These Stocks are on Fire: The Impact of Social Media on Mutual Funds' Performance, Flows, and Trading*

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

Abstract

The recent mania of Reddit darling GameStop raises questions about social media's impact on behaviours of institutional investors such as mutual funds. We examine whether and why mutual funds hold "sentimental" stocks that are heavily mentioned or with bullish views on social media. Based on a sample of 1,028 US domestic equity funds between 2011 and 2016, we find that mutual funds holding more of these stocks have short-lived outperformance over their peers. Such funds also attract substantially higher fund flows, especially those of individual investors. Our results suggest that mutual fund managers cater to investors' sentimental demand displayed on social media to boost short-term performance and to attract flows.

Keywords: Social media; Investor attention; Investor sentiment; Fund performance; Fund flows.
JEL Classification: G11; G23.

* This is a preliminary draft. Please consult the authors before citation. All errors are our own.
[1] Cardiff Business School, Cardiff University, UK
[2] Durham University Business School, Durham University, UK
[3] School of Finance, Southwestern University of Finance and Economics, China

1. Introduction

It has been well documented that social interaction affects investors' trading behaviour and in turn economic outcomes (e.g., Hong et al., 2004; Ivkovic and Weisbenner; 2007, Brown et al., 2008; Ozsoylev et al., 2014; Pool et al., 2015; Hirshleifer et al., 2021). Social media activities, just like any other forms of social interaction, allow for exchanges of information and misinformation. While the literature has found that social media can play an informational role in the stock market by conveying novel and value-relevant news (Chen et al., 2014; Bartov et al., 2018; Chen et al., 2019; Gu and Kurov, 2020; Farrell et al., 2020), we argue that interactions on social media, a type of social processes,[4] have the potential to make the presence of investor sentiment — broadly defined as demand that is unjustified by the facts at hand (Baker and Wurgler, 2006; McLean and Zhao, 2014) — and its impact more pronounced for at least several reasons. *First*, unlike traditional forms of social interaction (e.g., communication between a small group of people such as neighbours, colleagues, advisors and clients), social media platforms (e.g., Twitter, Facebook, stock forums) make the dissemination of information and opinions to a broad investment community faster and easier, and can amplify small bias into large effects (compounding) through interaction; *Second*, social media may exacerbate social transmission bias, making sentiment of social media users more difficult to predict (idiosyncrasy). This induces larger noise trader risk, creating room for stronger impact of investor sentiment (De Long et al., 1990). *Third*, interactions on social media can endogenously lead to swings in investor sentiment as a social emergent behaviour (Hirshleifer, 2020). A prominent recent example is an army of Reddit users skyrocketed the stock price of GameStop and unsettled professional short sellers such as hedge funds.

[4] Hirshleifer (2020) argues that social processes shapes both the economic thinking and behaviour of economic agents and can lead to biased opinions.

The broad definition of investor sentiment that we adopt allows us to consider the demand shock arising not only from shifts in bullishness of investors beliefs, but also from investor limited attention. On the one hand, when noise traders have more bullish view on returns of risky assets, they increase their demand for these assets (De Long et al., 1990). On the other hand, attention-constrained investors (such as individual investors) face a formidable search problem when buying, but less so when selling. Therefore, they tend to be net purchasers of attention-grabbing stocks (Barber and Odean, 2008). Both types of demand shock lead to temporary overpricing in the presence of limits to arbitrage. As mentioned above, social media can amplify these demand shocks. For example, the bullish/bearish views on social media can spread from one investor to others. Similarly, when an asset captures attention of some investors, their posts (and reposts) can attract further attention of other investors.

Meanwhile, existing literature has documented the relationship between investor sentiment (including investor attention) and various aspects of mutual funds. For example, Cooper et al. (2005), Lou (2012), and Kamstra et al. (2017), amongst others, show that mutual fund flows contain information about investor demands and/or sentiment. 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, hence, attract more fund flows. Solomon et al. (2014) conclude that the media coverage of stocks, a proxy of attention, affect investors' capital allocations to mutual funds.

This research studies a rarely explored question: how does social media affects mutual fund behaviours? More specifically, we investigate the impact of social media activities on mutual fund performance, flows, and fund managers’ trading behaviour. Our *first* hypothesis is that mutual funds which hold more stocks that are recently mentioned (especially heavily mentioned with bullish opinions) by investors on social media platforms are likely to have better short-term abnormal performance. This is based on the assumption that investors, especially individual investors, tend to be the net-purchasers of stocks that have a greater number of social media messages and/or more bullish views, as demands for these stocks are partly driven by excessive attention[5] and overoptimistic views — these stocks tend to outperform in the short term in the presence of limits to arbitrage. We believe that social media platforms are one of the most popular channels through which individual investors obtain information and opinions, as individual investors are generally unable to access many other credible information sources that are available to institutional investors.

Our second hypothesis posits that investors may chase the recent *social media active/bullish* stocks by investing in mutual funds that have greater holdings (in both absolute and relative terms) in stocks of such type. While investors can separately buy single *social media active/bullish* stocks, it is much costly than purchasing mutual funds holding a bunch of such stocks. Mutual funds also provide diversification, which is an important consideration for investors. A sub-hypothesis is that the effect of social media on the fund flows should be stronger among funds that are mostly held by individual investors, as they are more subject to excessive/limited attention and investor sentiment.

Finally, we hypothesise that fund managers' trading decisions are affected by the fund holdings' *social media activeness/bullishness* (*Hypothesis Three*). There are two sub-hypotheses — 1) managers cater to investors' preference of the recent *social media active/bullish* stocks by actively

[5] Traditional proxies for investor attention include traditional media coverage (Solomon et al., 2014), extreme price movements (Barber & Odean, 2008), and advertising expenses (Lou, 2014). These proxies, however, captures 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 mentioning can potentially capture both active and passive investor attention on these social media platforms.

trading such stocks to attract flows; 2) Although fund managers are professionals, they also hold overly bullish/bearish beliefs that are similar to those on social media and have limited attention. Therefore, they tend to be the net purchasers of the most recent *social media active/bullish* stocks. The first sub-hypothesis has support from previous studies and the industry. For example, Agarwal et al. (2020) show that mutual funds, especially those small, young, and poor-performing funds, hold lottery-type stocks not because of fund managers' own preferences, but because of managers' efforts to cater to investors' gambling preference and attract fund flows. In addition, following the Reddit-GameStop event, a 'social media sentiment ETF'[6] has been launched to cater to investors' preference to stocks that are most talked up with positive views on social media every month. Regarding the second sub-hypothesis, the prediction is that social media activeness/bullishness will have a larger impact on managers' buy than sell decisions. In contrast, such prediction cannot be made from the first sub-hypothesis.

Our sample consists of 1,105 mutual funds and 69,672 fund-month observations from 2011 to 2016. They are actively managed and diversified U.S. domestic equity funds, which have 80% or greater of assets held in US common stocks. The social media coverage and sentiment data for stocks are provided by *PsychSignal*, a third-party company that extracts social media coverage (i.e., number of messages) and sentiment data from Twitter and StockTwits. We find that social media mentioning is concentrated among the largest stocks, which tend to have higher individual ownership as the largest firms have much better name recognition among individual investors than small and mid-cap firms. Growth stocks, stocks with higher total risk or idiosyncratic volatility, and older stocks have more social media mentioning. Investors' views are more bullish for value stocks. The smallest stocks on average have extremely bullish views. Investors' opinions on larger

[6] VanEck Vectors Social Sentiment ETF

stocks are overall neutral or relatively bearish. Investors also tend to be more bullish about stocks with lower total risk or idiosyncratic volatility. It can be concluded that *social media active/bullish* stocks are not exactly lottery-type stocks,[7] which are low-priced stocks with high idiosyncratic volatility and high idiosyncratic skewness. We define a social media stock as a stock which has at least one social media message during a certain period (a month or quarter). Next, we compute a *Bull-Bear Score* (i.e., bullishness minus bearishness) for each SM stock. For every fund, we aggregate the data for every stock held by the fund and construct the following fund-level social media measures: the number of SM stocks, total number of messages, and the *Bull-Bear Score* (average or holding-adjusted). These variables represent the fund's *Social Media Feature*.

