Mkt***	***Beta_SMB***	***Beta_HML***	***Beta_UMD***	***Net Alpha (%)***
Low	3341.558	27.338	1.066	0.590	0.975	0.125	0.154	-0.053	-0.413
2	9662.673	27.443	0.919	0.530	0.977	0.065	0.019	-0.001	-0.322
3	5306.341	27.889	1.004	0.591	0.985	0.111	-0.034	0.020	-0.313
4	3579.733	27.358	1.061	0.628	0.987	0.195	-0.060	0.045	-0.284
High	1756.339	25.636	1.153	0.660	0.980	0.351	-0.045	0.056	-0.152
High-Low	-1585.219	-1.703	0.087	0.070	0.004	0.226	-0.198	0.109	0.261
t-stat	***-6.96***	***-5.60***	***9.30***	***3.05***	0.45	***7.84***	***-8.52***	***7.48***	**2.32**

Table 3: Correlation Coefficient Matrix

Notes: This table reports the correlation coefficients of fund-quarter key variables used in the empirical analysis.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) *No. of SA Stocks*	1.000													
(2) *No. of SA Articles*	0.648	1.000												
(3) *Net Bullishness Score*	0.018	0.022	1.000											
(4) *%Diff_No. of SA Stocks*	0.031	-0.012	-0.023	1.000										
(5) *%Diff_No. of SA Articles*	0.003	0.020	-0.022	0.677	1.000									
(6) *Diff_Net Bullishness (%)*	-0.011	-0.038	0.335	0.016	-0.028	1.000								
(7) *Newspaper Articles*	0.538	0.759	0.015	-0.036	-0.056	-0.033	1.000							
(8) *Newspaper_Tone*	-0.013	-0.004	0.177	-0.041	-0.018	0.019	-0.027	1.000						
(9) *ln(TNA)*	0.269	0.221	0.024	-0.036	-0.043	-0.009	0.267	-0.005	1.000					
(10) *ln(Age)*	-0.006	0.069	0.122	-0.033	-0.041	-0.044	0.173	-0.016	0.329	1.000				
(11) *Expense ratio* (%)	-0.450	-0.362	-0.003	0.030	0.035	0.016	-0.375	0.023	-0.559	-0.206	1.000			
(12) *Turnover*	-0.143	-0.132	-0.064	0.041	0.046	0.023	-0.136	0.034	-0.255	-0.146	0.261	1.000		
(13) *Net Alpha* (%)	0.017	-0.001	0.034	0.004	-0.024	0.079	0.002	0.014	0.037	0.002	-0.047	-0.066	1.000	
(14) *Gross Alpha* (%)	-0.168	-0.149	0.030	0.016	-0.008	0.080	-0.151	0.022	-0.194	-0.082	0.365	0.045	0.913	1.000

Table 4: Social Media Research and Quarterly Fund Flows

Notes: This table examines whether a fund's *SMR* variables affect mutual fund flows. The dependent variable is fund *i*'s percentage net flow in quarter *t+1*, and the primary explanatory variable is one of fund *i*'s adjusted *SMR* variables, *%Diff_No. of SA Stocks*, *%Diff_No. of SA Articles*, and *Diff_Net Bullishness* in quarter *t*. The definitions of the control variables are listed in Table A1 in the Appendix. All regressions control for the fund and year fixed effects. Coefficients are marked with *, **, or *** for the significance level of 10%, 5%, and 1%, respectively. Standard errors are clustered at the fund level and reported in parentheses.

