How Does Social Media Sentiment Impact Mass Media Sentiment? A Study of News in the Financial Markets

Jie Ren
Fordham University, NY, USA
jren11@fordham.edu

Hang Dong**
IE University, Madrid, Spain
hang.dong@ie.edu

Balaji Padmanabhan
University of South Florida, FL, USA
bp@usf.edu

Jeffrey V. Nickerson
Stevens Institute of Technology, NJ, USA
jnickers@stevens.edu

Abstract

Mass media sentiment of financial news significantly influences investment decisions of investors. Hence, studying how this sentiment emerges is important. In years past this was straightforward, often dictated by journalists who cover financial news, but this has become more complex now. In this paper we focus on how social media sentiment affects mass media sentiment. Using data from Sina Weibo and Sina Finance (around 60 million weibos and 6.2 million news articles) we show that social media does influence mass media sentiment emergence for financial news. The sentiment consistency between social media reaction and prior news articles amplifies the persistence of mass media sentiment over time. By contrast, we found limited evidence of social media reducing the persistence of mass media sentiment over time. The results have significant implications for understanding how two types of media, treated separately in the literature, may be connected.

Keywords: sentiment emergence, social media, IS/finance interface, financial information systems, financial markets.

** Contact author

1. INTRODUCTION

Sentiment is an important driver of investment decisions in the stock market (Deng, Huang, Sinha, & Zhao, 2018; Huang, Jiang, Tu, & Zhou, 2015). Mass media has long been the main source of this sentiment in the financial markets. Based on this link, many scholars have explored how mass media sentiment affects stock market performance (Tetlock, 2007, 2014; Tetlock, Saar‐Tsechansky, & Macskassy, 2008). For example, using deep learning techniques to tap into stock-relevant news, Hu et al. (2018) tried to predict stock trends (Hu, Liu, Bian, Liu, & Liu, 2018). Tetlock (2007) recorded the different reactions of Dow Jones returns to negative sentiment in mass media short-term versus long-term. Tetlock et al. (2008) further focused on negative words in firm-specific news and found that such negative sentiment predicts firms' low financial performance. But how does mass media sentiment about financial news emerge? In years past this was primarily driven by how reporters covered the news, but in recent years this has become more complex due to the role of social media. In this paper we examine exactly how social media impacts mass media sentiment in financial markets.

The sentiment expressed in mass media is certainly not easy to predict, even in a seemingly objective context such as financial news. Even in cases where the information about a stock is the same, the sentiment of mass media articles can be different. For instance, Alphabet, the parent company for Google, released its Q3 financial earnings on Oct 28, 2019. *The Independent*, a journal based in London, reported this announcement in a positive light: "Alphabet results: Google parent company announces 20 percent revenue increase." By contrast, some journals marked the same news as negative, including *Business Insider*, who entitled the report "Alphabet's Q3 earnings missed the Street's expectations by a mile."

Why does this happen? The fact that reader preferences are important (Mullainathan & Shleifer, 2005) suggests that social media might play a role in mass media sentiment emergence. In particular, Mullainathan and Shleifer (2005) argued that readers prefer to see news that are consistent with their prior beliefs and mass media may slant their news articles' sentiment to cater to such preference of readers.

Following this conceptual argument from Mullainathan & Shleifer (2005), we further argue, social media may act as a sentiment *filter* in affecting mass media sentiment in the stock market context. That is, social media reacts to news reports and modifies (amplifies or reduces) the persistence of mass media sentiment on the same topic over time. Based on this "demand-driven media bias" as illustrated by Mullainathan & Shleifer (2005) and Gentzkow & Shapiro (2010), and other alternative explanations (see Section 3 for details), there could be two competing mechanisms for the sentiment filter role.

One is this media bias: mass media caters somewhat to readers' preferences in setting the sentiment of news articles. The other is a mechanism of mass media where it evaluates and corrects social media's possible noise (for example, by rejecting rumors or correcting overreactions). In this case, given the same information and social media's discussion of the information, mass media may intentionally correct the social media's sentiment. This mechanism originates from the purported desire of mass media to pursue news objectivity. Our analysis in this paper provides important perspectives in this context by evaluating both mechanisms and reporting results.

