Short Sellers in the Realm of Social Media: Arbitrageurs or Manipulators?

Guilong Cai, R. David McLean, Tianyu Zhang, Mengxin Zhao

April 12, 2021

Abstract

Social media can change how short sellers impact stock prices. We study 75.1 million investment-related social media posts for 3,683 unique Chinese firms. Prior to high short interest, social media tone is abnormally positive. Once highly shorted, the tone flips and is abnormally negative. No such pattern exists with traditional media. Compared to firms that are just highly shorted, highly shorted firms with rise-then-fall patterns in social media tone have abnormal returns that are 1.8x higher before, and 3.3x lower after, the initiation of high short interest. Evidence from natural experiments involving China's introduction and subsequent suspensions of shorting also suggest social media manipulation. Our findings show that in the realm of social media, short sellers may profit more by creating mispricing than by correcting it.

In this paper, we study the interaction between short sellers and social media and the impact on stock prices. One hypothesis is that the combination of short sellers and social media leads to more efficient markets. The finance literature typically assumes that short sellers are informed traders that have a stabilizing effect on prices. Diamond and Verecchia (1987) contend that because short selling is costlier than buying, short sellers are more likely to be informed investors. Many empirical studies support this argument, showing that high short interest predicts low stock returns (e.g., Dechow et al. (2001), Duan, Hu, and McLean (2010), Boehmer, Jones, and Zhang (2008), Engelberg, Reed, and Ringgenberg (2012), and Chang, Lou, and Ren (2014)). A number of studies also find that social media plays an informative role with respect to stock prices (e.g., Chen, De, Hu, and Hwang (2014), Bartov, Faurel, and Mohanram (2017), Gianini, Irvine, and Shu (2017), and Tang (2018)). It could therefore be the case that short sellers use social media to share their information, letting other investors know that prices are too high. This in turn could encourage informed trading that brings prices more in line with fundamentals.

Alternatively, short sellers could use social media to manipulate stock prices. Although academics generally find a positive role for short sellers, regulators have expressed concern that short sellers may manipulate prices. A recent and salient example is the U.S. SEC's investigation into whether the 2021 rise and fall of GameStop's stock price was encouraged by social media manipulation.[1] The SEC has previously charged short sellers with spreading false rumors, and in 2008 issued emergency disclosure rules to limit such activities.[2] Other regulatory bodies have expressed similar concerns. In 2011, the European Securities and Markets Authority stated:

[1] See here: https://www.bloomberg.com/news/articles/2021-02-03/sec-hunts-for-fraud-in-social-media-posts-that-drove-up-gamestop.

[2] See here: https://www.sec.gov/news/press/2008/2008-209.htm

"While short-selling can be a valid trading strategy, when used in combination with spreading false market rumors this is clearly abusive."[3] Social media can be a very effective tool to spread rumors. Consistent with this idea, Jia, Redigolo, Shu, and Zhang (2020) provide evidence that Twitter exacerbates speculative merger rumors. Thus, our alternative hypothesis is that short sellers use social media to manipulate prices, and thereby make markets less efficient.

To test these competing hypotheses, we turn to China, as testing for these effects in China has several advantages relative to other countries. We have access to unique social media data. Our proprietary dataset includes 75.1 million posts on Guba, covering 3,683 unique firms, during the period 2009 to 2018. Guba is one of the oldest and most influential social media platforms that focuses on the Chinese capital market. Guba is different than general social media platforms such as Twitter, which is not focused on capital market topics like Guba is. Another benefit of Guba for our analysis is that the posts are aggregated under a section specific for each firm, so we do not need to match posts to firms through ticker or firm name as with other social media platforms. To the best of our knowledge, ours is the largest sample of investment-related social media posts used in the finance literature to date. In addition, we have access to a traditional Chinese media dataset includes 2.01 million articles, covering 3,603 unique firms. We can therefore tell whether social media posts reflect or portend actual news, or instead consist more of rumors and noise, using traditional news as a benchmark.

As in the U.S, studies of Chinese short sellers generally find that short sellers are sophisticated investors and that high short interest portends low returns (e.g., Chang, Lou, and Ren (2014)). One advantage of studying short sellers in China is that shorting was not allowed

[3] See here: https://www.investmentnews.com/countries-ban-short-selling-in-short-order-38151

prior to March of 2010, when the Chinese government began a pilot program that introduced short selling. We can therefore study how social media and other factors change when a firm enters the short selling program. In addition, since short selling was introduced in China there was period during which regulators made security lending very costly, which effectively halted short selling. This period further serves as a natural experiment for us to study how social media changes when the ability to short sell changes.

We begin our study by examining how social media tone and traditional media tone evolve when a firm is targeted by short sellers. We find that there is a rise-then-fall pattern in social media tone, but not in traditional media tone, around periods of high short interest. During the 30 days before a firm is highly shorted, its social media tone is abnormally positive. Once the firm is highly shorted, its social media tone turns abnormally negative. We find no such pattern with traditional media tone.

We then study the effects of entering the short selling program. Once a firm is selected, shorting is allowed the next day. However, whether a firm will be selected or not is unknown to the public beforehand. We find that during the first week that a firm enters the program, the tone of its social media posts become abnormally negative if it is targeted by short sellers. Before entering the program, the same firm's social media tone was not abnormal. This pattern is sensible, as short sellers did not know beforehand that the firm would enter the program.

We compare the monthly number of social media posts and the monthly volatility of social media tone before and after firms enter the pilot program. We find that both the number of posts and the volatility of tone are higher after the firm is in the program. These effects are not observed with traditional media tone. These findings suggest an attempt to manipulate stock prices via social media. This is because stock prices are first manipulated up, and then down, with abnormally positive, and then abnormally negative, social media tone. Thus, there is more volatility in the tone and more posts as compared to when shorting was not allowed and there was no manipulation. Short selling was then effectively halted during the months of August 2015 and March 2016. During these months, both the number of social media posts and the volatility of tone decreased significantly for firms that were in the shorting program. Again, these effects are not observed with traditional media.

We then study the impact that social media tone and short interest have on stock returns. We find that they interact. Firms targeted by short sellers have rise-then-fall patterns in stock returns around the initiation of high short interest i.e., stock returns are abnormally high (low) before (after) the initiation of high short interest. These effects are greater if the shorted firms also have rise-then-fall patterns in social media tone around the initiation of high short interest. Stocks that are both highly shorted and have rise-then-fall patterns in social media tone around the initiation of high short interest have abnormal stock returns that are 181% higher before, and 333% lower after, the initiation of high short interest. This suggests that short sellers are adept at using social media to influence stock prices.

We consider the idea that highly shorted stocks with rise-then-fall patterns in social media tone did not have any manipulation. Instead, these stocks have impending bad news and short sellers trade ahead of this, as is shown in Engelberg, Reed, and Ringgenberg (2012). If this is the case, then in the subsequent period the low returns will occur primarily on days with traditional news. Yet we find that opposite: the abnormally low returns of highly shorted stocks with rise-then-fall patterns in social media tone occur on days *without* traditional news.

Our findings contribute to several branches of literature. With respect to short selling, we are the only paper that we know of to provide systematic evidence that short sellers manipulate stock prices. Virtually all papers that we know of show that short sellers make prices more efficient, be it in the U.S. (e.g., Dechow et al. (2001), Duan, Hu, and McLean (2010), Boehmer, Jones, and Zhang (2008)), China (e.g., Chang, Lou, and Ren (2014)), or around the world (Bris, Goetzmann, and Zhu (2007)). Our paper shows that short sellers can play a destabilizing role when in the presence of social media, which is an increasingly common form of media and communication.

Our paper has ramifications for asset pricing theory, as it shows that social media can change the way that sophisticated investors (arbitrageurs) impact prices. Most theories in finance assume that arbitrageurs have a stabilizing effect on prices.[4] This is the case both in classical finance, where markets are assumed to be efficient, (e.g., Freidman (1953)), and in most behavioral finance theories, in which equilibrium prices do not reflect fundamentals, but arbitrageurs' trades still make markets more efficient (e.g., Figlewski (1979), De Long, Summers, Shleifer and Waldman (1990a), Shleifer and Summers (1990) and Barberis and Thaler (2003)). None of these papers include a role for social media, which did not exist when most of them were written. Our findings show that when social media is added to the mix, arbitrageurs may make markets less efficient. One paper in this spirit is De Long, Summers, Shleifer and Waldman (1990b) who find that arbitrageurs further destabilize prices when in the presence of feedback

[4] Like Shleifer and Summers (1990) we think of markets consisting of two types of investors: arbitragers who are the "smart money" or "rational speculators" and everyone else.

traders.[5] Another is van Bommel (2003), who creates a model in which investors spread rumors, causing other investors to trade and prices to diverge from fundamentals. The rumormongers then trade on and profit from the mispricing. Our findings are consistent with this framework.