	Dependent Var: *Flow* $_{i, t+1}$								
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
%Diff_No. of SA Stocks $_{i, t}$	0.024***	0.018***	0.017***						
	(0.006)	(0.006)	(0.006)						
%Diff_No. of SA Articles $_{i, t}$				0.086***	0.079***	0.078***			
				(0.007)	(0.007)	(0.007)			
Diff_Net Bullishness $_{i, t}$							-0.094***	-0.095***	-0.095***
							(0.006)	(0.006)	(0.006)
Newspaper_Article (ln) $_{i, t}$		0.219***	0.236***		0.204***	0.221***			
		(0.028)	(0.028)		(0.028)	(0.027)			
Newspaper_Tone $_{i, t}$								-0.017***	-0.018***
								(0.006)	(0.006)
Low $_{i, t}$		0.057	0.026		0.053	0.019		0.118	0.092
		(0.164)	(0.163)		(0.164)	(0.163)		(0.164)	(0.164)
Mid $_{i, t}$		0.006	0.001		0.005	-0.001		0.024	0.019
		(0.032)	(0.032)		(0.032)	(0.032)		(0.031)	(0.032)
High $_{i, t}$		0.506***	0.444**		0.498***	0.434**		0.506***	0.451***
		(0.172)	(0.174)		(0.172)	(0.173)		(0.171)	(0.173)
ln(TNA) $_{i, t}$	-0.534***	-0.543***	-0.530***	-0.527***	-0.535***	-0.524***	-0.524***	-0.519***	-0.509***
	(0.037)	(0.037)	(0.037)	(0.036)	(0.037)	(0.037)	(0.036)	(0.036)	(0.036)
ln(Age) $_{i, t}$	-0.222***	-0.185**	-0.197**	-0.225***	-0.190**	-0.202**	-0.222***	-0.218***	-0.231***

	(0.084)	(0.085)	(0.086)	(0.084)	(0.085)	(0.085)	(0.084)	(0.084)	(0.084)
Expense ratio $_{i,t}$	-0.277***	-0.281***	-0.280***	-0.275***	-0.280***	-0.279***	-0.274***	-0.274***	-0.273***
	(0.045)	(0.046)	(0.046)	(0.045)	(0.046)	(0.046)	(0.045)	(0.045)	(0.045)
Turnover $_{i,t}$	-0.066***	-0.064***	-0.066***	-0.068***	-0.066***	-0.068***	-0.065***	-0.064***	-0.064***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Beta_Mkt $_{i,t}$			-0.028***			-0.029***			-0.026***
			(0.010)			(0.009)			(0.010)
Beta_SMB $_{i,t}$			0.113***			0.108***			0.089***
			(0.019)			(0.019)			(0.018)
Beta_HML $_{i,t}$			-0.027**			-0.024*			-0.028**
			(0.012)			(0.012)			(0.012)
Beta_UMD $_{i,t}$			0.005			0.006			0.002
			(0.008)			(0.007)			(0.007)
Constant	0.024***	-0.005	0.003	0.027***	-0.000	0.008	0.024***	-0.014	-0.007
	(0.002)	(0.026)	(0.026)	(0.002)	(0.026)	(0.026)	(0.002)	(0.027)	(0.026)
Fund FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	33,185	33,185	33,185	33,185	33,185	33,185	33,186	33,186	33,186
R-squared	0.142	0.147	0.149	0.147	0.151	0.153	0.150	0.151	0.153

Table 5: Social Media Research and Monthly Fund Flows

Notes: This table examines whether a fund's *SMR* variables affect mutual fund flows. The dependent variable is fund *i*'s percentage net flow in month m+*1*, and the primary explanatory variable is one of fund *i*'s adjusted *SMR* variables, *%Diff_No. of SA Stocks*, *%Diff_No. of SA Articles*, and *Diff_Net Bullishness* in quarter *t*. The definitions of the control variables are listed in Table A1 the Appendix. All regressions control for the fund and year fixed effects. Coefficients are marked with *, **, or *** for the significance level of 10%, 5%, and 1%, respectively. Standard errors are clustered at the fund level and reported in parentheses.