We analyzed a large and unique dataset from Sina Finance and Sina Weibo in the Chinese stock market. Sina Finance as a financial news platform includes almost all financial news articles available on the Chinese stock market. Sina Weibo is a microblogging service similar to Twitter. Each microblog was restricted to 140 characters during the timeframe of our study. Our original dataset includes 6.2 million pieces of stock-related news and 59.36 million stock-related weibos from these two platforms for the years 2013 and 2014. The dataset allows us to trace the sentiment flow from one platform to another.

We present strong evidence in support of social media's filter role that is driven by the mechanism of "demand-driven media bias" in affecting mass media sentiment. Our results provide clear evidence for the *amplifying filter* role: the sentiment consistency between social media reaction and prior news articles amplifies the persistence of mass media sentiment over time. We found only limited, if any, evidence for the *reducing filter* role.

2. LITERATURE REVIEW

2.1 Media Biases

Mass media typically claim themselves as objective information channels. However, literature has shown that mass media may have a bias (Gentzkow, Shapiro, & Stone, 2015; Puglisi & Snyder Jr, 2015). This also holds in the stock market, often considered a utilitarian setting and less subject to any media bias (Tetlock, 2007, 2015). There are two sources of media bias, which we discuss below.

The first is caused by supply-side factors, such as pressure from ownership and governments (Ansolabehere, 2006; Larcinese, Puglisi, & Snyder, 2011; Puglisi & Snyder Jr, 2015). For example, Ahern and Sosyura (2014) argued that "the timing and content of financial media coverage may be biased by firms seeking to manipulate their stock price.". Dougal et al. (2012) found evidence that journalists' characteristics (e.g., optimism/pessimism about the market, experience, education, and industry expertise) are associated with the quality of their financial news articles. Reuter and Zitzewitz

(2006) found that some personal finance publications more frequently recommend a mutual fund that spends more on advertising, while this mutual fund does show better post-recommendation returns. The authors further concluded that the independence of mass media is compromised by the advertising. If these were the only factors, mass media sentiment emergence should be treated as random and exogenous.

The second media bias may be conceptualized as demand-driven media bias. Due to competition, media outlets compete for viewership in an effort to increase advertising revenues. Therefore, they may report the same information set with different sentiments catering to different audience groups (Gentzkow & Shapiro, 2010; Mullainathan & Shleifer, 2005). Our paper is more aligned with this stream of the literature and presents compelling evidence for this in the stock market.

To the best of our knowledge, there is no empirical evidence from prior literature on demand-driven media bias in the stock market. For investors, demand-driven media bias can be even more dangerous than supply-driven bias in the stock market. The supply-driven bias is generated randomly and the bias can decrease with the increasing competition (Gentzkow & Shapiro, 2008). By contrast, demand-driven media bias may be persistent and possibly lead to market bubbles and crises because it is caused by bias from investors, who consequently reinforce their bias by reading biased mass media news that they caused in the first place, creating an echo chamber effect.

2.2 The Role of Social Media in Financial Markets

There is also a large body of work showing how social media affects investors. Broadly, this suggests that social media may provide information to investors, but also plays an important role in investor sentiment.

One stream of work in this context relates to "information." It argues that social media reveals information that is unobtainable from traditional mass media on how customers perceive a firm's product and brand performance, both of which are associated with a firm's fundamental value (Luo, Zhang, & Duan, 2013). This stream of literature does not argue that social media affects its users, but serves as a proxy for a source of fundamental information.

A second stream of literature relates to "sentiment". It argues that social media affects users directly as social media users may be individual investors. For example, Deng et al. (2018) explicitly argued that social media captures the sentiment of market participants. They proposed that social media information, especially at the hourly level, drives social media users to incorrectly react in the stock market. An abundance of finance literature also claims this mechanism (Dong & Gil-Bazo, 2020;

Renault, 2017). For example, Renault (2017) documented that the association between social media and stock returns is mainly driven by the sentiment of novice traders. Another example shows that social media sentiment can affect bitcoin prices as social media users may be trading bitcoins (Mai, Shan, Bai, Wang, & Chiang, 2018).