We find that the social media feature of mutual funds is significantly correlated with the funds' short-term monthly performance. Funds holding more social media stocks, those having more messages about the stocks they hold, or those with higher level of aggregate bullishness perform significantly better[8] in the current and next months. The pattern regarding the aggregate bullishness is the most pronounced. Meanwhile, the strong performance is short-lived and generally disappears in the next 6 months or 12 months. The different dimensions of a fund's social media feature have different levels of influence on its performance. The effect of bullishness seems to outweigh that of the total number of messages or the number of social media stocks. For instance, among the funds with the highest level of bullishness, those with fewer total number of messages perform better. Overall, these findings strongly support our *first* hypothesis. Our results also lend strong support to the *second* hypothesis. After accounting for funds' performance and some fund characteristics, the number of social media stocks held by the funds, the total number of messages,

[7] Individual investors have strong preferences for lottery-type stocks (Kumar, 2009; Bali et al., 2011; Han and Kumar, 2013).
[8] Fund performance is measured by alphas adjusted against the Carhart (1997) four factors

and the *Bull-Bear* Score are all significantly positively associated with the contemporaneous and next-period monthly mutual fund flows. The positive effects of total number of messages and the *Bull-Bear Score* on fund flows are more pronounced among smaller funds and individual-investor-dominated funds. Regarding the *third* hypothesis, we confirm that fund managers' trading decisions are influenced by how often the fund holdings are mentioned in social media. After controlling for stock characteristics such as size, B/M ratio, age, etc., the total number of messages about a stock, or a dummy variable indicating whether a stock is often talked up on social media, is positively associated with the dollar amount of a fund's buys in that stock. The social media activeness of a stock also increases managers' sell amount. Meanwhile, the bullishness measure has limited impact on managers' trading decisions. Given that there is little evidence that social media activities have a larger impact on managers' buy than sell decisions, our results are more consistent with the *flow-catering* hypothesis — managers cater to investors' preference of the stocks that are more mentioned and positively talked up in social media by actively trading such stocks to attract flows.

Our paper contributes to the literature that examines the role of social media in the stock market. Previous studies show that the aggregate opinion (i.e., positivity/negativity) of firm-specific social media messages can convey value-relevant news (Chen et al., 2014; Bartov et al., 2018; Gu and Kurov, 2020), such as predicting a firm's forthcoming quarterly earnings and announcement returns (Bartov et al., 2018). Prior literature also provides evidence that social media activities and opinions affect trading volume and liquidity, leading to price predictability (Sprenger et al., 2014; Agrawal et al., 2018; Rakowski et al., 2021). We provide further evidence that social media may amplify investor sentiment of a large group of investors, especially individual investors, which is reflected in the superior short-term performance of mutual funds that have high holdings in recent *social media active/bullish* stocks and the increased flows to such funds.

Our paper also complements the studies that examine the impact of traditional media (newspapers) coverage on mutual fund flows (i.e., investors' investment behaviour), performance, and fund managers' investment decisions (e.g., Sirri and Tufano, 1998; Fang et al., 2014; Solomon et al.,2014; Kaniel and Parham; 2017). The differences between traditional and social media are distinct. While investors unilaterally obtain information and opinions from the mass print media, there is much more involvement and interaction on social media. Fang et al. (2014) show that mutual funds tend to buy than sell stocks covered by traditional media, suggesting that professional investors are subject to limited attention too. They also find no significant relationship between the media coverage of fund holdings and fund flows. In contrast with their evidence from traditional media, we provide strong evidence that social media activities and opinions of fund holdings affect flows. Sirri and Tufano (1998) and Kaniel and Parham (2017) base their analysis on mentions of mutual *fund names* in major newspapers, while we take a more detailed look by aggregating the number of messages and opinions for every stock held by a fund to form the fund-level measures.

The rest of the paper is organised as follows. Section 2 describes our mutual fund sample, social media data, and the relationship between stock characteristics and social media variables. Section 3 reports the results for testing the three main hypotheses and corresponding sub-hypotheses and discusses. Section 4 concludes.

2. Data Description

In this section, we describe the construction of our mutual fund sample, the social media dataset and variables, how stock characteristics are associated with social media variables, and the monthly fund-level social media feature.

2.1 Mutual Fund Sample

Our dataset combines information on mutual fund holdings with stocks' social media measures (explained in Section 2.2). We construct our main sample of mutual funds by merging CRSP Survivor-Bias-Free Mutual Fund Database with Thomson Reuters Mutual Fund Holdings Database. To make our study comparable with the existing literature, we include only actively managed, diversified U.S. domestic equity funds. We exclude index funds, international funds, municipal bond funds, bond and preferred stock funds, sector funds, and funds that cannot be linked to the CRSP mutual fund database via the Mutual Fund Links data set available from Wharton Research Data Services. Further, we require that funds have 80% or greater of assets held in US common stocks.

Thomson Reuters provides information on long positions in US common stock holdings of mutual funds on quarterly basis. For funds that do not follow the regular quarterly update on portfolio holdings, we use its most recent report to infer holdings in these months. We assume consistent holding throughout the report period as intra-quarter trading information is unavailable.

Lastly, we require that a fund must hold at least one stock that was once mentioned on social media during the sample period (2011 to 2016). Our final sample consists of 1,105 mutual funds. We exclude fund months with no social media data, ending up with 69,672 fund-month observations. In total, 1,705 out of 2,765 domestic (US) stocks held by these funds are identified with social media mentioning and sentiment data. While some funds also hold securities of other asset classes or foreign stocks, the proportion is small. Our social media data, described in the next section, mainly cover US stocks.

2.2 Social Media Measures

The social media coverage and sentiment data are obtained from *PsychSignal*, a provider of real-time social data and sentiment analytics for financial institutions and investment professionals. The database covers over 10,500 equities, ETFs, and futures. It extracts social media coverage (i.e., number of messages) and sentiment data from Twitter and StockTwits, dating back to 2009. The messages about individual stocks are identified by "CashTags", which indicate that the user is discussing a financial security (eg., $SPY, $GLD, $XOM, $AAPL). *PsychSignal* filters for relevant conversations to reduce the noise and accurately determine the author's sentiment surrounding stock symbols. *PsychSignal*'s proprietary algorithm derives sentiment measures by reading the language used in the conversation against its sentiment database compiled from thousands of interviews with financial trading professionals. *PsychSignal* boasts that the advantage of this linguistics-based approach is that its large, highly specialized lexicon of phrases is perfectly suited for interpreting short-form messages (e.g., Tweets).

For every stock that has once been held during the sample period by the mutual funds retrieved, we obtain the following daily data fields — the total number of messages, the number of bull-scored messages, and the number of bear-scored messages. Note that not all messages are classified as bull-scored or bear-scored messages by *PsychSignal*. These data fields have two versions, one generated from original posts only, and the other generated with both original posts and re-posts.