	Dependent Var: *Flow* $_{i,\,m+1}$								
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
%Diff_No. of SA Stocks $_{i,\,m}$	0.006*	0.006*	0.006*						
	(0.003)	(0.003)	(0.003)						
%Diff_No. of SA Articles $_{i,\,m}$				0.009**	0.010**	0.010***			
				(0.003)	(0.003)	(0.003)			
Diff_Net Bullishness $_{i,\,m}$							0.009***	0.006**	0.006**
							(0.003)	(0.003)	(0.003)
Newspaper_Article (ln) $_{i,\,m}$		-0.020	-0.025		-0.023	-0.028			
		(0.021)	(0.021)		(0.022)	(0.021)			
Newspaper_Tone $_{i,\,m}$								0.008**	0.008**
								(0.003)	(0.003)
Low $_{i,\,m}$		0.804***	0.788***		0.806***	0.790***		0.799***	0.783***
		(0.115)	(0.116)		(0.116)	(0.116)		(0.116)	(0.116)
Mid $_{i,\,m}$		0.026	0.027		0.026	0.027		0.026	0.027
		(0.019)	(0.019)		(0.019)	(0.019)		(0.019)	(0.019)
High $_{i,\,m}$		1.051***	1.044***		1.051**	1.044***		1.048***	1.041***
		(0.119)	(0.119)		(0.119)	(0.119)		(0.120)	(0.119)
ln(TNA) $_{i,\,m}$	-0.213***	-0.206***	-0.214***	-0.213***	-0.205***	-0.214***	-0.213***	-0.207***	-0.215***
	(0.030)	(0.030)	(0.031)	(0.030)	(0.030)	(0.031)	(0.030)	(0.030)	(0.031)
ln(Age) $_{i,\,m}$	-0.137*	-0.139*	-0.154**	-0.137*	-0.139*	-0.154**	-0.138*	-0.139*	-0.153**

	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)	(0.073)
Expense ratio $_{i,\,m}$	-0.187***	-0.187***	-0.178***	-0.187***	-0.187***	-0.178***	-0.187***	-0.187***	-0.177***
	(0.040)	(0.039)	(0.039)	(0.040)	(0.039)	(0.039)	(0.040)	(0.039)	(0.039)
Turnover $_{i,\,m}$	-0.045***	-0.043***	-0.042***	-0.045***	-0.043***	-0.042***	-0.045***	-0.043***	-0.042***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Beta_Mkt $_{i,\,m}$			-0.058***			-0.058***			-0.058***
			(0.013)			(0.013)			(0.013)
Beta_SMB $_{i,\,m}$			-0.047**			-0.047**			-0.046**
			(0.021)			(0.021)			(0.020)
Beta_HML $_{i,\,m}$			-0.036**			-0.036**			-0.036**
			(0.016)			(0.016)			(0.016)
Beta_UMD $_{i,\,m}$			0.034***			0.034***			0.034***
			(0.009)			(0.009)			(0.009)
Constant	-0.000	-0.171***	-0.167***	0.000	-0.171***	-0.167***	-0.000	-0.173***	-0.169***
	(0.003)	(0.020)	(0.020)	(0.003)	(0.020)	(0.020)	(0.003)	(0.020)	(0.020)
Fund FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	97,561	97,560	97,559	97,561	97,560	97,559	97,518	97,517	97,516
R-squared	0.134	0.139	0.141	0.134	0.139	0.141	0.134	0.139	0.141

Table 6: Event study and DID analysis

Notes: This table reports the event study results on mutual fund flows in Panel A and the Difference-in-Differences regression testing the mutual fund flows response to SA stock holdings in Panel B. The monthly indicator variable I(TREAT$_{i,m}$) is set to be 1 if fund *i*'s *%Diff_No. of SA stocks, %Diff_No. of SA articles*, or *Diff_Net Bullishness* measurement ranks in the top 30% in the calendar quarter, and 0 if the measurement is in the bottom 30%. The monthly binary variables I(Lead_Lag $_{i,M+n}$) (where n = -1, 0, 1) are assigned a value of 1 to signify that the specific month $M+n$ was n periods away from the SEC filing month M at the end of any quarter (i.e., March, June, September, and December), and 0 otherwise. The monthly indicator variable I(POST$_{i,m}$) is designated a value of 1 for the first month following a filing month M, and 0 otherwise. The dependent variable in both panels is the fund flows in month $m+1$. The regression models also include a complete list of control variables employed in the baseline analysis, in addition to fund and time fixed effects. Coefficients are marked with *, **, or *** for the significance level of 10%, 5%, and 1%, respectively. Standard errors are clustered at the fund level and reported in parentheses.