The emerging consensus here is that, not surprisingly, social media sentiment directly affects social media users. Our paper follows this stream of literature and defines social media sentiment as representing the sentiment of market participants. Accordingly, our paper provides a different mechanism for how social media affects the stock market. We argue that there may be a flow of sentiment from social media to mass media. Our focus here is on the impact of social media sentiment on the sentiment of the news, not that on the information content in the news. Consequently, by affecting future mass media sentiment, social media sentiment not only affects its own users but also affects mass media readers, a much larger group.

3. HOW SOCIAL MEDIA AFFECTS MASS MEDIA SENTIMENT

How does social media affect mass media sentiment? In general, social media may play two roles in affecting mass media sentiment emergence: 1) acting as a sentiment *generator*, and 2) acting as a sentiment *filter*. As a generator, social media could create new sentiment, which mass media later reports. As a filter, social media reacts to news reports and modifies (amplifies or reduces) the persistence of mass media sentiment on the same topic over time. Pew Research (2019)[1] showed about 43% of American adults receive, share and comment on news from Facebook.

In a general market, mass media may monitor the sentiment of social media events and set the sentiment of news articles on related topics – this is the sentiment generator role of social media. For example, in the entertainment industry, mass media monitors the emotions of fans of certain celebrities and report news on these celebrities using the same suggested sentiment (Liang & Shen, 2016), in part to attract more eyeballs and thus more advertising revenues.

However in the financial market, mass media plays a dominant role in diffusing sentiment (Ahern & Sosyura, 2015; Tetlock, 2015). Hence we posit that it is unlikely for social media to act as a sentiment generator for financial news. Therefore, in this paper, we focus on the other more apparent sentiment role of social media in the financial market, which serves a *sentiment filter*: social media may discuss the prior news articles, developing a sentiment; mass media may react to this sentiment in setting the

[1] https://www.pewresearch.org/fact-tank/2019/05/16/facts-about-americans-and-facebook/

sentiment of later news articles on the same topic. If this is the case, future mass media articles on the same stock/topic will reflect the new curated sentiment.

In the next two subsections, we argue that social media amplifies or reduces the persistence of mass media sentiment over time, driven by two possible mechanisms. One mechanism argues mass media plays a role in information diffusion by catering to their readers' preferences. The other mechanism argues mass media may play a role to correct misinformation in social media in the stock market.

3.1 Mechanism #1: Demand-driven Media Bias

In the first mechanism, mass media is driven by the demand of readers – this is the theoretical argument of "demand-driven media bias" (Gentzkow & Shapiro, 2010; Mullainathan & Shleifer, 2005). People are subject to some psychological vulnerabilities, known as bounded rationality (Simon, 1982). In particular, people may experience psychological stress when they receive information that is contradictory to their existing opinions and, therefore, they may prefer information that confirms their previous beliefs, known as "cognitive dissonance theory" or "confirmation bias" (Festinger, 1957; Park, Konana, Gu, Kumar, & Raghunathan, 2013). In order to attract readers' attention and increase the viewership of the news articles, which can further attract advertising revenues, mass media outlets may want to provide news sentiment that is consistent with readers' beliefs, which can be proxied and traced by social media (Mullainathan & Shleifer, 2005).

In other words, the desire to attract viewership leads mass media outlets to slant news sentiment toward catering to social media sentiment. For example, as cognitive dissonance theory shows, people are more likely to comment and share sentiment that they agree with (Harmon-Jones & Mills, 1999). Such sentiment consistency between mass media sentiment and social media sentiment can amplify the mass media sentiment persistence over time. In other words, mass media sentiment at $t-1$ is more persistent toward t when mass media sentiment at $t-1$ and social media sentiment at $t-1$ are consistent (both positive or both negative). On the other hand, if social media users disagree with news sentiment, such disagreement can reduce the mass media sentiment persistence. That is, mass media at $t-1$ is less persistent toward t when mass media sentiment at $t-1$ and social media sentiment at $t-1$ are inconsistent (one is positive and the other is negative, or vice versa).