Our paper contributes to a nascent yet growing literature concerned with how social media impacts stock prices. As we mention above, a number of studies find that social media plays an informative role with respect to stock prices (e.g., Chen, De, Hu, and Hwang (2014), Bartov, Faurel, and Mohanram (2018), Gianini, Irvine, and Shu (2017), Tang (2018), and Farrell, Green, Russell, and Markov (2020)). Our study shows that social media can also do the opposite, and enable stock price manipulation. In this respect our paper is consistent with Jia, Redigolo, Shu, and Zhang (2020), who find that Twitter can exacerbate merger rumors. An important difference is that where our paper suggests manipulation, Jia et al. (2020) document rumors, which are noisy but not need not be purposefully manipulative.

Finally, our paper is relevant for regulation. As we mention above, earlier papers do show that social media can help with price discovery and make asset prices more informative. Our results show that social media can also play a dark role, and can be used by some investors to make abnormal profits at the expense of others. This finding is relevant in the discussion as to whether and how social media platforms should be monitored and regulated. Our findings suggest that if regulators are concerned about manipulation from either short sellers or social

[5] Feedback trading can be caused by extrapolative expectations about prices, trend chasing, and even stop-loss orders. De Long et al. point out that feedback trading is recognized as far back as Bagehot (1897) and that feedback trading is perhaps the most well-documented type of "noise trading". Feedback trading is perhaps the most well-documented behavioral finance bias (see Andraessen and Kraus (1988), Frankel and Froot (1988), Case and Shiller (1988) and Shiller (1988)). De Long et al. (1990b) are inspired by Soros (1987), who claims to have traded in this spirit during the 1960s conglomerate bubble and the 1970s REIT bubble.

media, then they should look for the combination of short interest surrounded by a rise-then-fall pattern in social media tone.

1. Data, Sample, and Variables

1.1. Social Media Data

We use posts on the Guba of East Money, which is one of the oldest and most influential social media platforms focusing on the capital market in China. We design an automatic crawler to get all the main posts, ignoring the replying posts, for each firm.[6] We require each firm to have at least three social media posts per day to avoid errors in our measurement of daily social media tone. Our social media dataset includes about 75.1 million posts, covering 3,683 unique firms over the period 2009 to 2018.

1.2. Traditional media data

The traditional media data is an updated version of that used in Piotroski et al. (2017). We use articles published in official newspapers and non-official newspapers focusing on financial and economic news. We collect data from Wisenews[7], a database archives all historical articles published by varieties of newspapers and magazines in Chinese. Our traditional media dataset includes about 2.01 million articles, covering 3,603 unique firms from 2009 to 2018.

1.3. Firm-Sample

[6] We filter the posts, such as news articles, that are automatically posted by the platform by tracing the hyperlink.
[7] https://www.wisers.com.cn/hk/home/index.html

We begin our sample with all available firm-day short interest for the period March 31, 2010 to December 31, 2018. This yields 1,248,302 observation. Short-selling was prohibited in China prior to March 31, 2010. We then drop 84,778 observations in financial industries and observations with missing data that are needed to construct our main variables (38,317 observations). Our final sample includes 1,125,207 firm-day observations, with 1,013 unique firms.

1.4. Variables

The primary variables in this paper are concerned with the measurement of short selling and the tone of the posts from social media and news articles from traditional media. With respect to social media, we utilize machine learning techniques to construct the tones of news articles from traditional media and posts from social media. The resulting data are also used in Wang, Wong and Zhang (2021). A team of research assistants, including undergraduate and postgraduate business school students labeled the tone of each sentence of 50,000 articles randomly picked from our sample as negative, positive, and neutral. Using these manually labeled training materials, we train a support vector machine (SVM) model to classify each sentence into positive, neutral, or negative and check the out-of-sample classification accuracy using a subset of manually labeled sentences that the model has not seen. The out-of-sample validation using 10,000 randomly selected sentences shows that the accuracy rate of our model is above 90%.

The tone of the article is measured by the relative weight of positive sentences to negative sentences in the article. In addition, we also consider the importance of sentences from different positions within an article. We weigh the sentences from the first and last paragraphs as 2, the first and last sentences of the first and last paragraphs as 3, and other sentences from the article as 1. The tone of the article's body equals (#of positive sentences−# of negative sentences)/(#of positive sentences+# of negative sentences+1).[8] The overall tone of the article, in the end, is defined as (tone of text body*0.7+tone of title*0.3).

For the posts from Guba, we label the training set at post level rather than at sentence level because the post are normally short. In total we label 50,000 randomly selected posts as the training set and then classify all the posts into positive, negative and neutral using SVM. Given the linguistic feature, we also consider the emoji of social media in our modeling.

We measure short interest each day by taking the ratio of the number of shares shorted and dividing it by the number of shares outstanding. We rank all firms for which shorting is possible according to their short interest for each year. If at any trading day during the year that a firm crosses the 90 percentile, we refer to the day as highly shorted for the firm. In unreported tests we get similar findings using the 95th percentile as our cutoff. We have also used continuous short interest in our tests, and get similar findings. Table 1 shows that the mean value of short interest in our sample is 0.011 and the standard deviation is 0.016. For firms that are highly shorted the mean level of short interest is 0.047 and the standard deviation is 0.023.

[Insert Table 1 here]

[8] We put more weight on the title and certain sentences in the text following Njølstad et al. (2014) and Yang et al. (2014).

We also use several firm-level variables in our tests. These data are obtained from the Chinese Securities Market and Accounting Research (CSMAR). To control the firm fundamentals, we include firm size (*SIZE*), leverage (*LEV*), book-to-market ratio (*BM*), return on total assets (*ROA*), yearly stock returns (*PRERET*). We also include an indicator variable, *SOE*, to capture the political bias for state-owned enterprises. The construction of these variables is detailed in the appendix. We report the sample distribution by year and by industry in Table 2.

[Insert Table 2 here]

2. Main Results

In this section of the paper, we discuss our findings regarding the evolution of media tone around short selling. We test whether the evolution of several social media variables around the initiation of high short interest are consistent with manipulative trading. We also ask whether any patterns in social media are mirrored by similar patterns in traditional media. A positive correlation between social media and traditional media suggests that social media is reflecting actual news, rather than manipulative trading.

2.1. Short Selling and Social Media Tone

The regressions reported in Table 3 test how social media tone evolves around the initiation of high short interest. The unit of observation is firm-day. The dependent variable is the average daily social media tone measured over various periods. We measure social media tone over the intervals of *t-30* to *t* and *t-5* to t, and then *t+1* to *t+5* and *t+1* to *t+30*. We regress social media tone on a dummy variable equal to 1 if the firm becomes highly shorted on day *t*, and zero otherwise. We control for size, profitability, leverage, book-to-market, lagged stock returns, and whether the firm is a state-owned-enterprise. We also control for traditional media tone, measured over the same horizon as the social media tone. The regressions include firm and time fixed effects, and the standard errors are clustered on firm.

[Insert Table 3 here]

In the first regression, the dependent variable is the average social media tone over the period *t-30* to *t*. The coefficient for the day *t* high short interest dummy is 0.008 (*t*-statistic = 3.10). This shows that social media tone is abnormally positive over the 30 days before a firm becomes highly shorted. Regression 2 studies social media tone over days *t-5* to *t* and finds the same effect. With respect to economic significance, the mean of social media tone is -0.214 over the *t-30* to *t* period (see Table 1). The dummy variable in regression 1 thus shows that the social media tone is higher by about 3.7% relative to the mean, while in regression 2 the effect is 4.2% higher relative to *t-5* to *t* period's mean.