Panel A: Event study

Dependent Var: *Flow* $_{i, m+1}$

	(1) *%Diff_No. of SA Stocks*	(2) *%Diff_No. of SA Articles*	(3) *Diff_Net Bullishness*
I (Lead_Lag $_{i,M-1}$) * *Fwd.* I (TREAT $_{i,m}$)	0.031***	0.016	-0.015
	(0.012)	(0.011)	(0.011)
I (Lead_Lag $_{i,M}$) * I (TREAT $_{i,m}$)	-0.011	-0.031*	-0.019
	(0.016)	(0.016)	(0.016)
I (Lead_Lag $_{i,M+1}$) * *Lag.* I (TREAT $_{i,m}$)	0.111***	0.109***	0.077***
	(0.014)	(0.014)	(0.012)
Controls	Y	Y	Y
Fund/Year FE	Y	Y	Y
Observations	42,859	42,867	43,344
R-squared	0.145	0.154	0.153

Panel B: Difference-in-Differences analysis

Dependent Var: *Flow* $_{i, m+1}$

	(1) *%Diff_No. of SA Stocks*	(2) *%Diff_No. of SA Articles*	(3) *Diff_Net Bullishness*
I (POST $_{i,m}$) * *Lag.* I (TREAT $_{i,m}$)	0.098***	0.099***	0.084***
	(0.010)	(0.010)	(0.009)
Controls	Y	Y	Y
Fund/Year FE	Y	Y	Y
Observations	55,637	55,687	55,734
R-squared	0.143	0.152	0.150

Table 7: Social Media Research Interacting with Fund Characteristics

Notes: In each quarter, funds are classified into three groups based on the 30th and 70th percentiles of certain fund or fund holding characteristics, ordered from low to high (Group 1 to Group 3): *TNA, Age, Expense Ratio, Beta_SMB, Beta_HML*, *Price* (holding-weighted), and *MAX* (holding-weighted). The fund portfolios are rebalanced every quarter. The same regression analysis in Table 4 is performed for each group across the entire sample period, incorporating the full set of control variables. This table reports the coefficients of *%Diff_No. of SA Stocks* (Panel A), *%Diff_No. of SA Articles* (Panel B), and *Diff_Net Bullishness* (Panel C), along with the corresponding t-statistics in the parentheses. The Chow (1960) test is performed to assess the equality of the coefficients between Group 1 and Group 3, with the p-values of the F-statistics reported. Coefficients and p-values are marked with *, **, or *** for the significance level of 10%, 5%, and 1%, respectively.