Since we find that the two versions of data series have very high correlations (greater than 0.95), we simply adopt the version without retweets.

We define a *social media (SM) stock* as a stock which has at least one message during a certain period (a month or quarter)[9]. Next, we compute a *Bull-Bear Score* for each SM stock as follows:

$$Bull\ Bear\ Score = \frac{\text{Number of Bull Messages} - \text{Number of Bear Messages}}{\text{Number of Bull Messages} + \text{Number of Bear Messages}} \qquad (1)$$

For every fund, we aggregate the data for every stock held by the fund and construct the following social media variables: the number of SM stocks, total number of messages and the *Bull-Bear Score* (average or weighted by the fraction of holdings in the fund).[10] These variables represent the fund's social media feature (SMF).

2.3 Stock Characteristics vs. Social Media Variables

In this section we examine the relationship between some characteristics of stocks and the social media indicators (i.e., total number of messages and the *Bull-Bear Score*). We would like to know which types of stocks attract higher investor attention or receive more bullish comments on social media.

Every year between 2011-2016, stocks with no message are classified into the 'no message' group. Stocks with at least one message are sorted by their total number of messages (or the *Bull-Bear Score*) into deciles from the lowest (group 1) to the highest (group 10). We then calculate the

[9] Our analysis consists of samples with different frequency. Section 3.1 and 3.2 provides estimations using fund-month level data while Section 3.3 presents result with fund-stock-quarter level data, given the nature of quarterly update on fund stock holding information.
[10] We measure social media variables in both monthly and quarterly frequencies in different parts of analysis. Detailed variable definitions can be found in Appendix A.

yearly average standard deviation (using monthly returns), idiosyncratic volatility,[11] market capitalization, B/M ratio, and age for each group. Lastly, we calculate the average values of these variables for all years (2011 to 2016). Figure 1 (a) and (b) present these values for each group. Values of all characteristic variables are adjusted so that the value for group 1 is set to be 100.

[Insert Figure 1]

In Figure 1(a), the most obvious pattern is the positive relationship between the market capitalisation of a stock and the number of messages mentioning that stock on social media platforms. Larger stocks have greater number of messages. Stocks in group 10, the group with the greatest number of messages, are significantly larger in size than stocks in all other groups, indicating that social media discussion is concentrated among the largest stocks, which tend to have a broad investor base as the largest firms have much better name recognition among investors, including individual investors. This finding is consistent with those found for traditional media (e.g., Fang and Peress, 2009; Hillert et al., 2014; Liu and Han, 2020) — larger stocks have greater media coverage. Figure 1(a) also shows that stocks with lower B/M ratio (growth stocks), higher total risk (standard deviation) or idiosyncratic volatility, and older stocks have more social media mentioning.

In Figure 1(b), the most obvious pattern is that investors are more bullish about stocks with higher B/M ratio (value stocks). It seems that relatively small stocks tend to have more extreme sentiment, either bearish or bullish, though the smallest stocks tend to be more extremely bullish. Investors' opinions on larger stocks are overall neutral or relatively bearish. The finding regarding the relationship between size and average opinion is in line with the finding for traditional media

[11] Calculated every year with daily return data regressed against the Carhart (1997) 4-factor model.

stories (see Liu and Han, 2020). Investors also tend to be more bullish about stocks with lower total risk (standard deviation) or lower idiosyncratic volatility.

2.4 Summary Statistics of the Mutual Fund Sample

Table 1 reports descriptive statistics for our mutual fund sample. Panel A tabulates the fund characteristics. In our sample, there are very small funds with total net asset less than 1 million, as well as big funds with around half billion ($498 Million) assets. The average turnover ratio at 61.7% and the percentage of common stocks holding at 86.7% indicate that the selected equity funds are actively managed. The average annual expense ratio of our sample is 1%, ranging from 0% to 2.3%. Load is the summation of front and rear load charges. Management fee normally includes the cost obtained from the Statement of Operations, which can be offset by fee waivers and/or reimbursements and lead to negative values (minimum management fee is -0.02%). The average monthly return before adjusting for expenses is 0.9%. In our sample, funds have negative monthly flow of -0.6%.

Panel B presents the descriptive statistics of the funds' monthly SMF. The average number of SM stocks held by a fund in a month is around 41, with a standard deviation of greater than 77, and the maximum value is 1,379. The mean of *Bull-Bear Score* (average) is 0.457, while the mean of *Bull-Bear Score* (holding-adjusted) is much smaller (0.139), indicating that investors' opinions on smaller holdings (i.e., SM stocks with small weights) tend to be more bullish. This finding is consistent with the evidence shown in Section 2.3 that investors' opinions on the smallest stocks, which normally have small weights in a fund, tend to be more extremely bullish and opinions on larger stocks are overall neutral or relatively bearish.

[Insert Table 1]

3. Results and Discussions

This section presents the results and discussions for testing the three main hypotheses and the corresponding sub-hypotheses mentioned in the introduction. In the first sub-section, we examine how social media activities and sentiment of mutual fund holdings are associated with short- to medium-term fund performance. In the second sub-section, we investigate if investors' fund investment decisions are affected by the level of *social media activeness/bullishness* of stocks that a fund holds. In the last sub-section, we test if fund managers' trading volumes (buys and sells) are affected by stocks' social media measures.

3.1 Social Media and Mutual Fund Performance

As mentioned earlier, for every fund-month observation, we have obtained the fund's number of SM stock held, total number of messages, and the *Bull-Bear Score* (average or holding-adjusted). These variables represent the fund's SMF, which indicates if the fund holds many stocks that are mentioned on social media, if these SM stocks are on average heavily discussed by investors, and how bullishness investors' opinions are towards these stocks. In this section, we would like to examine the relationship between the funds' SMF variables and their short- to medium-term performance.

3.1.1 Single Sort

For every month *t*, we sort funds with at least one SM stock into quintiles, from low to high, by the number of SM stocks, total number of messages, and the *Bull-Bear Score* (average or holding-adjusted), respectively. We then examine each quintile's performance in month *t*, *t*+6 and *t*+12.

Within each quintile, we obtain each fund's alpha by regressing its return against the Carhart (1997) 4-factor model. The alphas of all funds are then equally weighted to obtain the quintile's alpha.

Table 2 reports the results. Column (a) shows that when funds are sorted by the number of SM stock they hold, the alphas of quintiles are generally increasing from Q1 to Q5. The return differences between Q5 and Q1 in month *t*[12], *t*+6 and *t*+12 are all significantly positive at the 1% level, although the magnitude is relatively small—1.89%, 1.74%, and 0.75% in annual terms. Column (b) shows that when funds are sorted by the total number of messages, the return differences between Q5 and Q1 in month *t* and *t*+6 are significantly positive at the 1% level (1.34% and 1.83% annually), but the return difference in month *t*+12 is insignificant. Column (c) and (d) show that as the *Bull-Bear Score* (average or holding-adjusted) increases from Q1 to Q5, the alpha in month *t* considerably increases. For example, the alpha of Q1 (Q5) for *Bull-Bear Score* (average) is -4.57% (2.81%) in annual terms, and the difference between Q5 and Q1 is 7.38% per annum. However, the positive return difference disappears in month *t*+6 and *t*+12.

[Insert Table 2]

Overall, results in this section show that the SMF of mutual funds in a month is strongly associated with the funds' contemporaneous performance. Funds holding more SM stocks, those having more social media mentioning about their holdings, or those with higher level of aggregate bullishness perform significantly better. The pattern regarding the aggregate bullishness is the most pronounced. Meanwhile, the strong performance generally vanishes in the next 6 months or 12 months, suggesting that the initial outperformance is only short-lived.