Regressions 3 and 4 study social media tone after the firm becomes highly shorted, over 5 -day and 30-day horizons, respectively. We find that the tone flips. In both specifications, the effect of being highly shorted is associated with abnormally negative social media tones. In regression 3, the coefficient is -0.005 (*t*-statistic = -3.89). This shows that the tone is abnormally low by about 2.3% relative to the mean over the 5 days after a firm becomes highly shorted. Regression 4 shows that the effect grows, and is 4.3% lower relative to the mean over the 30-day horizon. Hence, once short sellers target a stock there is a negative and statistically significant change in its social media tone.

The results in Table 3 also show that social media tone is more positive if the firm has high past stock returns and is a glamour stock (low book-to-market ratio), which is sensible. The traditional media tone coefficient is positive and significant in all specifications. This is also sensible, it shows that when the traditional news tone is more positive, social media tone is also more positive. It also shows that our results cannot be explained by social media tone simply reflecting traditional media tone, as traditional media tone is controlled for.

2.2. Short Selling and Traditional Media Tone

The results thus far show a rise-then-fall pattern in social media tone around high levels of short interest. That is, if a firm has a high level of short interest on day *t*, the tone of the social media concerning the firm was abnormally positive over the days leading up to and including day *t*, and then abnormally negative over the days following day *t*. This pattern is consistent with short sellers attempting to manipulate stock prices via social media. However, the pattern could also reflect social media participants discussing actual news about the firm. Perhaps the firm had good news before it was highly shorted and then bad news afterwards? Although we control for traditional media in our social media tests reported in Table 3, we explore this issue further here.

We re-estimate the regression reported in Table 3, however, we replace social media tone with traditional media tone as the dependent variable. If the same rise-then-fall pattern in social media tone around high short interest is not observed with traditional media tone, then it is unlikely that the social media tone patterns in Table 3 reflect actual news.

The first two regressions in Table 4 study traditional media tone over periods *t-30* to *t* and *t-5* to *t*. In both specifications, the high short interest coefficient is insignificant. Regressions 3 and 4 study the traditional media tone over the 5-day and 30-day periods after the firm is highly shorted. The high short interest coefficient is positive but insignificant over the 5-day horizon, and then negative and marginally significant over the 30-day horizon. The coefficient over the 30-day horizon is -0.009 (*t*-statistic = -1.67), showing that the traditional media tone is lower 2.8% over the period. This is not surprising, as highly shorted stocks are expected to have some bad news, however the effect is 35% less than that measured in Table 3 with social media tone over the same horizon, and the *t*-statistic is also much smaller (1.67 vs. 5.45 in Table 3). Moreover, the rise-then-fall pattern in social media tone observed in Table 3 is not observed in Table 4 with traditional media.

The control variables in Table 4 show that larger and more profitable firms have more positive traditional news tone. Table 3 shows that social media tone is unrelated to both of these variables. As with social media tone, traditional media tone is more positive for growth stocks and stocks with high past returns. Finally, as we saw in Table 3, traditional media tone and social media tone are positively correlated, as the social media tone coefficient is positive and significant in all of the specifications.

[Insert Table 4 here]

2.3. The Short Selling Pilot Program

Short selling began in China in March 2010 with a pilot program that allowed shorting of selected firms. Once a firm enters the program shorting is allowed, however which firms will be selected on which dates is not known ahead of time by market participants. In addition, during the months August 2015 to March 2016 regulators increased the cost of short selling significantly, which effectively halted short selling during these periods.[9] Thus, the introduction of short selling and this temporary stoppage serve as natural experiments with which we can test for the effects that shorting has on social media. Thus far our results suggest that short sellers use social media to manipulate stock prices, and these events allow us to further test this hypothesis.

2.3.1. The Short Selling Pilot Program and Social Media Tone

We begin by studying how a firm's social media tone changes once it enters the short selling pilot program. Our sample again consists of firm-day observations; however, the observations are limited to the first five days that a firm is in the shorting program. We expect that the abnormally positive social media tone prior to high short interest, documented in Table 3, will not appear for firms that are new entrants to the shorting program. This is because short sellers don't know which firms will enter the program, so they cannot manipulate via social media beforehand. We do though, expect to find the abnormally negative social media tone for highly shorted firms, as was documented in in Table 3.

We report the findings from these tests in Table 5. The first two regressions in Table 5 study social media over the periods *t-30* to *t* and *t-5* to *t*, where day *t* refers to the day that short interest is measured. Day *t* is limited to the first 5 days that the firm is in the short selling program. The results in Table 5 show that there is no difference in social media tone between firms that

[9] On August 4th, 2015, the exchange changes the trading rules for short sell from t+0 to t+1. The new rule was that the short seller cannot repay the borrowed shares on the same day, increasing the cost of short sell significantly. On the second day, majority of the brokers suspended their business of lending shares, which was restored until March 2016.

become highly shorted and firms that do not become highly shorted. This makes sense, as it is not known ahead of time that a firm will enter the program and that shorting will be allowed.

[Insert Table 5 here]

Columns 3 and 4 report the effects of high short interest on social media tone during the periods *t+1* to *t+5* and *t+1* to *t+30*. The results show that the social media tone becomes abnormally negative for firms that are targeted by short sellers. In the third column, which measures social media tone during the period *t+1* to *t+5*, the high short interest coefficient is -0.036 (*t*-statistic = -2.52). This reflects a 17.56% decline in social media tone relative to the mean value of social media tone reported in Table 1. The high short interest coefficient in the fourth column reflects a 15.24% decline in social media tone during the period *t+1* to *t+30*.

Taken together, the results in Table 5 show that once stocks become shortable and targeted by short sellers, their social media tone turns negative. This finding, taken together with those in Tables 3 and 4, are consistent with the idea that short sellers use social media tone to manipulate stock prices.

2.3.2. The Short Selling Pilot Program and Traditional Media Tone

Table 6 is like Table 5, only it studies traditional media tone. Our sample again consists of firm-day observations that are limited to the first five days that a firm is in the shorting program. In the first two columns, the high short interest coefficient is positive but insignificant in both the *t-5* to *t* window, and in the *t-30* to *t* window. Similarly, the high short interest coefficient is positive and insignificant in the *t+1* to *t+5* window, and negative but insignificant in the *t+1* to *t+30* window. Hence, unlike social media tone, traditional media tone does not change once a firm enters the program and becomes highly shorted. The abnormally negative social media effects for highly shorted firms reported in Table 5 therefore cannot be explained by social media tone reflecting news items in traditional media.

[Insert Table 6 here]

2.3.3. The Number of Social Media Posts and Volatility of Social Media Tone

Table 7 further studies the effects of entering the short selling pilot program. We now use monthly data, and study how social media and traditional media change when a firm enters the shorting program. Our tests so far suggest that short sellers may use social media to manipulate stock prices. If short sellers use social media to manipulate stock prices, then we would expect there to be more social media posts for firms in the program, and for the volatility of the tone of the posts to increase, as the pattern we have observed around shorting is abnormally positive tone followed by abnormally negative tone.

We measure the number of social media posts and the standard deviation of the posts' tone for each firm-month observation. In China, firms enter the shorting pilot program in different batches (staggered events), and there are still many firms that cannot be shorted. The sample in Table 7 includes all firm-month observations for all listed firms from 2010 January to 2018 December, including firms that never entered the program. We create a dummy variable equal to 1 if the firm-month is shortable and zero otherwise. Once a firm enters the program it tends to stay in, i.e., the dummy variable remains equal to 1. Our regressions include firm and time fixed effects, so technically we are estimating Difference-in-Difference models, i.e., the coefficient is the difference between the pre and post for firms that enter the pilot program, compared to firms that never enter the program.

The first regression in Table 7 shows that the number of social media posts increases once a firm enters the program. The dependent variable is the log of one plus the number of social media posts in month *t*. The coefficient for the shorting dummy is 0.084 (*t*-statistic = 4.16), showing that the number of posts is significantly higher once a firm enters the program. The regressions have time and firm fixed effects, so the coefficient can be interpreted as showing that the number of posts is significantly higher for firm *i* after it enters the program as compared to before. The dependent variable is a log, and exponentially transforming the short selling coefficient shows that the number of posts is 8.76% higher after shorting is allowed compared to before.

The regression reported in column 2 uses the standard deviation of media tone as its dependent variable. We control for the number of social media posts. The short selling dummy in this specification is 0.003 (*t*-statistic = 4.40), showing that volatility of social media tone increases after a firm enters the short selling pilot program.