Dependent Var: *Flow* $_{i,t+1}$												
	Panel A: *%Diff_No. of SA Stocks* $_{i,t}$				**Panel B:** *%Diff_ No. of SA Articles* $_{i,t}$				**Panel C:** *Diff_Net Bullishness* $_{i,t}$			
Rank	1	2	3	P-val of diff (1 vs 3)	1	2	3	P-val of diff (1 vs 3)	1	2	3	P-val of diff (1 vs 3)
TNA	0.009	0.011	0.024**	0.175	0.063***	0.061***	0.117***	0.000***	-0.069***	-0.093***	-0.122***	0.005***
	(0.772)	(1.183)	(2.089)		(4.786)	(6.020)	(10.242)		(-6.026)	(-9.961)	(-11.592)	
Age	0.006	0.027***	0.018	0.151	0.063***	0.092***	0.084***	0.027**	-0.076***	-0.097***	-0.121***	0.003***
	(0.640)	(2.657)	(1.420)		(5.867)	(8.242)	(5.927)		(-7.678)	(-8.807)	(-11.285)	
Expense ratio	0.016	0.023**	0.008	0.430	0.115***	0.068***	0.065***	0.002***	-0.136***	-0.092***	-0.070***	0.000***
	(1.242)	(2.096)	(0.771)		(7.991)	(6.276)	(5.550)		(-10.504)	(-9.895)	(-6.783)	
Beta_SMB	0.055***	0.022*	0.009	0.011**	0.131***	0.086***	0.047***	0.000***	-0.157***	-0.133***	-0.034***	0.000***
	(3.825)	(1.786)	(1.065)		(9.277)	(8.135)	(4.297)		(-11.690)	(-14.268)	(-3.787)	
Beta_HML	0.021*	0.018*	0.016	0.463	0.058***	0.089***	0.091***	0.514	-0.136***	-0.104***	-0.045***	0.000***
	(1.683)	(1.684)	(1.494)		(4.666)	(8.071)	(6.830)		(-12.320)	(-9.980)	(-3.889)	
Price	0.011	0.024**	0.011	0.288	0.049***	0.093***	0.106***	0.000***	-0.024***	-0.122***	-0.182***	0.000***
	(1.212)	(2.153)	(0.728)		(4.409)	(7.653)	(8.387)		(-2.596)	(-12.239)	(-14.811)	
MAX	0.044***	0.025**	0.003	0.017**	0.125***	0.097***	0.038***	0.000***	-0.141***	-0.116***	-0.041***	0.000***
	(3.178)	(2.552)	(0.278)		(9.137)	(9.276)	(3.322)		(-11.054)	(-12.437)	(-4.023)	

Table 8: Social Media Research and future fund performance

Notes: This table examines future fund performance, measured by cumulative net alphas across different investment horizons, against the *SMR* variable and the same set of control variables used in the fund flow regressions, along with the current-quarter *Net Alpha*. The analysis covers both short-term and mid-to-long-term fund performance, specifically over the window periods of [*t*+1, *t*+2], [*t*+3, *t*+4], and [*t*+1, *t*+4], where *t* represents the current quarter. Alphas for all investment periods are annualised and normalised. All regressions control for the fund and year fixed effects. Coefficients are marked with *, **, or *** for the significance level of 10%, 5%, and 1%, respectively. Standard errors are clustered at the fund level and reported in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	*Cum_Alpha (t+1, t+2)* $_{i,t}$	*Cum_Alpha (t+3, t+4)* $_{i,t}$	*Cum_Alpha (t+1, t+4)* $_{i,t}$	*Cum_Alpha (t+1, t+2)* $_{i,t}$	*Cum_Alpha (t+3, t+4)* $_{i,t}$	*Cum_Alpha (t+1, t+4)* $_{i,t}$	*Cum_Alpha (t+1, t+2)* $_{i,t}$	*Cum_Alpha (t+3, t+4)* $_{i,t}$	*Cum_Alpha (t+1, t+4)* $_{i,t}$
%Diff_No. of SA stocks $_{i,t}$	0.019***	-0.023***	-0.004						
	(0.006)	(0.006)	(0.005)						
%Diff_No. of SA articles $_{i,t}$				0.013**	-0.027***	-0.008			
				(0.006)	(0.007)	(0.006)			
Diff_Net Bullishness $_{i,t}$							-0.021***	0.004	-0.014***
							(0.006)	(0.006)	(0.004)
Newspaper_Article (ln) $_{i,t}$	-0.024	0.050*	-0.008	-0.024	0.053*	-0.006			
	(0.031)	(0.030)	(0.040)	(0.031)	(0.030)	(0.040)			
Newspaper_Tone $_{i,t}$							-0.008	0.003	-0.005
							(0.007)	(0.008)	(0.008)
Net Alpha $_{i,t}$	-0.075***	-0.019**	-0.062***	-0.074***	-0.020**	-0.062***	-0.073***	-0.020**	-0.061***
	(0.008)	(0.009)	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.009)	(0.008)
ln(TNA) $_{i,t}$	-0.517***	-0.406***	-0.657***	-0.516***	-0.408***	-0.658***	-0.515***	-0.405***	-0.655***
	(0.039)	(0.038)	(0.053)	(0.039)	(0.038)	(0.053)	(0.039)	(0.038)	(0.053)
ln(Age) $_{i,t}$	-0.252***	-0.202**	-0.273**	-0.254***	-0.198**	-0.273**	-0.250***	-0.203**	-0.271**
	(0.092)	(0.090)	(0.125)	(0.092)	(0.090)	(0.125)	(0.092)	(0.090)	(0.125)
Expense ratio $_{i,t}$	-0.127***	-0.070*	-0.133**	-0.126***	-0.071*	-0.133**	-0.126***	-0.071*	-0.133**
	(0.040)	(0.039)	(0.053)	(0.040)	(0.039)	(0.053)	(0.040)	(0.039)	(0.053)
Turnover $_{i,t}$	-0.099***	-0.064***	-0.114***	-0.099***	-0.064***	-0.113***	-0.099***	-0.065***	-0.114***
	(0.016)	(0.015)	(0.020)	(0.016)	(0.015)	(0.020)	(0.016)	(0.015)	(0.020)