---

[12] Although not reported in Table 2, the pattern of fund performance for quintiles in month *t+1* is qualitatively similar to that in month *t*.

3.1.2 Double Sorts

For every month *t*, funds are *independently* double sorted in to terciles by any two SMF variables (denoted as A and B). Hence, we form a set of 3 × 3 fund groups. In Table 3, we report two sets of results—number of SM stocks ×*Bull-Bear Score* (holding-adjusted) and the total number of messages × *Bull-Bear Score* (holding-adjusted). We report alphas adjusted against the Carhart (1997) 4-factor model in both gross and net terms. Group A1, A2, and A3 are terciles sorted by the number of SM stock (Panel (a)) or total number of messages (Panel (b)), and group B1, B2, and B3 are terciles sorted by *Bull-Bear Score* (holding-adjusted).

[Insert Table 3]

In panel (a) of Table 3, alphas of A1 to A3 (B1 to B3) consistently increase within all B (A) groups. The return differences between A3 and A1 are significantly positive at the 1% level within B1 and B2—1.58% and 1.50% annually (gross) or 1.92% and 1.79% (net), but insignificant within B3. The return differences between B3 and B1 are consistently significantly positive at the 1% level within all A groups—3.76%, 4.23%, and 2.62% (gross) or 3.68%, 4.12%, and 2.40% (net). Among all 3×3 groups, group A3-B3 has the highest alpha.

In panel (b) of Table 3, alphas of A1 to A3 are increasing within B1 and B2, and the return differences between A3 and A1 are significantly positive at the 1% level—2.51% and 1.56% (gross) or 2.81% and 1.85% (net). However, alphas of A1 to A3 decrease within B3, indicating that among the funds with the highest level of bullishness, those with fewer total number of messages perform better— the return difference between A3 and A1 is -2.25% (gross) or -2.03% (net). Such results might arise from the characteristics of the stocks held by these funds. For example, these funds may hold many smallest stocks, while Section 2.3 already shows that smallest stocks tend to have high level of bullishness but fewer number of messages. Alphas of B1 to B3 consistently increase within all A groups. The return differences between B3 and B1 are consistently significantly positive at the 1% level within all A groups—6.41%, 3.49%, and 1.67% (gross) or 6.38%, 3.43%, and 1.54% (net). Among all 3×3 groups, the A1-B3 group has the highest alpha.

In short, we find that SMF of funds' have significant impact on their performance. Funds that hold more stocks that appear on social media and funds by which the stocks held are more heavily discussed by investors or receive more bullish opinions perform better contemporaneously. Meanwhile, the different dimensions of a fund's SMF have different levels of influence on its performance. Bullishness seems to outweigh the total number of messages and the number of SM stocks.

3.2 Social Media and Mutual Fund Flows

In this sub-section, we employ regression analysis to test whether social media activities influence investors' fund investment decisions, and if yes, whether the impact is greater for individual investors, who are more subject to investor sentiment and limited attention than institutional investors.

3.2.1 The impact of SMF on fund flows

Table 4 reports the results investigating the influence of fund's SMF on its contemporaneous monthly flow. Fund flow is measured as the growth rate of total net asset, after adjusting for the appreciation of the fund's assets.[13] Panel A reports the simple regression results of monthly fund flows regressed on the four SMF variables, respectively. After accounting for the effect of a vector

[13] The monthly fund flow is calculated as $\frac{TNA_{i,t} - TNA_{i,t-1} * (1 + ret_{i,t})}{TNA_{i,t-1} * (1 + ret_{i,t})}$.

of control variables, we find that the number of SM stocks held by a fund, the total number of messages, and the *Bull-Bear Score* are all significantly positively associated with the same-month fund flows. The net monthly mutual fund inflow increases by 0.0013 for holding one more social media stock. Fund flow also significantly increases if stocks held by the fund have more mentions on social media. The positive effect of the average *Bull-Bear Score* is strongly significant, but the effect of the holding-adjusted *Bull-Bear Score* is inconclusive (comparing panel A and B). Overall, investors favour funds holding more SM stocks and/or with more bullish SM sentiment, while the weight of such stocks within the fund's AUM is less important.

Panel B investigates whether the effect of the SMF variables varies in the cross-section of mutual funds with different sizes. We include the interaction terms of the SMF variables and fund size into the regression. The findings are consistent with those in Panel A. The insignificant coefficients in Column (5) implies that the number of SM stocks held by funds do not have varying effects on flows of funds with different sizes. The negative coefficient of the interaction term in Column (6) shows that the positive effect of total number of messages on fund flows is stronger among smaller funds. The effect of the *Bull-Bear Score* is also more pronounced and greater in magnitude for smaller funds.

We also employ the next-month's fund flows as the dependent variable and replicate the above analysis. Results are qualitatively similar, although the effects of SMF variables are slightly smaller in economics terms. [14]

In contrast, Fang et al. (2014) find no significant relationship between the traditional media coverage of fund holdings and fund flows. While investors unilaterally obtain information and

[14] These results are available upon request.

opinions from the mass print media, there is much more involvement and interaction on social media, through which investors' excessive attention and biased opinions on SM stocks are reflected in the increase of fund flows.

[Insert Table 4]

3.2.2 Individual investor vs. Institutional investor

We further investigate whether the effect of social media activities on fund flows varies between individual and institutional investors. Compared with institutional investors, individual investors are more subject to limited attention (Barber and Odean, 2008) and should be more exposed to the biased belief arising from social media messages. Therefore, we posit that the effect of SMF variables on the monthly mutual fund flows would be stronger among funds that are mostly held by individual investors. To test our prediction, we create a dummy variable *Indiv* that equals 1 if over 80% of fund assets are categorised as individual (retail) share classes and 0 otherwise. We add both *Indiv* and its interaction term with the SMF variables into our regression analysis. Table 5 presents the results. Panel A and Panel B respectively shows the contemporaneous and predictive relationship between SMF variables and fund flows, conditional on investor types.

Comparing with Panel B of Table 4, Panel A of Table 5 shows that the effect of SMF variables on contemporaneous fund flows remains robust. The coefficients of *Indiv* are significantly negative in Panel A except in Column (1), indicating that individual-investor-dominated mutual funds generally have lower monthly flows. The insignificant coefficient of the interaction term in Column (1) shows that the positive impact of the number of SM stocks held by a fund on flows are not significantly different between individual-dominated and institution-dominated mutual funds. The significant and positive coefficients of the interaction terms in Column (2) to (4) imply that the total number of messages and the *Bull-Bear Score* of social media stocks have a greater impact on individual investors' fund investment decisions than institutional investors. The findings in Table 5 are consistent with those in Table 4 in that the effects of SMF variables tend to be more pronounced among small funds, which tend to have greater fraction of individual shares. We also conduct sub-sample regressions by separating the individual and institutional investors' mutual fund shares. Our conclusions regarding the varying impact of social media on the fund investment decisions of different investor types remain consistent. [15]

Panel B reports the relationship between the one-month-ahead mutual fund flows and the SMF variables. Overall, the effect of SMF variables on the next-period fund flows seems to be weaker than that on the contemporaneous flows, as reflected by the statistical significance of their coefficients. Consistent with Panel A, the total number of messages and *Bull-Bear Score* still have a greater impact on individual investors' fund investment decisions than institutional investors', although the results are less significant than those for the contemporaneous flows.