The regressions reported in columns 3 and 4 report what happens when short selling was temporarily suspended among firms that were shortable. As we explain earlier, between August 2015 and March 2016, most security lenders temporarily stopped lending shares in China, so short selling was effectively halted. For these specifications, we limit our sample to firms that can be shorted, and test whether the number of posts and volatility of social media tone dropped during the suspension months.

In both regressions 3 and 4, the dummy for the short selling suspension is negative and significant, showing that both the number of posts and the volatility of social media tone declined when short selling was temporarily halted. In regression 3, exponentially transforming the no-shorting coefficient shows that the number of social media posts declined by 12.28% during the months in which short selling was halted. The no-shorting coefficient in regression 4 reflects about a 1.25% decrease in the volatility of social media tone during the no-shorting months.

[Insert Table 7 here]

2.3.4. The Number of Traditional News Articles and Volatility of Traditional Media Tone

Table 8 repeats the same regressions that are reported in Table 7, only we replace social media with traditional media. The first two regressions in Table 8 show that the number of traditional media articles and the standard deviation of the articles' tone did not increase for firms that entered the short selling pilot program. This is in contrast to the results in Table 7, which show that both the number of social media posts and the volatility of social media tone increased significantly for firms that entered the shorting program. Hence, the changes in social media documented in Table 7 are not a reflection of traditional media news stories.

Regression 3 shows that when short selling was suspended, the number of traditional news stories did not change significantly. This is again in contrast to Regression 3 in Table 7, which shows that there was a significant decline in social media posts when shorting was suspended. Regression 4 shows that the volatility of tone in traditional media actually increased when short selling was temporarily halted. This is opposite to the findings in Table 7, which show that the volatility of social media tone declined when shorting was suspended.

[Insert Table 8 here]

3. Short Selling, Social Media Tone, and Stock Returns

In this section of the paper, we study how shorting and media tone impact stock returns. A number of earlier studies show that high levels of short interest portend low stock returns in China and other countries (e.g., Dechow et al. (2001), Duan, Hu, and McLean (2010), Boehmer, Jones, and Zhang (2008), Engelberg, Reed, and Ringgenberg (2012), and Chang, Lou, and Ren (2014)). This is typically interpreted as showing that short sellers are informed investors. We build on this and look for evidence of short-seller manipulation.

3.1. Short Selling, Social Media Tone, and Stock Returns: General Results

Our results thus far show that social media tone is abnormally positive before stocks are highly shorted and abnormally negative once highly shorted. These effects are not explained by social media reflecting actual news, as traditional media follows no such pattern. If such patterns reflect manipulation, then firms that have such rise-then-fall patterns in social media tone should also have contemporaneous rise-then-fall patterns in stock returns.

To more carefully test for these effects, we regress the period's cumulative abnormal return (size adjusted daily stock return[10]) on a high short interest dummy, a dummy variable that we refer to as *Target*, the high short interest dummy interacted with *Target*, and controls. The

[10] At the beginning of each year, we divide all the stocks into quintiles based on the firm's market value. The daily abnormal return equals a firm's daily stock return minus the mean value of the stocks in the portfolio. Our results are basically the same when we use the market adjusted daily stock returns (daily stock return minus the daily market return)

variable *Target* is equal to 1 if the firm has social media tone that is above the sample median before and on the day that short interest is measured, and below the sample median after that day, and zero otherwise. As an example, in Regression 1 of Table 9, we study abnormal stock returns over the period *t-30* to *t*. *Target* is equal to 1 if the average daily social media tone was above the sample median over the period *t-30* to *t* and then below the sample median over the period *t+1* to *t+30*.

In Regression 1, the coefficients for the high short interest dummy, *Target*, and the interaction between the two are positive and significant. This shows that highly shorted firms, firms with rise-then-fall patterns in social media, and especially firms that have both of these effects have abnormally high stock returns during the 30 days before becoming highly shorted. The effects are economically meaningful. The coefficient for the high short interest variable is 0.026, for *Target* it is 0.039, while the *Target*-high short interest interaction coefficient is 0.008. Thus, for a firm that is both highly shorted and has a positive value of *Target*, the overall effect is the sum of the coefficients, which is equal to 0.073. Thus, the abnormal return effect of having both social media manipulation and high short interest is 181% greater than just having high short interest. Similar findings are reported Regression 2, which studies the effects over a 5-day window. In this regression, abnormal stock returns (the stock return associated with being highly shorted) are 217% higher for highly shorted firms that also have positive values of *Target* as compared to highly shorted firms that do not.

Regressions 3 and 4 examine the post shorting windows. Regression 3 examines returns over the period *t+1* to *t+5*. In this regression, the coefficients for the high short interest dummy, *Target*, and their interaction are all negative and significant. Thus, highly shorted firms have low stock returns, and the effects are greater for highly shorted firms with rise-then-fall patterns in social media tone. Regression 4 studies stock returns over the 30 days subsequent to being highly shorted. The coefficients suggest that stock returns are 333% lower for firms that are both highly shorted and have rise-then-fall social media tones, as compared to firms that are just highly shorted.

Table 9 in its entirety shows that stocks that are targeted by short sellers and have rise-then-fall patterns in social media tone have especially large runups and then declines in stock prices. These results are also shown in Figure 1, which displays the cumulative abnormal return for the 30 days before and 30 days after for the initiation of high short interest. Figure 1 shows that stock returns are significantly higher before and significantly lower after for highly shorted firms that have positive values of *Target* as compared to highly shorted firms that do not. Overall, the findings here are consistent with the idea that short sellers manipulate stock prices via social media.

[Insert Table 9 here]

[Insert Figure1 here]

3.2. Short Selling, Social Media Tone, and Stock Returns: The Effects of News Events

In this section of the paper, we conduct further tests of whether highly shorted firms with rise-then-fall patterns in social media tone have low returns due to stock price manipulation. As we explain earlier, most academic studies find that short sellers are informed traders. Christophe, Angel, and Ferri (2004) find that short sellers trade ahead of earnings announcements, anticipating which stocks will have poor earnings news. Boehmer, Jones, Wu, and Zhang (2019) find that event days with earnings news or analyst-related information account for 24% of short seller's abnormal returns, even though these days only account for 12% of trading days. More generally, Engelberg, Reed, and Ringgenberg (2012) find that the abnormal returns of highly shorted stocks are twice as large on news days as compared to non-news days. Our results thus far show that highly shorted stocks with rise-then-fall patterns in social media tone have abnormally low stock returns. If this somehow reflects informed trading that is anticipating subsequent bad news, then such stocks should have their abnormal returns concentrated on days with news. If instead, the low stock returns reflect social media manipulation, then we would not expect news days to play an outsized rule in accounting for the abnormal returns.

We report results from these tests in Table 10. To allow for ample time for news to come out, we study returns over the 30-day period after a firm becomes highly shorted.[11] In columns 1 and 2, the dependent variable is the daily abnormal stock return (stock return – size adjusted return). In column 1, the sample is limited to days with traditional news stories, while in column 2 the sample is limited to days with no traditional news stories. As in Tables 8 and 9, we regress stock returns on the high short interest dummy, the variable *Target*, which is equal to 1 if the firm has social media tone above the sample media in the pre-window and below the sample median in the post-window, and an interaction between the high short interest dummy and *Target*.

In column 1, where the sample is limited to days with news, the interaction coefficient is insignificant. That is, highly shorted stocks with rise-then-fall patterns in social media tone do not have abnormally low stock returns on days with traditional news. In column 2, however, the

[11] We get similar results using a 5-day window.

interaction is negative and significant, showing that the abnormally low returns earned by highly shorted stocks with rise-then-fall patterns in social media tone arise entirely on days *without* traditional news. This finding is not consistent with the abnormal returns reflecting the anticipation of bad news about the firm's fundamentals on the part of short sellers. Instead, the pattern suggests that the rise-then-fall pattern in social tone reflects stock price manipulation, which was traded on and perhaps caused by the short sellers.

The regressions reported in columns 3 and 4 tell a similar story. In these tests, we sort firms into two groups based on the number of traditional news stories over the 30-day period. The *More News* group consists of firms having the number of news days (size-adjusted) above the sample median, while the *Less News* group consists of firms having the number of news days (size-adjusted) below the median.[12] Here again, the interaction coefficient between high short interest and *Target* is insignificant in the more news specification, and negative and significant in the less news specification. That is, highly shorted firms with rise-then-fall patterns in social media tone earn abnormally low returns when there is less actual news. When there is more actual news, the abnormally low returns disappear. This is the opposite of what we would expect if the abnormally low returns reflected the anticipation of bad news by short sellers.