Beta_Mkt $_{i,\,t}$	-0.052***	0.027**	-0.031**	-0.052***	0.028**	-0.031**	-0.053***	0.028**	-0.032**
	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.013)
Beta_SMB $_{i,\,t}$	-0.044*	-0.083***	-0.121***	-0.044*	-0.081***	-0.121***	-0.042	-0.088***	-0.122***
	(0.026)	(0.026)	(0.027)	(0.026)	(0.026)	(0.027)	(0.026)	(0.026)	(0.027)
Beta_HML $_{i,\,t}$	0.090***	0.084***	0.111***	0.090***	0.084***	0.111***	0.089***	0.085***	0.111***
	(0.016)	(0.016)	(0.019)	(0.016)	(0.016)	(0.019)	(0.016)	(0.016)	(0.019)
Beta_UMD $_{i,\,t}$	-0.086***	-0.043***	-0.084***	-0.087***	-0.043***	-0.084***	-0.087***	-0.042***	-0.084***
	(0.013)	(0.010)	(0.013)	(0.013)	(0.010)	(0.013)	(0.013)	(0.010)	(0.013)
Constant	0.020***	0.011***	0.028***	0.020***	0.011***	0.028***	0.019***	0.014***	0.029***
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
Fund FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	31,832	29,732	29,346	31,832	29,732	29,346	31,833	29,733	29,347
R-squared	0.162	0.136	0.282	0.162	0.136	0.282	0.162	0.135	0.282