[Insert Table 5]

3.3 Social Media and Fund Trading

In this sub-section, we examine our third hypothesis — fund managers' trading decisions are affected by the fund holdings' social media activeness/bullishness. We calculate the total buy and sell of each SM stock in each active fund at the quarterly frequency and scale them by the lagged fund TNA. The resulting measures are employed as the dependent variables.

[15] These results are available upon request.

$$\frac{\$Buy_{i,k,q}}{TNA_{i,q}} = \frac{prc_{k,q} * (shr_{i,k,q} - shr_{i,k,q-1})}{TNA_{i,q}}, if\ shr_{i,k,q} \geq shr_{i,k,q-1}; \quad (2A)$$

$$\frac{\$Sell_{i,k,q}}{TNA_{i,q}} = -\frac{prc_{k,q} * (shr_{i,k,q} - shr_{i,k,q-1})}{TNA_{i,q}}, if\ shr_{i,k,q} < shr_{i,k,q-1}. \quad (2B)$$

where $prc_{k,t}$ is stock k's price at the end of quarter q, $shr_{i,k,t}$ and $shr_{i,k,t-1}$ are fund i's holdings in stock k at the end of quarters q and *q-1*. The SMF variables of interest in this test include a stock's total number of messages and the *Bull-Bear Score*. We also construct a dummy variable *Coverage rank* which equals 1 if the stock's total number of messages is above the median in the quarter, and 0 otherwise. Following the framework of Fang et al. (2014), we include the following stock characteristics as control variables for the fund managers' buys and sells of the SM stocks — size (market cap), squared market cap, past returns, turnover, B/M ratio, and age. Table 6 report the regression results. The dependent variable of regressions in Panel A (B) is manager's buy (sell) volume relative to the fund size, as shown in equation 2A (2B).

Both Column (1) and (2) show that fund managers' buy volumes are greater for stocks with greater number of mentions (i.e., messages) on social media or higher coverage rank. Similar patterns have been found for the selling volumes of stocks, as shown in Column (4) and (5). As we assume that social media mentioning (coverage) is a proxy of investor attention, we conclude that social media attention on a particular stock increases fund managers' both buying and selling amount of that stock. The effect of *Bull-Bear Score* is marginally significant for fund managers' buying behaviour but insignificant for selling. Manager's buying volume of a SM stock relative to fund size is greater for stocks with higher social media bullishness. These results remain qualitatively and quantitatively similar when we employ the *next-quarter* buy or sell volumes as the dependent variable. As the impact of a stock's *social media activeness/bullishness* on managers' trading behaviour does not significantly differ for buys and sells, there is little evidence to support the sub-hypothesis that managers are constrained by limited attention and thus tend to be the net purchasers of the most recent *social media active/bullish stocks*, which are believed to be more likely to attract investor attention than general stocks. Our results are more in line with the sub-hypothesis that managers cater to investors' preference of the recent *social media active/bullish* stocks by actively trading such stocks to attract flows.

[Insert Table 6]

4. Conclusions

We show that mutual funds with higher SMF measures, namely, funds holding more stocks that are mentioned on social media, those having more messages about the stocks they hold, or those with higher level of aggregate bullishness perform significantly better in the short term. The outperformance, however, is short-lived and generally disappears in the medium term. The short-term outperformance is likely to be driven by excessive attention and higher investor sentiment.

We also confirm that investors' decision to buy a fund is significantly affected by a fund's SMF. Fund flows are greater in the current and next month for funds with higher monthly SMF measures, after controlling for fund characteristics and contemporaneous returns. The positive effects of the total number of social media mentions and aggregate bullishness of fund holdings on fund flows are more pronounced among smaller funds and individual-investor-dominated funds.

A stock's social media activeness, as measured by the stock's number of messages, increases fund managers' both buying and selling amount of that stock. The effect of social media bullishness is marginally significant for fund managers' buying behaviour but insignificant for selling. As the impact of social media variables on managers' trading behaviour does not significantly differ for buys and sells, there is little evidence to support the hypothesis that managers are constrained by limited attention. Managers cater to investors' preference of the recent *social media active/bullish stocks* by actively trading such stocks to attract flows.

Overall, the various aspects of social media have significant impact on mutual fund performance, flows, and trading. There is strong evidence that investors' decision to buy a fund is at least partly driven by excessive attention and/or higher sentiment towards social media stocks that a fund holds, and fund managers take advantage of investor biases to boost short-term performance and/or attract flows.

References

Agarwal, V., Jiang, L., & Wen, Q. (2020). Why do mutual funds hold lottery stocks?. *Journal of Financial and Quantitative Analysis*, 1-47.

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

Bali, T. G., Cakici, N., & Whitelaw, R. F., 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, 99, 427-446.

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.

Brown, J. R., Ivkovic, X., Smith, P. A., & Weisbenner, S. (2008). Neighbors Matter: Causal Community Effects and Stock Market Participation. *Journal of Finance*, 63(3), 1509–1531.

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

Chen, H., De, P., Hu, Y. J., & Hwang, B.-H. (2014) Wisdom of crowds: the value of stock opinions transmitted through social media. *Review of Financial Studies*, 27(5), 1367-1403.

Chen, H., B.H. Hwang, & B. Liu (2019) The Emergence of 'Social Executives' and Its Consequences for Financial Markets. *Working paper*. *Available at SSRN 2318094.*

Cooper, M. J., Gulen, H., & Rau, P. R. (2005). Changing names with style: Mutual fund name changes and their effects on fund flows. *The Journal of Finance*, *60*(6), 2825-2858.

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.

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. (2020). The democratization of investment research and the informativeness of retail investor trading. *Working paper*. *Available at SSRN 3222841*.

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

Han, B., Kumar, A., (2013). Speculative retail trading and asset prices. *Journal of Financial and Quantitative Analysis*, 48, 377-404.

Hillert, A., Jacobs, H., & Müller, S. (2014). Media makes momentum. *Review of Financial Studies*, 27, 3467–3501.

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

Hirshleifer, D. A., Peng, L., & Wang, Q. (2021). Social Networks and Market Reactions to Earnings News. *Available at SSRN 3824022*.

Hong, H., Kubik, J. D., & Stein, J. C. (2004). Social interaction and stock-market participation. *Journal of Finance*, 59, 137–163.

Ivkovic, Z., & Weisbenner, S. (2007). Information Diffusion Effects in Individual Investor's Common Stock Purchases: Covet They Neighbors' Investment Choices, *Review of Financial Studies*, 20(4), 1327–1357.

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?. *The Journal of Finance*, *64*(4), 1889-1933.

Liu, S. and J. Han (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. *The Review of Financial Studies*, *27*(6), 1797-1829.

McLean, R. D., & Zhao, M. (2014). The business cycle, investor sentiment, and costly external finance. *The Journal of Finance*, *69*(3), 1377-1409.

Ozsoylev, H. N., Walden, J., Yavuz, M. D., & Bildik, R. (2014). Investor networks in the stock market. *Review of Financial Studies*, 27(5), 1323–1366.

Pool V., N. Stoffman, and S. E. Yonker (2015). The People in Your Neighborhood: Social Interactions and Mutual Fund Portfolios. *Journal of Finance*, 70 (6), 2679-2732.

Rakowski, D., Shirley, S. E., & Stark, J. R. (2021). Twitter activity, investor attention, and the diffusion of information. *Financial Management*, *50*(1), 3-46.

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.