Focusing on the high short interest coefficient, the results in Table 10 show that highly shorted stocks have abnormally low returns on both news days and non-news days. The coefficients suggest slightly larger abnormal returns on news days, however the returns on non-news days are highly significant as well. Engelberg, Reed, and Ringgenberg (2012) find that in the

[12] We regress the number of news stories on firm size, and take the residual. We then sort firms into the two groups based on being above or below the median value of the residual. We do this because large firms tend to have more news stories.

U.S., the abnormal returns of highly shorted stocks are twice as large on news days as compared to non-news days. The results here show that in China the difference between the two is not as large. This suggests that short selling may be more informed in the U.S. than in China, although Chinese short selling is still informed.

4. Conclusion

Earlier studies find that short sellers play a stabilizing role in stock markets. The common narrative is that short sellers are informed investors that target overvalued firms. Short sellers' trades therefore encourage market efficiency. This narrative is consistent with roles that arbitrageurs play in classical finance and most behavioral finance models as well. In contrast, our paper shows that short sellers can play a destabilizing role in stock markets. One factor that makes our study different from earlier studies is the inclusion of social media data. Social media is relatively new, and most of the literature on arbitrage and short selling was written before social media had such large presence.

We find that firms that are targeted by short sellers tend to have rise-then-fall patterns in social media tone around the initiation of short interest. The patterns in social media tone are mirrored by patterns in abnormal stock returns, i.e., stock returns are abnormally high before shorting and abnormally low after shorting, and this effect is stronger for firms that have abnormally positive social media tone before shorting and abnormally negative social media tone after shorting. Our findings suggest that once social media is added to the mix, it can be more profitable for sophisticated investors to manipulate prices and exacerbate mispricing, rather than trade against it. Social media is an increasingly populator form of media and communication, so our findings are relevant for academic theories of price formation, the regulation of social media, the regulation of short sellers, and for how practitioners may view stock price dynamics in the presence of intense social media postings.

References

Bartov, E., Faurel, L., Mohanram, P.S. 2017. Can Twitter help predict firm-level earnings and stock returns? *Accounting Review* 93, 25-57.

Boehmer, E., Jones, C., Zhang, X. 2008. Which shorts are informed? *Journal of Finance* 63, 491-527.

Boehmer, E., Jones, J. Wu, C., Zhang, X. 2019. What do short sellers know? *Review of Finance*, forthcoming.

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, 1367-1403.

Christophe, S., Ferri, M., Angel, J. 2004. Short-selling prior to earnings announcements. *Journal of Finance* 59, 1845-1875.

Dechow, P., Hutton, A., Meulbroek, L., Sloan, R. 2001. Short sellers, fundamental analysis and stock returns. *Journal of Financial Economics* 61, 77–106.

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

De Long, J.B., Shleifer, A., Summers, L., Waldmann, R. 1990. Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance* 45, 379-395.

Diamond, D., Verrecchia, R. 1987. Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics* 18, 277–311.

Farrell, M., Green, T. C., Russell Jame, Markov, S. 2020. The Democratization of Investment Research and the Informativeness of Retail Investor Trading. *Working Paper*.

Figlewski, S. 1979. Subjective information and market efficiency in a betting market. *Journal of Political Economy* 87, 75-88.

Giannini, R., Irvine, P., Shu, T. 2018. Nonlocal disadvantage: An examination of social media sentiment. *Review of Asset Pricing Studies* 8, 293–336.

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, 1-25.

Piotroski, J., Wong, T.J., Zhang, T. 2017. Political bias in corporate news: The role of conglomeration reform in China. *Journal of Law and Economics* 60, 173-207.

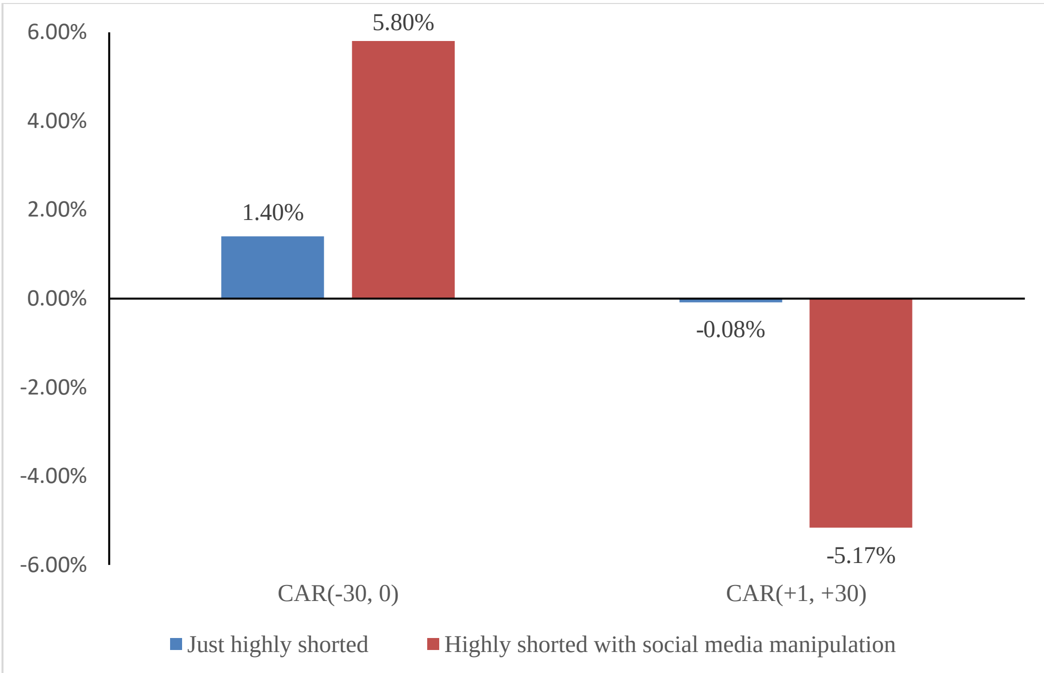
Soros, G. 1987. The alchemy of finance. Simon and Schuester, New York.

Tang, V.W. 2018. Wisdom of crowds: cross-sectional variation in the informativeness of third-party-generated product information on Twitter. *Journal of Accounting Research* 56, 989-1034.

Van Bommel, J. 2003. Rumors. *Journal of Finance* 58, 1499-1519.

Wang, E., Wong, T.J., Zhang, T. 2021. Do Chinese social media correct the optimistic bias of traditional media? *Working Paper*.

Figure 1
Stock returns before and after highly shorting for stocks with and without social media manipulation



Notes: This figure displays the stock returns before and after the initiation of high short interest. We divide the highly shorted firms into two groups: with and without social media manipulation. Social media manipulation is defined as having social media tone that is both above the sample median before the initiation of high short intertest, and below the sample median after the initiation of high short interest.

Table 1
Descriptive Statistics

Variable definitions are provided in the appendix.