Appendix

Table A1: List of Variable Definitions

Variable	Definition
Age	The number of years since the issuance of the earliest share class in the fund.
Beta_HML	The beta loading of the value factor (high-minus-low) estimated from regressing 60-month rolling window net-of-expense fund returns on the Carhart (1997) four factors.
Beta_Mkt	The beta loading of the market factor estimated from regressing 60-month rolling window net-of-expense fund returns on the Carhart (1997) four factors.
Beta_SMB	The beta loading of the size factor (small-minus-big) estimated from regressing 60-month rolling window net-of-expense returns on the Carhart (1997) four factors.
Beta_UMD	The beta loading of the momentum factor (up-minus-down) estimated from regressing 60-month rolling window net-of-expense fund returns on the Carhart (1997) four factors.
Common Stock (%)	The common stock percentage holding of a fund.
Diff_Net Bullishness	The simple first difference of *Net Bullishness*.
Expense ratio	The annualised *TNA*-weighted averages of expense ratio across all fund share classes.
Fund Flow (monthly)	Calculated as: $\frac{TNA_{i,m} - TNA_{i,m-1} * (1 + ret_{i,m})}{TNA_{i,m-1} * (1 + ret_{i,m})}$, where $ret_{i,m}$ is the *TNA*-weighted averages of monthly raw returns across all fund share classes.
Fund Flow (quarterly)	Calculated as: $\frac{TNA_{i,t} - TNA_{i,t-1} * (1 + ret_{i,t})}{TNA_{i,t-1} * (1 + ret_{i,t})}$, where $ret_{i,t}$ is the *TNA*-weighted averages of quarterly raw returns (holding period returns computed from monthly returns) across all fund share classes.
Gross Alpha	Quarterly alphas estimated from the Carhart (1997) four-factor model using daily gross fund returns within each quarter.
Low	For each quarter (or month), we assign all funds fractional ranks (*Rank*), according to their Carhart (1997) four-factor alpha, which are uniformly distributed between 0 (worst performance) and 1 (best performance). The variable Low_i is defined as MIN (0.2, *Rank*), *Mid* is defined as MIN (0.6, *Rank* - *Low*), and *High* is computed as *Rank* - *Low* - *Mid*.
Mid	
High	
I(Lead_Lag $_{i,M+n}$)	I(Lead_Lag $_{i,M+n}$) (where n = -1, 0, 1) are assigned a value of 1 to signify that the specific month M+n was n periods away from the SEC filing month M at the end of any quarter (i.e., March, June, September, and

	December), and 0 otherwise.
I($POST_{i,m}$)	I($POST_{i,m}$) is designated a value of 1 for the first month following a filing month *M*, and 0 otherwise.
I($TREAT_{i,m}$)	I($TREAT_{i,m}$) is set to be 1 if fund *i*'s *%Diff_No. of SA stocks, %Diff_No. of SA articles*, or *Diff_Net Bullishness* measurement ranks in the top 30% in the calendar quarter, and 0 if the measurement is in the bottom 30%.
MAX	The holding-weighted average of the stock-level MAX measure, which is the maximum daily returns of a stock within a month and averaged over a quarter.
Net Alpha	Quarterly (monthly) alphas estimated from the Carhart (1997) four-factor model using daily net-of-expense fund returns within each quarter (month).
Net Bullishness_Stock	The daily sentiment score for each *SA* stock is calculated as: $\frac{No.of\ positive\ words - No.of\ negative\ words}{Total\ No.of\ words}$ The quarterly or monthly sentiment score for each *SA* stock is determined by averaging the daily scores within the respective period.
Net Bullishness (monthly)	The holding-weighted average of the monthly stock-level sentiment score, using a fund's previous-quarter holdings as weights.
Net Bullishness (quarterly)	The holding-weighted average of the quarterly stock-level sentiment score, using a fund's current-quarter holdings as weights.
Newspaper_Article	The total number of newspaper articles from *The New York Times, The Washington Post, USA Today*, and *The Wall Street Journal* for stocks held by a fund.
Newspaper_Tone (monthly)	The holding-weighted average of the monthly stock-level RavenPack Composite Sentiment Score, using a fund's previous-quarter holdings as weights.
Newspaper_Tone (quarterly)	The holding-weighted average of the quarterly stock-level RavenPack Composite Sentiment Score, using a fund's current-quarter holdings as weights.
No. of SA Articles	The total number of *SeekingAlpha.com* articles for stocks held by a fund.
No. of SA Stocks	The total number of *SA* stocks (a stock with at least one article on *SeekingAlpha.com* during a certain period) held by a fund.
Price	The holding-weighted average of the prices (monthly closing prices averaged over a quarter) of stocks held by a fund.

TNA	The total net assets in $ million of a fund by summing the total net assets across all share classes of a fund.
Turnover	The annual *TNA*-weighted averages of turnover ratio across all fund share classes.
%Diff_No. of SA stocks	The percentage first difference of *No. of SA Stocks*.
%Diff_No. of SA articles	The percentage first difference of *No. of SA Articles*.