3.2 Mechanism #2: Mass Media Correcting Social Media's Noisy Reaction to Prior News

Motivated by recent studies on rumors and fake news (Allcott & Gentzkow, 2017; Shu, Sliva, Wang, Tang, & Liu, 2017; Vosoughi, Roy, & Aral, 2018), we argue that in the second mechanism social media users may intentionally or accidentally spread rumors or fake news in response to the news articles of certain companies or stocks reported from mass media professionals. Previous literature documented that social media diffuses non-fundamental information, defined as any information that is not associated with the fundamental value of the company (Deng et al., 2018; Jiao, Veiga, & Walther, 2020). Mass media with the purpose of providing objective information may work to correct social media sentiment: when social media users overreact or underreact to mass media sentiment at *t-1*, mass media sentiment at *t* may be inconsistent with social media sentiment at *t-1* as mass media attempts to correct/reduce the perceived incorrect social media sentiment at *t-1*. In this mechanism, we predict a different result of the filter role: mass media sentiment persistence over time (from *t-1* to *t*) increases (decreases) when mass media sentiment at *t-1* and social media sentiment at *t-1* are inconsistent (consistent).

Both mechanisms above predict that social media plays a sentiment filter role, having two subroles: the amplifying filter role and the reducing filter role and suggesting the impact on mass media sentiment emergence, but with different directions. Therefore, we propose two sets of competing hypotheses (see Table 1). The first set (H1a and H1b) is motivated by "the demand-driven media bias" argument. The second set (H2a and H2b) is motivated by the "mass media correcting social media's noises" argument, as described above.

H1a. The sentiment consistency between social media reaction and prior news articles (both positive or both negative) amplifies the persistence of mass media sentiment over time.

H1b. The sentiment inconsistency between social media reaction and prior news articles (one positive and the other negative, or vice versa) reduces the persistence of mass media sentiment over time.

H2a. The sentiment inconsistency between social media reaction and prior news articles (one positive and the other negative, or vice versa) amplifies the persistence of mass media sentiment over time.

H2b. The sentiment consistency between social media reaction and prior news articles (both positive or both negative) reduces the persistence of mass media sentiment over time.

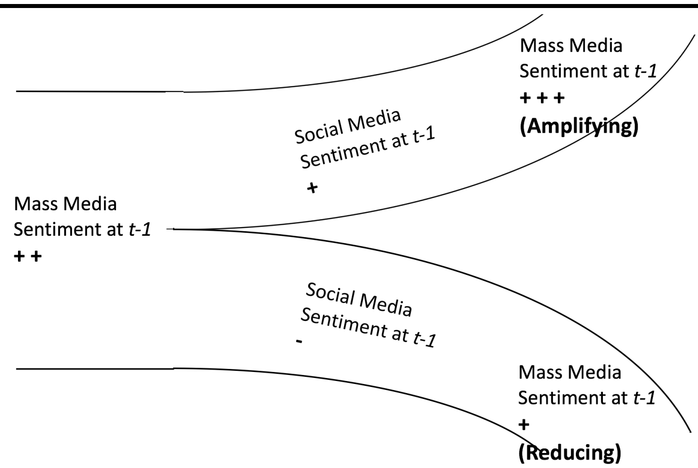
Table 1. Two Sets of Competing Hypotheses for the Sentiment Filter Role of Social Media

Hypotheses	The working mechanism	Brief explanation
H1a (the amplifying filter role) H1b (the reducing filter role)	Demand-driven media bias	Mass media slants toward readers' preferences (here social media sentiment) in setting their news article sentiment.
H2a (the amplifying filter role) H2b (the reducing filter role)	Mass media correcting social media's noisy reaction to prior news	Assuming that social media may be a source of noises, mass media in the financial market is filtering out the noises and correcting any "incorrect" sentiment.

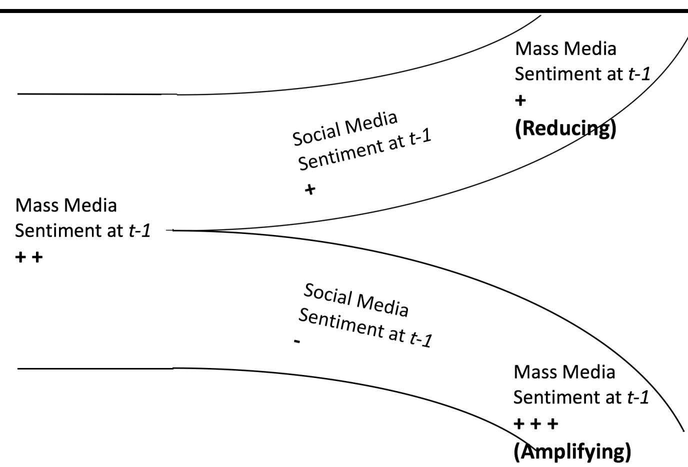
Figure 1 illustrates the two mechanisms when mass media sentiment is positive at *t-1* as an example.