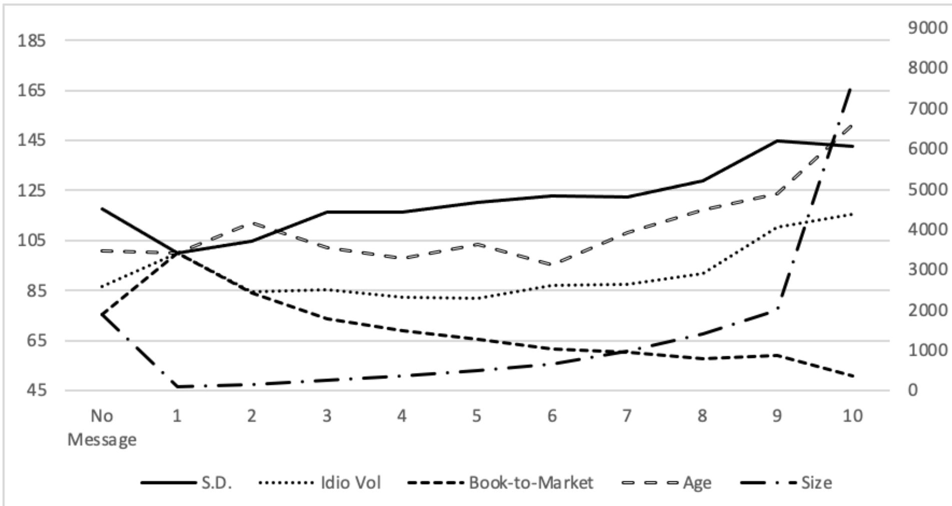
Sirri, E. R., & Tufano, P. (1998). Costly Search and Mutual Fund Flows. *Journal of Finance*, 53, 1589-1622.

Appendix A List of Variable Definitions

Variable Name	**Description**
Number of SM Stock	The natural log of total number of social media stocks (a stock with at least one message) held by fund *i* in month *t*.
Total Number of Messages	The natural log of total number of messages for social media stocks held by fund *i* in month *t*.
Bull-Bear Score (average)	The average of fund *i*'s *Bull-Bear Score* from holding social media stocks in month *t*.
Bull-Bear Score (holding-adjusted)	The measure of fund *i*'s *Bull-Bear Score* weighted by the fraction of holdings of social media stocks in month *t*.
Flow	The growth rate of the total net asset of fund *i* in month *t*, after adjusting for the appreciation of the fund's assets, calculated as: $\frac{TNA_{i,t} - TNA_{i,t-1} * (1 + ret_{i,t})}{TNA_{i,t-1} * (1 + ret_{i,t})}$
Size	The natural log of fund *i*'s total net asset in month *t*.
12b-1	The 12b-1 fee [actual_12b1] of fund *i* in month *t*
Load	The sum of front-end and rear-end load of fund *i* in month *t*.
Return	The monthly gross return of fund *i* in month *t*, calculated as the monthly net return [mret] plus the monthly expense ratio.
Expense ratio	The expense ratio of fund *i* in month *t*, calculated as the nearest available annualized expense ratio [exp ratio] divided by 12.
Turnover ratio	The turnover ratio [turn_ratio] of fund *i* in month *t*.
Management fee	The management fee [mgmt_fee] of fund *i* in month *t*.
Indiv	A dummy variable which equals 1 if over 80% of fund *i*'s assets are categorised as individual (retail) share classes.
Coverage rank	A dummy variable which equals 1 if stock *k* has social media message number above the median in the quarter *q*, and 0 otherwise.
Bull-Bear Score	The sentiment of social media stock *k* is calculated as: $\frac{\text{Number of Bull Messages} - \text{Number of Bear Messages}}{\text{Number of Bull Messages} + \text{Number of Bear Messages}}$

Figure 1: Social Media Stock Characteristics

(a) Sorted by total number of messages



(b) Sorted by the *Bull-Bear Score*

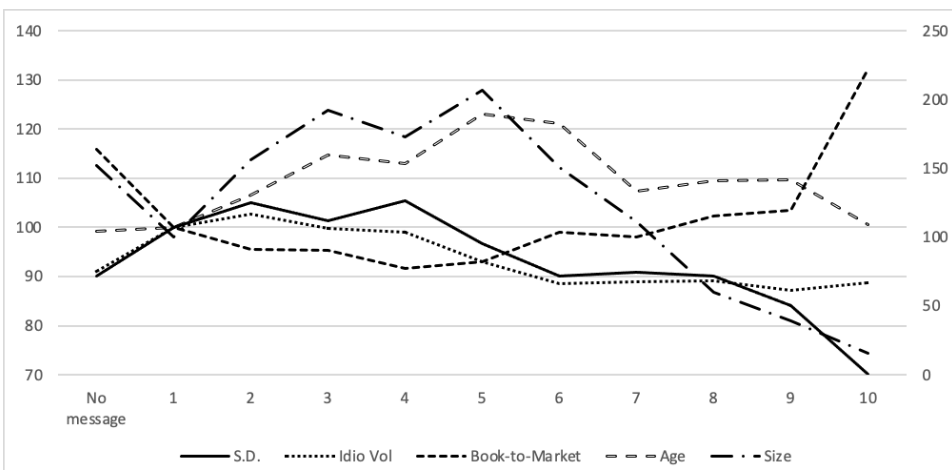


Table 1: Summary Statistics of Funds

This table presents the summary statistics for mutual funds in our sample. Panel A reports descriptive statistics for common fund characteristics, such as fund size, expense ratio, management fee, fund flows, etc. Panel B reports descriptive statistics for the funds' social media features.

Variable	Obs	Mean	Std. Dev.	Min	Max
Panel A. Fund Characteristics (monthly)					
Fund Size ($ Million)	69,672	4.152	16.255	0.001	498.117
Common Stock (%)	69,672	86.778	7.70754	80	99.98
12b-1	69,672	0.002	0.0021	0	0.01
Expense ratio	69,672	0.010	0.0038	0.001	0.023
Load	69,672	0.00	0.012	0	0.084
Management Fee (%)	69,672	0.682	0.263	-0.021	1.368
Turnover Ratio	69,672	0.617	0.5373	0.02	2.94
Return (gross)	69,672	0.009	0.041	-0.274	0.2772
Fund Flow	69,671	-0.006	0.028	-0.120	0.116
Panel B. Social Media Feature (monthly)					
Number of SM Stock	69,672	41.407	77.771	1	1379
Total Number of Messages	69,672	32838	57180.67	1	1169074
SM Stock Number (ln)	69,672	3.197	0.947	0.693	7.229
Total Message Number (ln)	69,672	9.335	1.595	0.693	13.972
Bull-Bear Score (Average)	69,672	0.457	0.116	-1	1
Bull-Bear Score (Holding-adjusted)	69,672	0.139	0.062	-0.667	1

Table 2: Fund Performance (Single Sort)

This table presents the summary statistics of the one-way sort of mutual funds' monthly abnormal returns on the funds' social media feature variables. We sort the mutual fund into quintiles based on the number of SM stocks held by the funds, the total number of messages, the average and holding-adjusted *Bull-Bear score*, respectively. Panel A summarize the contemporaneous abnormal returns. Panel B and Panel C report the results for abnormal returns in 6 and 12 months, respectively. The mean abnormal return, significance level and the standard errors are presented for each quintile and for the differential of the top and bottom quintiles. The monthly abnormal return (Alpha) is in percentage and adjusted against the Carhart (1997) four factors. Alphas are marked with *, **, or *** for the significance level of 10%, 5%, and 1%, respectively.