	Variable	N	Mean	P50	SD	Min	Max
Social Media Tone	*AVGSMT(-30,0)*	1,125,207	-0.214	-0.222	0.158	-0.551	0.220
	AVGSMT(-5,0)	1,125,207	-0.212	-0.222	0.183	-0.607	0.287
	AVGSMT(+1,+5)	1,125,207	-0.205	-0.215	0.197	-0.631	0.334
	AVGSMT(+1,+30)	1,125,207	-0.210	-0.218	0.129	-0.483	0.159
Traditional Media Tone	*AVGTMT(-30,0)*	1,125,207	0.232	0.000	0.388	-0.750	0.976
	AVGTMT(-5,0)	1,125,207	0.178	0.000	0.360	-0.711	0.975
	AVGTMT(+1,+5)	1,125,207	0.158	0.000	0.346	-0.698	0.974
	AVGTMT(+1,+30)	1,125,207	0.319	0.316	0.399	-0.768	0.975
Short Selling	*SHORT INTEREST*	1,125,207	0.011	0.005	0.016	0.000	0.089
	SHORT INTEREST for the Top 10%	112,509	0.047	0.043	0.023	0.014	0.089
Firm Fundamentals	*SIZE*	1,125,207	23.968	23.827	1.092	22.011	27.337
	ROA	1,125,207	0.041	0.034	0.056	-0.188	0.204
	LEV	1,125,207	0.488	0.498	0.201	0.078	0.894
	BM	1,125,207	0.626	0.619	0.279	0.110	1.209
	PRERET	1,125,207	-0.003	-0.013	0.113	-0.277	0.372
Stock Returns	*CAR(-30, 0)*	1,125,207	-0.003	-0.013	0.114	-0.282	0.375
	CAR(-5, 0)	1,125,207	-0.001	-0.005	0.051	-0.130	0.179
	CAR(+1, +5)	1,125,207	-0.001	-0.005	0.046	-0.120	0.164
	CAR(+1, +30)	1,125,207	-0.003	-0.012	0.111	-0.279	0.362

Table 2
Sample Distribution

Panel A: Year distribution

Year	Freq.	Percent	Cum.
2010	11,695	1.04	1.04
2011	19,083	1.70	2.74
2012	55,059	4.89	7.63
2013	111,460	9.91	17.53
2014	157,961	14.04	31.57
2015	179,890	15.99	47.56
2016	185,667	16.50	64.06
2017	200,084	17.78	81.84
2018	204,308	18.16	100.00
Total	1,125,207	100.00	

Panel B: Industry distribution

Industry	Freq.	Percent	Cum.
Computer and Communications	92,684	8.24	8.24
Pharmaceutical	87,234	7.75	15.99
Real Estate	82,726	7.35	23.34
Chemical Products	52,542	4.67	28.01
Electrical Manufacture	49,850	4.43	32.44
Specialized Equipment Manufacture	46,206	4.11	36.55
Software and Information Technology	41,438	3.68	40.23
Metallic Product Manufacture	39,597	3.52	43.75
Automotive	38,607	3.43	47.18
Construction	36,334	3.23	50.41
Electricity and Heat Supply	34,862	3.10	53.51
Alcoholic Beverage, Non-alcoholic Beverage and Tea	34,151	3.04	56.54
Retails	31,311	2.78	59.33
Coal Mining and Washing	31,121	2.77	62.09
Non-metallic Mineral	30,899	2.75	64.84
Wholesale	30,055	2.67	67.51
General Equipment Manufacture	27,525	2.45	69.96
Non-Ferrous Metal Smelting	22,467	2.00	71.95
Transportation Equipment Manufacture	21,875	1.94	73.90
Ferrous Metal Smelting	18,290	1.63	75.52
Business Service	15,297	1.36	76.88
Water transportation	15,043	1.34	78.22
News and Publishing	15,095	1.34	79.56
Aero Transportation	12,583	1.12	80.68
Internet Service	12,348	1.10	81.78
Food Manufacture	11,508	1.02	82.80
Others	193,559	17.20	100.00
Total	1,125,207	100.00	

Table 3
Short-selling and social media tone

VARIABLES	(1) *AVGSMT(-30,0)*	(2) *AVGSMT(-5,0)*	(3) *AVGSMT(+1,+5)*	(4) *AVGSMT(+1,+30)*
TOP10%SHORT	0.008***	0.009***	-0.005***	-0.009***
	(3.10)	(3.19)	(-3.89)	(-5.45)
AVGTMT	0.017***	0.027***	0.021***	0.015***
	(13.25)	(22.27)	(21.70)	(14.06)
AVGSMT(-30,0)			0.522***	0.395***
			(100.11)	(70.51)
SIZE	0.006	0.007*	0.004*	0.002
	(1.51)	(1.71)	(1.93)	(0.84)
ROA	0.038	0.035	0.016	0.024
	(1.28)	(1.16)	(1.05)	(1.24)
LEV	0.024	0.024	0.008	0.016
	(1.41)	(1.44)	(0.98)	(1.52)
BM	-0.130***	-0.132***	-0.068***	-0.082***
	(-12.24)	(-12.36)	(-12.18)	(-11.54)
PRERET	0.055***	0.041***	0.008***	0.010***
	(13.63)	(9.42)	(2.63)	(2.86)
SOE	-0.009	-0.004	-0.002	-0.007
	(-0.73)	(-0.34)	(-0.31)	(-0.76)
Constant	-0.342***	-0.365***	-0.193***	-0.150**
	(-3.37)	(-3.57)	(-3.65)	(-2.23)
Year-Month	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes
Observations	1,125,207	1,125,207	1,125,207	1,125,207
Adj-R^2	0.292	0.140	0.208	0.400

Notes: This table examines how the social media tone behaves before and after the initiation of high short interest (the top 10% of short interest). *AVGSMT* is the average daily social media tone; *AVGTMT* is the average daily traditional media tone; *TOP10%SHORT* equals 1 if the daily short interest of a firm ranks in the top 10% of short interest (sorted by year) and zero otherwise; short interest is calculated as the daily unbalanced short-selling divided by outstanding shares; we control traditional media tones at the same window of calculating the social media tones (*AVGTMT*). Standard errors are adjusted for heteroskedasticity and clustered by firm. t-values are reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels (two-tail tests), respectively.

Table 4
Short-selling and traditional media tone

VARIABLES	(1) *AVGTMT(-30,0)*	(2) *AVGTMT(-5,0)*	(3) *AVGTMT(+1,+5)*	(4) *AVGTMT(+1,+30)*
TOP10%SHORT	-0.002	0.003	0.004	-0.009*
	(-0.33)	(0.85)	(1.45)	(-1.67)
AVGSMT	0.193***	0.098***	0.074***	0.180***
	(13.12)	(22.12)	(21.69)	(13.18)
AVGTMT(-30,0)			0.051***	0.070***
			(22.56)	(15.09)
SIZE	0.070***	0.049***	0.043***	0.068***
	(6.60)	(7.59)	(7.91)	(6.66)
ROA	0.520***	0.211***	0.142***	0.493***
	(7.24)	(5.34)	(4.25)	(6.98)
LEV	0.052	0.038*	0.025	0.035
	(1.26)	(1.68)	(1.27)	(0.94)
BM	-0.092***	-0.089***	-0.073***	-0.094***
	(-3.36)	(-5.50)	(-5.32)	(-3.67)
PRERET	0.035***	0.032***	0.026***	0.032***
	(2.99)	(5.03)	(4.53)	(2.79)
SOE	0.026	-0.007	-0.007	0.020
	(0.84)	(-0.36)	(-0.42)	(0.67)
Constant	-1.313***	-0.885***	-0.778***	-1.311***
	(-5.13)	(-5.71)	(-5.98)	(-5.41)
Year-Month	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes
Observations	1,125,207	1,125,207	1,125,207	1,125,207
Adj-R^2	0.211	0.186	0.187	0.217

Notes: This table examines how traditional media tone behaves before and after the initiation of high short interest (the top 10% of short interest). *AVGSMT* is the average daily social media tone; *AVGTMT* is the average daily traditional media tone; *TOP10%SHORT* equals 1 if the daily short interest of a firm ranks in the top 10% of short interest; short interest is calculated as the daily unbalanced short-selling divided by outstanding shares; we control traditional social tones at the same window (*AVGSMT*). Standard errors are adjusted for heteroskedasticity and clustered by firm. t-values are reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels (two-tail tests), respectively.