Figure 1. The Two Mechanisms (When Mass Media Sentiment at *t-1* is Positive as an Example)

Driven by the Demand-driven Media Bias Mechanism



Driven by the Correcting Mechanism



Whichever mechanism is dominant, our research has significant implications. If the former mechanism is dominant (related to H1a and H1b), mass media, even in the financial market, may be subject to a systematic demand-driven media bias in setting sentiment, which may lead individual investors, who are usually the news consumers, to reinforce their (biased, if any) beliefs on stocks. If the latter mechanism is dominant (related to H2a and H2b), assuming that social media may be a source of noises, mass media in the financial market is filtering out the noises and correcting any "wrong" sentiment – this suggests a more unbiased role of mass media in financial markets where mass media strives to maintain the "correct" sentiment of news.

4. DATA AND VARIABLES

We analyzed data from Sina Finance – a finance news platform that distributes stock-related news and Sina Weibo, a microblogging service in China. The two platforms are independent. The Chinese stock market is an ideal market for our study because most recipients of information in this market are individuals. According to the China Securities Depository and Clearing Corporation Limited, as of the end of 2014, more than 99.6% of investors in the Chinese stock market, or 72.71 million investors, were individuals, and 85.37% of trading volume corresponded to trades of individual investors. Another reason is that in China, by law, companies can only release their announcements via mass media (for details, please see Subsection 5.2.2) – we can leverage this unique characteristic of the Chinese stock market to relieve the endogeneity issue, i.e., sentiment flows from mass media to social media, to the extent possible. All authors do not receive monetary incentives from Sina and claim no conflict of interest.

4.1 Social Media Sentiment

The microblogging data from Sina Weibo is from 2013 to 2014. As reported in the 2015 annual report of Sina, as of December 2015, Sina Weibo had 235.7 million active monthly users and an average of 106.3 million active daily users. Weibo is considered the most influential microblogging platform in China (Harwit, 2014). Sina provided us with the original sentiment data.

Sina used Ticker and *Jiancheng* (literally: the short name of stocks in Chinese), to extract all weibos (microblogs) that mentioned Chinese stocks during 2013 and 2014. This method of using Ticker to identify social media posts on relevant stocks is commonly used in the literature (Da, Engelberg, & Gao, 2011). An interesting feature in the Chinese stock market is that Ticker is typically a series of numbers. We also provided some examples in Table A.5. The Tickers do not have practical meanings in daily life. As such, Sina may perfectly match Weibo posts with the stocks they mentioned. We argue that our social media sentiment captures sentiment of market participants in the stock market, rather than sentiment of consumers as documented in Luo et al., (2013). Then Sina used its proprietary sentiment dictionary to measure the sentiment of microblogs – Sina gives a sentiment score for each word and sums all the scores of words that appreared in each weibo. If the overall sentiment score for a weibo is positive, netural, or negative, Sina considers the weibo as positive (1), neutral (0), or negative (-1). We cannot retrieve the content of the Weibo posts. As such, we validated Sina's sentiment measure to ensure the sentiment word dictionary that Sina used was robust. Specifically, we randomly selected 40,000 news items from December 2014. We used BosonNLP, a third-party commercial software and used in Zheng et al. (2019), to calibrate the tones of these news items. We found that 82.67% of the sample provided by Sina has the same sentiment as that determined by the third-party software.