	(a)		(b)		(c)		(d)	
	No. of Social Media Stock		**Total No. of Message**		**Bull-Bear Score (average)**		**Bull-Bear Score (holding-adjusted)**	
	Alpha in %	S.E.	Alpha in %	S.E.	Alpha in %	S.E.	Alpha in %	S.E.
Panel A: Alpha at *t*								
Q1	-0.1484***	(0.0241)	-0.1329***	(0.0243)	-0.3809***	(0.0205)	-0.2886***	(0.0212)
Q2	-0.1200	(0.0183)	-0.0145	(0.02)	-0.1404***	(0.0167)	-0.1558***	(0.0167)
Q3	-0.0922	(0.0186)	-0.0100	(0.0164)	-0.0379**	(0.0186)	-0.0599***	(0.0195)
Q4	-0.0382	(0.016)	0.0119*	(0.0136)	0.0549***	(0.02)	-0.0395**	(0.0183)
Q5	0.0093**	(0.0121)	-0.0210	(0.0107)	0.2341***	(0.0252)	0.1515***	(0.0234)
Q5-Q1	0.1577***	-0.027	0.1119***	(0.0266)	0.615***	(0.0324)	0.4401***	(0.0315)
Panel B: Alpha at *t*+6								
Q1	-0.1637***	(0.0262)	-0.1877***	(0.0267)	-0.0721***	(0.0178)	-0.1238***	(0.0188)
Q2	-0.1328***	(0.0193)	-0.0484**	(0.0213)	-0.0438***	(0.0181)	-0.0951***	(0.0148)
Q3	-0.0998***	(0.0203)	-0.0663***	(0.0185)	-0.1095***	(0.0186)	-0.0865***	(0.0172)
Q4	-0.0744***	(0.0169)	-0.0120	(0.0159)	-0.101***	(0.0194)	-0.0495***	(0.019)
Q5	-0.0187	(0.0125)	-0.0352***	(0.0117)	-0.1194***	(0.0243)	-0.1098***	(0.0238)
Q5-Q1	0.1450***	-0.029	0.1525***	(0.0291)	-0.0473	(0.0301)	0.0140	(0.0303)
Panel C: Alpha at *t*+12								
Q1	-0.0548***	(0.0214)	0.0129	(0.0205)	-0.0207	(0.0223)	0.0239	(0.0187)
Q2	-0.0184	(0.0186)	-0.0016	(0.0213)	0.0021	(0.0167)	-0.0132	(0.0159)
Q3	-0.0228	(0.0183)	-0.0376*	(0.019)	0.0117	(0.0167)	-0.0387***	(0.0160)
Q4	-0.0143	(0.0153)	-0.0418***	(0.0142)	-0.0297	(0.0196)	-0.0507***	(0.0172)
Q5	0.0080	(0.0111)	-0.0041	(0.0114)	-0.0862***	(0.0241)	-0.0416*	(0.0255)
Q5-Q1	0.0628***	(0.0241)	-0.0169	(0.0235)	-0.0655**	(0.0328)	-0.0655**	(0.0317)

Table 3: Fund Performance (Double Sorts)

This table presents the summary statistics of the two-way sort of mutual funds' monthly abnormal returns on the funds' social media feature variables. We first sort the funds into terciles based on their holding-adjusted *Bull-Bear score* and then further sort each bin into terciles by the number of SM stocks held (Panel A) or by the total number of messages (Panel B). The mean abnormal return, significance level and the standard errors are presented for each quintile and for the differential of the top and bottom quintiles. The monthly abnormal return (Alpha) is in percentage and adjusted against the Carhart (1997) four factors. Alphas are marked with *, **, or *** for the significance level of 10%, 5%, and 1%, respectively.

Panel (a) *Bull-Bear Score* (holding-adjusted)

		B1		B2		B3		B3-B1	
	Gross Returns	**Alpha**	**t-stat**	**Alpha**	**t-stat**	**Alpha**	**t-stat**	**Alpha**	**t-stat**
Number of SM Stock	A1	-0.2671	-9.16	-0.1534	-5.63	0.0464	1.40	0.3135	***7.20***
	A2	-0.2798	-10.65	-0.0775	-3.07	0.0758	2.62	0.3556	***8.99***
	A3	-0.1350	-7.98	-0.0285	-1.53	0.0833	3.29	0.2183	***7.14***
	A3-A1	0.1320	***3.93***	0.1249	***3.79***	0.0369	0.88		
		B1		B2		B3		B3-B1	
	Net Returns	**Alpha**	**t-stat**	**Alpha**	**t-stat**	**Alpha**	**t-stat**	**Alpha**	**t-stat**
Number of SM Stock	A1	-0.3665	-12.34	-0.2531	-9.16	-0.0596	-1.76	0.3068	***6.94***
	A2	-0.3713	-14.05	-0.1690	-6.67	-0.0222	-0.76	0.3490	***8.73***
	A3	-0.2061	-11.88	-0.1042	-5.51	-0.0060	-0.24	0.2001	***6.45***
	A3-A1	0.1603	***4.67***	0.1489	***4.45***	0.0536	1.26		

Table 3 (continued)

Panel (b) *Bull-Bear Score* (holding-adjusted)

		B1		B2		B3		B3-B1	
	Gross Returns	**Alpha**	**t-stat**	**Alpha**	**t-stat**	**Alpha**	**t-stat**	**Alpha**	**t-stat**
Total Number of Messages	A1	-0.3138	-9.92	-0.1211	-3.63	0.2207	6.53	0.5345	***11.57***
	A2	-0.2109	-10.56	-0.0185	-0.82	0.0796	3.06	0.2905	***9.02***
	A3	-0.1054	-6.20	0.0088	0.55	0.0334	1.71	0.1388	***5.08***
	A3-A1	0.2084	***5.95***	0.1299	***3.48***	-0.1873	***-4.90***		
		B1		B2		B3		B3-B1	
	Net Returns	**Alpha**	**t-stat**	**Alpha**	**t-stat**	**Alpha**	**t-stat**	**Alpha**	**t-stat**
Total Number of Messages	A1	-0.4147	-12.98	-0.2215	-6.69	0.1165	3.38	0.5313	***11.36***
	A2	-0.3026	-15.06	-0.1104	-4.82	-0.0172	-0.65	0.2854	***8.80***
	A3	-0.1804	-10.31	-0.0672	-4.18	-0.0524	-2.65	0.1280	***4.60***
	A3-A1	0.2344	***6.60***	0.1543	***4.12***	-0.1690	***-4.36***		

Table 4: Fund Flow Analysis

Panel A of this table presents the results for monthly fund flows regressed on the funds' social media feature variables, namely the number of SM stocks, the total number of messages, the average and holding-adjusted *Bull-Bear score*, and a vector of control variables. Panel B includes the interaction term of social media variables and fund size in the regressions. The p-values are reported in parentheses. We control for the month and fund fixed effects. The standard errors are clustered at the fund level. Coefficients are marked with *, **, or *** for the significance level of 10%, 5%, and 1%, respectively.