Table 5
Short-selling in the first week of being listed in pilot program and social media tone

VARIABLES	(1) *AVGSMT(-30,0)*	(2) *AVGSMT(-5,0)*	(3) *AVGSMT(+1,+5)*	(4) *AVGSMT(+1,+30)*
TOP10%SHORT	0.006	-0.024	-0.036**	-0.032***
	(0.52)	(-1.37)	(-2.52)	(-3.00)
AVGTMT	0.019*	0.013	0.039***	0.019**
	(1.89)	(0.92)	(3.17)	(1.98)
AVGSMT(-30,0)			0.585***	0.503***
			(13.04)	(15.15)
SIZE	0.022***	0.026***	0.000	0.006
	(4.50)	(3.66)	(0.04)	(1.15)
ROA	0.240***	0.144	0.052	-0.002
	(2.80)	(1.06)	(0.43)	(-0.02)
LEV	0.044*	0.010	-0.009	-0.007
	(1.67)	(0.26)	(-0.25)	(-0.30)
BM	-0.115***	-0.128***	-0.064**	-0.073***
	(-5.70)	(-4.57)	(-2.44)	(-4.04)
PRERET	0.023	0.094**	0.160***	0.101***
	(0.85)	(2.19)	(3.74)	(3.88)
SOE	0.022***	0.022*	0.018	0.011
	(2.72)	(1.86)	(1.55)	(1.40)
Constant	-0.726***	-0.802***	-0.053	-0.184*
	(-6.78)	(-5.16)	(-0.35)	(-1.68)
Observations	3,736	3,736	3,736	3,736
Adj-R^2	0.092	0.041	0.155	0.277

Notes: This table examines social media tone before and after the initiation of high short interest (the top 10% of short interest) during the first week that a firm is in the shorting program. The sample consists of firm-day observations limited to the first five days when a firm is in the shorting program. *AVGSMT* is the average daily social media tone; *AVGTMT* is the average daily traditional media tone; *TOP10%SHORT* equals 1 if the daily short interest of a firm ranks in the top 10% of short interest; short interest is calculated as the daily unbalanced short-selling divided by outstanding shares; we control traditional media tones at the same window (*AVGTMT*). Standard errors are adjusted for heteroskedasticity and clustered by firm. t-values are reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels (two-tail tests), respectively.

Table 6
Short-selling in the first week of being listed in pilot program and traditional media tone

VARIABLES	(1) *AVGTMT(-30,0)*	(2) *AVGTMT(-5,0)*	(3) *AVGTMT(+1,+5)*	(4) *AVGTMT(+1,+30)*
TOP10%SHORT	0.038	0.052	0.022	-0.005
	(1.27)	(1.55)	(0.67)	(-0.16)
AVGSMT	0.198*	0.050	0.150***	0.205**
	(1.89)	(0.92)	(3.60)	(2.09)
AVGTMT(-30,0)			0.067***	0.123***
			(2.62)	(3.48)
SIZE	0.011	0.079***	0.090***	0.058***
	(0.70)	(5.63)	(7.10)	(4.12)
ROA	0.174	-0.288	-0.167	0.250
	(0.57)	(-1.08)	(-0.73)	(0.88)
LEV	0.136	-0.095	-0.108*	-0.069
	(1.59)	(-1.26)	(-1.73)	(-0.81)
BM	-0.093	0.018	-0.005	-0.013
	(-1.46)	(0.34)	(-0.11)	(-0.20)
PRERET	0.122	0.178**	-0.128*	-0.110
	(1.36)	(2.19)	(-1.75)	(-1.22)
SOE	0.009	0.004	-0.010	-0.001
	(0.33)	(0.20)	(-0.51)	(-0.03)
Constant	0.092	-1.619***	-1.858***	-0.990***
	(0.26)	(-5.35)	(-6.72)	(-3.20)
Observations	3,736	3,736	3,736	3,736
Adj-R^2	0.013	0.036	0.059	0.043

Notes: This table examines traditional media tone before and after the initiation of high short interest (the top 10% of short interest) during the first week that a firm is in the shorting program. The sample consists of firm-day observations limited to the first five days when a firm is in the shorting program. *AVGSMT* is the average daily social media tone; *AVGTMT* is the average daily traditional media tone; *TOP10%SHORT* equals 1 if the daily short interest of a firm ranks in the top 10% of short interest; short interest is calculated as the daily unbalanced short-selling divided by outstanding shares; we control social media tones at the same window (*AVGSMT*). Standard errors are adjusted for heteroskedasticity and clustered by firm. t-values are reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels (two-tail tests), respectively.

Table 7
Short-selling pilot program and the number of social media posts and volatility of social media tone

	Whole sample		Short-selling sample	
VARIABLES	*POSTNUM*	*SD SMTONE*	*POSTNUM*	*SD SMTONE*
	(1)	(2)	(3)	(4)
SHORT	0.084***	0.003***		
	(4.16)	(4.40)		
CLOSEWINDOW			-0.131***	-0.004***
			(-9.73)	(-4.79)
SIZE	0.190***	0.003***	0.159***	0.006***
	(11.36)	(5.84)	(4.57)	(4.87)
ROA	-0.425***	0.0167***	-0.041	0.019**
	(-4.35)	(5.07)	(-0.21)	(2.49)
LEV	0.095*	0.003	-0.013	-0.007
	(1.73)	(1.56)	(-0.11)	(-1.63)
BM	0.245***	-0.005***	0.045	-0.014***
	(4.83)	(-3.52)	(0.47)	(-4.05)
RETURN	-0.205***	0.000	-0.211***	-0.003***
	(-17.84)	(0.94)	(-10.41)	(-3.77)
SDAR	35.900***	0.520***	30.660***	0.681***
	(148.70)	(32.34)	(64.64)	(22.65)
POSTNUM		-0.048***		-0.058***
		(-160.99)		(-105.06)
Constant	0.285	0.535***	1.891**	0.514***
	(0.76)	(49.14)	(2.29)	(17.31)
Firm	YES	YES	YES	YES
Year-month	YES	YES	YES	YES
Observations	249,963	249,963	56,982	56,982
Adj-R^2	0.564	0.447	0.605	0.518

Notes: This table examines how short selling affects the number of social media posts and the volatility of social media tone. Column (1) and (2) use all firm-month observations for all the non-financial listed firms; Column (3) and (4) only include firms that can be shortable. *POSTNUM* is the intensity of social media posts for a firm in a month, which is measured as the log value of one plus number of social media posts in a month; *SD SMTONE* is the monthly standard deviation of daily social media tone; *SHORT* equals 1 after a firm is listed in the pilot program, and 0 otherwise (many firms may be removed out of the list, for these firms, *SHORT* equals 0 after they are excluded); *CLOSEWINDOW* equals 1 for the period when short selling was temporarily suspended in China (from 2015, Aug. to 2016 Mar.), and 0 for other periods. Standard errors are adjusted for heteroskedasticity and clustered by firm. t-values are reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels (two-tail tests), respectively.

Table 8
Short-selling pilot program and the number of news articles and the volatility of traditional media tone

	Whole sample		Short-selling sample	
VARIABLES	*ARTICLENUM*	*SD TMTONE*	*ARTICLENUM*	*SD TMTONE*
	(1)	(2)	(3)	(4)
SHORT	-0.001	-0.001		
	(-0.06)	(-0.38)		
CLOSEWINDOW			0.006	0.021***
			(0.41)	(3.83)
SIZE	0.232***	-0.001	0.289***	0.001
	(14.48)	(-0.24)	(7.59)	(0.08)
ROA	-0.196**	-0.086***	-0.157	-0.140**
	(-2.12)	(-3.24)	(-0.70)	(-2.23)
LEV	0.086*	-0.022*	0.078	-0.015
	(1.72)	(-1.83)	(0.59)	(-0.51)
BM	-0.296***	0.034***	-0.314***	0.028
	(-6.75)	(3.27)	(-3.22)	(1.36)
RETURN	-0.105***	-0.011***	-0.119***	-0.010**
	(-9.68)	(-3.98)	(-6.42)	(-2.46)
SDAR	15.510***	1.263***	16.780***	1.095***
	(50.92)	(13.34)	(31.36)	(6.37)
ARTICLENUM		-0.005***		-0.009***
		(-2.73)		(-3.32)
Constant	-3.490***	0.420***	-4.421***	0.410**
	(-9.56)	(5.74)	(-4.80)	(2.28)
Firm	YES	YES	YES	YES
Year-month	YES	YES	YES	YES
Observations	82,842	82,842	26,254	26,254
Adj-R^2	0.627	0.062	0.742	0.101

Notes: This table examines how short selling affects the number of news articles and the volatility of traditional media tone. Column (1) and (2) use all firm-month observations for all the non-financial listed firms; Column (3) and (4) only include firms that can be shortable. *ARTICLENUM* is the coverage intensity of traditional media for a firm in a month, which is measured as the log value of one plus the number of news articles in traditional media; *SD TMTONE* is the monthly standard deviation of daily traditional media tone; *SHORT* equals 1 if a firm is in the short list pilot program, and 0 if not; *CLOSEWINDOW* equals 1 for the period when short selling was temporarily suspended in China (from 2015, Aug. to 2016 Mar.), and 0 for other periods. Standard errors are adjusted for heteroskedasticity and clustered by firm. t-values are reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels (two-tail tests), respectively.