Our original dataset includes nearly 60 million microblogs from January 2013 to December 2014, each of which refers to at least one stock. There are very few weibos (around 0.007%) mentioning more than one stock. We dropped those observations. We also dropped 96 stocks that are delisted or inactive (the number of trading days is more than one third, or 161 days, of the overall trading days during our sample period). The final data includes 58,705,064 Weibo posts covering 2,501 stocks during the 2013 and 2014 period. Our dataset covers 96.3% of the stocks traded in the stock market (2597 stocks) and 97.8% of the market capitalization. We calculated social media sentiment the same way as Antweiler and Frank (2004) and Deng et al. (2018) did:

$$SentWeibo_{i,t} = ln\,(\frac{1 + PW_{i,t}}{1 + NW_{i,t}})$$

where $SentWeibo_{i,t}$ is the proxy for social media sentiment on stock *i* at day *t*, and $PW_{i,t}$ ($NW_{i,t}$) is the number of microblogs with positive (negative) sentiment on stock *i* at day *t*.[2]

4.2 Mass Media Sentiment

Sina directly provided us with more than 6.2 million financial news articles for the same time period (2013 and 2014). We only kept the news articles mentioning stocks, which left more than 1 million news articles in our dataset. We then dropped news articles that included more than one stock

[2] Following the standard financial literature (Antweiler & Frank, 2004; García, 2013), *t* indicates natural/calendar days.

because a news article with more than one stock may discuss the stock market rather than the individual stocks. We confirmed that the news articles that we dropped are at the industry and market level rather than at the individual stock level. Finally, our database included 462,984 news articles. This data processing procedure is standard in the literature (Tetlock, 2011).

Sina Finance provided news sentiment in the same way as in how Weibo provided social media sentiment. That is, a news article can be positive (coded as 1), neutral (0), or negative (-1). Similar to the variable definition of social media sentiment, we calculated the news sentiment ("*SentNews*") for each stock. Specifically,

$$SentNews_{i,t} = ln\,(\frac{1 + PN_{i,t}}{1 + NN_{i,t}})$$

where $SentNews_{i,t}$ is the proxy for mass media sentiment on stock *i* at day *t*, and $PN_{i,t}$ ($NN_{i,t}$) is the number of news articles with a positive (negative) sentiment on stock *i* at day *t*.

We used stock and day to merge the two Sina datasets and the stock market data to create a panel dataset. We obtained the stock market data from the *Resset* dataset, which is a commonly used academic database on the Chinese market (Li, Song, & Wu, 2015). The dataset provides stock characteristics on the 2,501 Chinese stocks during the time period of 2013 and 2014.

4.3 Differentiating Positive Sentiment from Negative Sentiment

We differentiated positive sentiment from negative sentiment for both social media and mass media. We aimed to test how mass media reacts to prior social media sentiment. As such, the coefficient of our main variables of interests – the coefficient of the interaction term between prior social media sentiment and prior mass media sentiment for the sentiment filter role of social media – may vary depending on whether the sentiment is positive or negative (for details see the next section.)

Following the definitions in Sections 4.1 and 4.2, we defined positive social media sentiment ($SentWeiboP_{i,t}$), negative social media sentiment ($SentWeiboN_{i,t}$), positive mass media sentiment ($SentNewsP_{i,t}$), and negative mass media sentiment ($SentNewsN_{i,t}$) as the natual logarithm of one plus the volume of the positive (negative) Weibo posts or news articles:

$$SentWeiboP_{i,t} = ln\,(1 + PW_{i,t})$$

where $SentWeiboP_{i,t}$ is the proxy for the positive social media sentiment on stock *i* at day *t*, and $PW_{i,t}$ is the volume of positive social media posts on stock *i* at day *t*.

$$SentWeiboN_{i,t} = ln\,(1 + NW_{i,t})$$

where $SentWeiboN_{i,t}$ is the proxy for the negative social media sentiment on stock *i* at day *t*, and $NW_{i,t}$ is the volume of negative social media posts on stock *i* at day *t*.

$$SentNewsP_{i,t} = ln\,(1 + PN_{i,t})$$

where $SentNewsP_{i,t}$ is the proxy for the positive mass media sentiment on stock *i* at day *t*, and $PN_{i,t}$ is the volume of positive mass media news articles on stock *i* at day *t*.

$$SentNewsN_{i,t} = ln\,(1 + NN_{i,t})$$

where $SentNewsN_{i,t}$ is the proxy for the negative mass media sentiment on stock *i* at day *t*, and $NN_{i,t}$ is the volume of negative mass media news articles on stock *i* at day *t*.