	Panel A				Panel B			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of SM stock (ln)	0.0033***				0.0037***			
	(0.000)				(0.001)			
Number of SM stock (ln) * Size					-0.0005			
					(0.508)			
Total number of messages (ln)		0.0014***				0.0023***		
		(0.000)				(0.000)		
Total number of messages (ln) * Size						-0.0014***		
						(0.000)		
Bull-Bear Score (average)			0.0057***				0.0079***	
			(0.000)				(0.000)	
Bull-Bear Score (average) * Size							-0.0036***	
							(0.005)	
Bull-Bear Score (holding-adjusted)				0.0064				0.0164***
				(0.146)				(0.001)
Bull-Bear Score (holding-adjusted) * Size								-0.0164***
								(0.001)
size	0.0041*	0.0043*	0.0044**	0.0044**	0.0061	0.0198***	0.0062**	0.0072***
	(0.062)	(0.051)	(0.047)	(0.047)	(0.135)	(0.000)	(0.011)	(0.003)
12b-1	-1.6193***	-1.5894***	-1.5750***	-1.5813***	-1.6323***	-1.7197***	-1.5824***	-1.5990***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.004)	(0.003)	(0.006)	(0.006)
load	-0.0131	-0.0123	-0.0130	-0.0126	-0.0130	-0.0109	-0.0130	-0.0124
	(0.291)	(0.326)	(0.301)	(0.314)	(0.294)	(0.386)	(0.300)	(0.320)
return	0.0724***	0.0716***	0.0695***	0.0714***	0.0724***	0.0717***	0.0697***	0.0720***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
expense ratio	0.3214	0.3063	0.2870	0.2929	0.3495	0.5315	0.2976	0.3365
	(0.417)	(0.438)	(0.468)	(0.459)	(0.380)	(0.177)	(0.451)	(0.393)
turnover ratio	0.0003	0.0004	0.0004	0.0004	0.0003	0.0006	0.0004	0.0004
	(0.726)	(0.656)	(0.702)	(0.695)	(0.723)	(0.543)	(0.693)	(0.680)
management fee	-0.0092**	-0.0087*	-0.0085*	-0.0085*	-0.0094**	-0.0104**	-0.0085*	-0.0087**
	(0.039)	(0.055)	(0.057)	(0.058)	(0.036)	(0.022)	(0.055)	(0.050)
constant	-0.0052	-0.0067	0.0021	0.0038	-0.0066	-0.0183***	0.0009	0.0016
	(0.352)	(0.245)	(0.648)	(0.419)	(0.290)	(0.002)	(0.847)	(0.733)
Month FE	YES	YES	YES	YES	YES	YES	YES	YES
Fund FE	YES	YES	YES	YES	YES	YES	YES	YES
Obs	69671	69671	69671	69671	69671	69671	69671	69671
R^2	0.0275	0.0274	0.0271	0.0268	0.0275	0.0294	0.0273	0.0273
N_clust	1105	1105	1105	1105	1105	1105	1105	1105

Table 5: Fund Flow Analysis of Individual and Institutional Investors

This table present the effect of social media stocks on fund flow among individual and institutional investors. Panel A presents the results of contemporaneous mutual fund monthly flow regressed on the funds' SMF variables, the interaction terms of individual investor dummy and SMF variables, and a vector of control variables. The SMF variables include the number of SM stocks, the total number of messages, the average and holding-adjusted *Bull-Bear Score*. *Indiv*, the individual investor dummy, equals 1 if over 80% of the fund assets are categorized as individual share classes and 0 otherwise. Panel B presents the predictive regression results of the next-month fund flows. The standard errors are reported in parentheses. We control for the month and fund fixed effects. The standard errors are clustered at the fund level. Coefficients are marked with *, **, or *** for the significance level of 10%, 5%, and 1%, respectively.

	Panel A. Contemporaneous Analysis				**Panel B. Predictive Analysis**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of SM stock (ln)	0.0032***				0.0033***			
	(0.0011)				(0.0011)			
Number of SM stock (ln) * Indiv	0.0002				-0.0001			
	(0.0011)				(0.0011)			
Total number of messages (ln)		0.0009**				0.0007		
		(0.0005)				(0.0005)		
Total number of messages (ln) * Indiv		0.0009**				0.0008*		
		(0.0004)				(0.0004)		
Bull-Bear Score (average)			0.0033**				0.0031**	
			(0.0015)				(0.0015)	
Bull-Bear Score (average) * Indiv			0.0061**				0.0047*	
			(0.0026)				(0.0025)	
Bull-Bear Score (holding-adjusted)				0.0017				0.0027
				(0.0055)				(0.0057)
Bull-Bear Score (holding-adjusted) * Indiv				0.0149*				0.0112
				(0.0078)				(0.0076)
Indiv	-0.0044	-0.0127***	-0.0066***	-0.0058***	0.0000	-0.0078*	-0.0024	-0.0019
	(0.0039)	(0.0042)	(0.0016)	(0.0016)	(0.0040)	(0.0042)	(0.0016)	(0.0016)
Control Variables	YES	YES	YES	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES	YES	YES	YES
Fund FE	YES	YES	YES	YES	YES	YES	YES	YES
Obs	69671	69671	69671	69671	68515	68515	68515	68515
R^2	0.0281	0.0282	0.0278	0.0275	0.0388	0.0387	0.0385	0.0382
N_clust	1105	1105	1105	1105	1101	1101	1101	1101

Table 6: Funds' Buys and Sells of Social Media Stocks

The dependent variable of regressions in Panel A (B) is manager's buy (sell) volume relative to the fund size. Specifically, they are calculated as the total buy and sell of each SM stock in each active fund at the quarterly frequency scaled by the lagged fund TNA. All independent variables are measured at the stock level. Coverage rank is a dummy variable which equals 1 if the stock's total number of messages is above the median in the quarter, and 0 otherwise. Control variables include some stock characteristics. The p-values are reported in parentheses. We control for quarter and fund fixed effects. The standard errors are clustered at the stock level. Coefficients are marked with *, **, or *** for the significance level of 10%, 5%, and 1%, respectively.

	Panel A: $Buy / TNA			Panel B: $Sell / TNA		
	(1)	(2)	(3)	(4)	(5)	(6)
Total number of message (ln)	0.0050*			0.0081***		
	(0.067)			(0.001)		
Coverage rank		0.0104***			0.0121***	
		(0.007)			(0.000)	
Bull-Bear Score			0.0059*			-0.0018
			(0.089)			(0.561)
Stock Size	-0.0415	-0.0547**	-0.0474*	-0.0263	-0.0461**	-0.0390**
	(0.127)	(0.038)	(0.076)	(0.202)	(0.019)	(0.048)
Stock Size Squared	0.0035***	0.0040***	0.0038***	0.0020***	0.0028***	0.0026***
	(0.000)	(0.000)	(0.000)	(0.004)	(0.000)	(0.000)
Past Return	0.0036***	0.0036***	0.0035***	0.0028***	0.0028***	0.0029***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Stock Turnover	0.0092***	0.0103***	0.0109***	0.0044***	0.0067***	0.0072***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Stock B/M	0.0036	0.0035	0.0037	-0.0129**	-0.0132***	-0.0133***
	(0.518)	(0.530)	(0.509)	(0.011)	(0.010)	(0.009)
Stock Age	-0.0005	-0.0006	-0.0006	-0.0003*	-0.0004*	-0.0004*
	(0.129)	(0.113)	(0.122)	(0.082)	(0.057)	(0.064)
Constant	0.1474	0.2113	0.1630	0.1965	0.1686	0.1268
	(0.805)	(0.266)	(0.396)	(0.292)	(0.254)	(0.393)
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Clustered by Stock	Yes	Yes	Yes	Yes	Yes	Yes
N	233901	233901	233901	199373	199373	199373
R^2	0.0687	0.0674	0.0598	0.0572	0.0647	0.0658
N_clust	2070	2070	2070	1872	1872	1872