Table 9
Short Selling, Social Media Tone, and Stock Returns

VARIABLES	(1) *CAR(-30, 0)*	(2) *CAR(-5, 0)*	(3) *CAR(+1, +5)*	(4) *CAR(+1, +30)*
TOP10%SHORT×TARGET	0.008**	0.002**	-0.005***	-0.007***
	(2.44)	(2.24)	(-7.44)	(-2.79)
TOP10%SHORT	0.026***	0.006***	-0.002***	-0.015***
	(10.99)	(12.28)	(-6.66)	(-8.10)
TARGET	0.039***	0.011***	-0.013***	-0.043***
	(34.46)	(44.54)	(-65.62)	(-44.68)
SIZE	0.020***	0.005***	0.003***	0.012***
	(6.57)	(7.71)	(7.18)	(4.64)
ROA	0.072***	0.011***	0.009***	0.084***
	(3.48)	(2.79)	(2.74)	(4.30)
LEV	0.022**	0.004**	0.003*	0.027***
	(2.02)	(1.98)	(1.92)	(2.63)
BM	-0.157***	-0.030***	-0.025***	-0.137***
	(-19.95)	(-20.29)	(-20.13)	(-19.67)
PRERET	-0.095***	-0.010***	-0.006***	-0.050***
	(-21.34)	(-9.76)	(-7.19)	(-11.51)
SOE	0.003	0.001	0.000	0.003
	(0.26)	(0.38)	(0.21)	(0.36)
Constant	-0.450***	-0.104***	-0.073***	-0.263***
	(-6.27)	(-7.36)	(-6.47)	(-4.24)
Year-Month	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes
Observations	1,125,207	1,125,207	1,125,207	1,125,207
Adj-R^2	0.080	0.020	0.024	0.074

Notes: This table examines stock returns for highly shorted firms, firms that are likely to have manipulated social media (TARGET = 1), and the interaction between the two variables. *CAR* is cumulative abnormal returns, where daily abnormal return is calculated as firms' daily stock return minus the average return of stocks in the portfolio with similar firm size; *TARGET* equals 1 if the average social media tone (*AVGSMT*) in the prior window is larger than the sample median and *AVGSMT* in the post window is smaller than the sample median, and 0 otherwise. We define *TARGET* using the same window as that for the *CAR* window. *TOP10%SHORT* equals 1 if the daily short interest of a firm ranks in the top 10% of short interest; *SHORT INTEREST* is calculated as the daily unbalanced short-selling divided by outstanding shares. Standard errors are robust and clustered by firm. t-values are reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels (two-tail tests), respectively.

Table 10 Short Selling, Social Media Tone, and Stock Returns: The Effects of News Events

	CAR(+1, +30)			
VARIABLES	*News event$_t$=1*	*News event$_t$=0*	More news during (+1,+30)	Less news during (+1,+30)
	(1)	(2)	(3)	(4)
TOP10%SHORT× TARGET	-0.003	-0.007***	-0.005	-0.008**
	(-0.89)	(-2.62)	(-1.33)	(-2.54)
TOP10%SHORT	-0.016***	-0.015***	-0.017***	-0.017***
	(-6.22)	(-7.63)	(-6.50)	(-6.77)
TARGET	-0.051***	-0.041***	-0.044***	-0.041***
	(-28.80)	(-42.18)	(-31.80)	(-33.17)
SIZE	0.006*	0.015***	0.040***	0.016***
	(1.76)	(5.30)	(7.57)	(3.50)
ROA	0.017	0.093***	0.059**	0.084**
	(0.51)	(4.65)	(2.38)	(2.47)
LEV	0.027*	0.027***	-0.003	0.064***
	(1.71)	(2.64)	(-0.21)	(3.94)
BM	-0.169***	-0.134***	-0.139***	-0.155***
	(-17.88)	(-18.04)	(-12.22)	(-13.66)
PRERET	-0.046***	-0.051***	-0.062***	-0.058***
	(-6.74)	(-11.49)	(-10.19)	(-10.28)
SOE	0.018**	-0.000	0.003	-0.026
	(2.05)	(-0.00)	(0.24)	(-0.99)
Constant	-0.111	-0.321***	-0.884***	-0.362***
	(-1.27)	(-4.90)	(-7.23)	(-3.30)
Year-Month	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes
Observations	196,404	928,803	549,517	549,496
Adj-R^2	0.088	0.074	0.102	0.089

Notes: This table examines whether highly shorted firms with rise-then-fall patterns in social media tone are more likely to have low returns on news days. *CAR* is cumulative abnormal returns, where daily abnormal return is calculated as firms' daily stock return minus the average return of stocks in the portfolio with similar firm size. We define news events as articles published in all the traditional media. *News event$_t$* equals one for days with news events, and zero otherwise. We count the number of news events over the window of (+1, +30) and adjusted by firm size, then divide the whole sample into two groups based on the sample median (column 3 and 4). *TARGET* equals 1 if the average social media tone (*AVGSMT*) in the prior window is larger than the sample median and *AVGSMT* in the post window is smaller than the sample median, and 0 otherwise. We define *TARGET* using the same window as that for the *CAR* window. *%SHORT* equals 1 if the daily short interest of a firm ranks in the top 10% of short interest. All the standard errors are adjusted for heteroskedasticity and clustered by firm. t-values are reported in parentheses. ***, **, and * denote significance at 1%, 5%, and 10% levels (two-tail tests), respectively.

Appendix
Variable definition

Variables	Definition
SHORT INTEREST	= daily unbalanced short-selling divided by outstanding shares
TOP10%SHORT	= 1 for the top 10% of short interest; = 0 otherwise
AVGSMT(-30 , 0)	= the average daily social media tone over the period t-30 to t, with missing value replaced by zero
AVGSMT(-5 , 0)	= the average daily social media tone over the period t-5 to t, with missing value replaced by zero
AVGSMT(+1 , +5)	= the average daily social media tone over the period t+1 to t+5, with missing value replaced by zero
AVGSMT(+1 , +30)	= the average daily social media tone over the period t+1 to t+30, with missing value replaced by zero
AVGTMT(-30 , 0)	= the average daily traditional media tone over the period t-30 to t, with missing value replaced by zero
AVGTMT(-5 , 0)	= the average daily traditional media tone over the period t-5 to t, with missing value replaced by zero
AVGTMT(+1 , +5)	= the average daily traditional media tone over the period t+1 to t+5, with missing value replaced by zero
AVGTMT(+1 , +30)	= the average daily traditional media tone over the period t+1 to t+30, with missing value replaced by zero
SIZE	= the log value of the market value of firms at the end of the fiscal year
LEV	= the leverage ratio at the end of the fiscal year, which is calculated as total debt divided by total assets
ROA	= return on total assets, which is net income divided by total assets
BM	= book-to-market ratio at the end of the fiscal year
PRERET	= 30-day cumulative abnormal returns prior to shorting days (skip 1 month, that is, from day t-60 to day t-30)
SD SMTONE	= the monthly standard deviation of daily social media tone

SD TMTONE	= the monthly standard deviation of daily traditional media tone
SHORT	= 1 if a firm is in the short list pilot program, = 0 if not
CLOSEWINDOW	= 1 for the period that most of the security firms temporarily close their business of lending stocks to the market (2015, Aug. to 2016 Mar.), = 0 for other periods
RETURN	= stock return for the current fiscal year
SDAR	= the monthly standard deviation of stock return
POSTNUM	= the intensity of social media posts for a firm in a month, which is measured as the log value of one plus number of social media posts in a month
ARTICLENUM	= the coverage intensity of traditional media news articles for a firm in a month, which is measured as the log value of one plus number of articles in traditional media in a month
TARGET	=1 if the firm has social media tone above the sample media in the pre-window and below the sample median in the post window; =0 otherwise
CAR(-30, 0)	= the cumulative abnormal stock returns over the period of t-30 to t, where daily abnormal stock return is the size adjusted daily return
CAR(-5, 0)	= the cumulative abnormal stock returns over the period of t-5 to t, where daily abnormal stock return is the size adjusted daily return
CAR(+1, +5)	= the cumulative abnormal stock returns over the period of t+1 to t+5, where daily abnormal stock return is the size adjusted daily return
CAR(+1, +30)	= the cumulative abnormal stock returns over the period of t+1 to t+30, where daily abnormal stock return is the size adjusted daily return