Table A2: Mutual Fund Quintile Portfolios Sorted on Adjusted *SMR* Variables

Funds are classified into quintile portfolios based on *%Diff_No. of SA Stocks*, *%Diff_ No. of SA Articles*, or *Diff_ Net_Bullishness* from the formation quarter. Quintile 1 (Low) comprises funds with the lowest value of the selected variable, whereas Quintile 5 (High) includes funds with the highest value. These quintile portfolios are rebalanced every calendar quarter. Following this classification, we calculate the average values of *TNA, Age, Expense ratio (%), Turnover*, fund betas, and *Net Alpha* for the funds within each quintile during the formation quarter. Next, we compute and report the mean of the time-series of these average values for each characteristic across each quintile. We also calculate the differences between the High (Quintile 5) and Low (Quintile 1) quintiles, providing the Newey-West (1987) adjusted t-statistics for these differences.

	TNA ($ Million)	*Age (Year)*	*Expense ratio (%)*	*Turnover*	*Beta_Mkt*	*Beta_SMB*	*Beta_HML*	*Beta_UMD*	*Net Alpha (%)*
Panel A: *%Diff_No. of SA Stocks*									
Low	2690.692	26.730	1.109	0.662	0.979	0.243	0.027	0.017	-0.382
2	5487.850	27.865	0.997	0.555	0.985	0.112	-0.004	0.012	-0.340
3	7310.564	27.921	0.954	0.514	0.985	0.078	-0.007	0.008	-0.332
4	6359.835	27.559	0.994	0.570	0.984	0.141	-0.003	0.017	-0.265
High	2599.965	26.378	1.115	0.667	0.976	0.275	0.020	0.023	-0.302
High-Low	-90.727	-0.352	0.006	0.004	-0.004	0.033	-0.007	0.006	0.081
t-stat	-0.58	-1.06	0.82	0.44	-0.79	0.84	-0.56	0.75	0.70
Panel B: *%Diff_ No. of SA Articles*									
	TNA ($ Million)	*Age (Year)*	*Expense ratio (%)*	*Turnover*	*Beta_Mkt*	*Beta_SMB*	*Beta_HML*	*Beta_UMD*	*Net Alpha (%)*
Low	2338.723	26.531	1.120	0.667	0.979	0.248	0.026	0.013	-0.365
2	5572.980	28.099	1.003	0.558	0.983	0.122	0.002	0.011	-0.261
3	8655.659	27.847	0.926	0.511	0.985	0.065	-0.004	0.011	-0.311
4	5355.290	27.540	1.004	0.565	0.985	0.135	-0.006	0.018	-0.355
High	2300.588	26.427	1.120	0.669	0.978	0.279	0.015	0.024	-0.336
High-Low	-38.135	-0.104	0.000	0.003	-0.001	0.031	-0.011	0.011	0.030
t-stat	-0.25	-0.46	-0.02	0.22	-0.26	0.99	-0.77	1.08	0.23
Panel C: *Diff_Net Bullishness*									
	TNA ($ Million)	*Age (Year)*	*Expense ratio (%)*	*Turnover*	*Beta_Mkt*	*Beta_SMB*	*Beta_HML*	*Beta_UMD*	*Net Alpha (%)*
Low	2425.915	26.383	1.121	0.620	0.980	0.264	0.033	0.013	-0.573
2	5053.190	28.098	1.004	0.588	0.984	0.124	-0.010	0.020	-0.398
3	9346.484	27.607	0.919	0.539	0.987	0.076	-0.006	0.016	-0.256
4	4928.851	27.945	1.011	0.604	0.987	0.122	-0.007	0.014	-0.244
High	2469.043	26.412	1.118	0.619	0.971	0.263	0.023	0.014	-0.156
High-Low	43.128	0.029	-0.003	-0.002	-0.008	-0.001	-0.010	0.001	0.417
t-stat	0.40	0.17	-0.65	-0.21	-1.70	-0.04	-0.64	0.14	***3.75***