All variables' definitions, summary statistics, and correlations can be found in the Appendix. In Table A.4 in the Appendix, we provided the distribution of Weibo posts and News articles across the stocks. In Table A.5, we gave examples of the top 10 stocks mentioned in Weibo posts and news articles during our sample period.

5. METHODOLOGY AND RESULTS

We examined how social media sentiment moderates the persistence of mass media sentiment over time. Specifically, we regressed mass media sentiment at *t* on mass media sentiment at *t-1*, social media sentiment at *t-1,* and their interaction terms. To relieve the endogeneity issue to the extent possible, we focused on a scenario where mass media sentiment at *t-1* is calculated using companies' announcements in the media at *t-1* because, in China, by law, companies' announcements should first be published through mass media as mentioned above. In other words, mass media sentiment of announcements at *t-1* when the announcements were published for the first time should be irrelevant to mass media sentiment at *t-2* and social media sentiment at *t-2*. Therefore, announcements can be viewed as external shocks in our setting. In the following subsections, we present the details of our tests.

5.1 Methodology

For analysis, we tested the following model.

$$SentNews_{i,t} = \alpha_i + \alpha_t + \beta_1 SentNewsP_{i,t-1} \times SentWeiboP_{i,t-1} + \beta_2 SentNewsN_{i,t-1} \times SentWeiboP_{i,t-1} + \beta_3 SentNewsP_{i,t-1} \times SentWeiboN_{i,t-1} + \beta_4 SentNewsN_{i,t-1} \times SentWeiboN_{i,t-1} + \beta_5 SentNewsP_{i,t-1} + \beta_6 SentNewsN_{i,t-1} + controls + \varepsilon_{i,t} \quad (1)$$

where $SentNews_{i,t}$ is mass media sentiment on stock *i* at day *t*; α_i and α_t are stock fixed effects and day fixed effects; $SentNewsP_{i,t-1}$ is the positive mass media sentiment on stock *i* at *t-1*; $SentNewsN_{i,t-1}$ is the negative mass media sentiment on stock *i* at *t-1*; $SentWeiboP_{i,t-1}$ is the positive social media sentiment on stock *i* at day *t-1*; $SentWeiboN_{i,t-1}$ is the negative social media sentiment on stock *i* at day *t-1*. Positive and negative social media sentiment at *t-1* ($SentWeiboP_{i,t-1}, SentWeiboN_{i,t-1}$), social media sentiment at *t-2* and *t-3* ($SentWeibo_{i,t-2}$, $SentWeibo_{i,t-3}$), and mass media sentiment at *t-2* and *t-3* ($SentNews_{i,t-2}$, $SentNews_{i,t-3}$) are control variables. The reason of controlling for social media sentiment at *t-2* and *t-3* is that we want to include at least one control variable from the trading days (e.g., the Mondays sample should include the sentiment on Fridays). The standard errors are clustered at the stock and day level.

Social media's sentiment filter role suggests that social media may amplify or reduce the persistence of mass media sentiment over time. Specifically, to explain the equation in detail, we used the working mechanism where demand-driven media bias is functioning as one example: the amplifying filter role suggests β_1 in Equation 1 to be positive and β_4 to be negative – as past social media sentiment is positive (negative) and past mass media sentiment is positive (negative), future mass media sentiment is also more positive (negative) – mass media sentiment is more persistent over time. The reducing filter role suggests that β_2 is positive and β_3 is negative – as past social media sentiment is positive (negative) and past mass media sentiment is negative (positive), future mass media sentiment is less negative (less positive), which is more conflicting with past mass media sentiment.

Figure 2 conceptually demostrates our predictions driven by the demand-driven media bias mechanism. Here when social media sentiment at *t-1* and mass media sentiment at *t-1* are consistent (both positive or both negative), the persistence of mass media sentiment (from *t-1* to *t*) is amplified, with future mass media sentiment being more positive or more negative. Otherwise, the persistence of mass media sentiment over time is reduced. Please note in this figure we only focused on the slope changes (positive changes or negative changes) and did not need to clarify the intercepts.

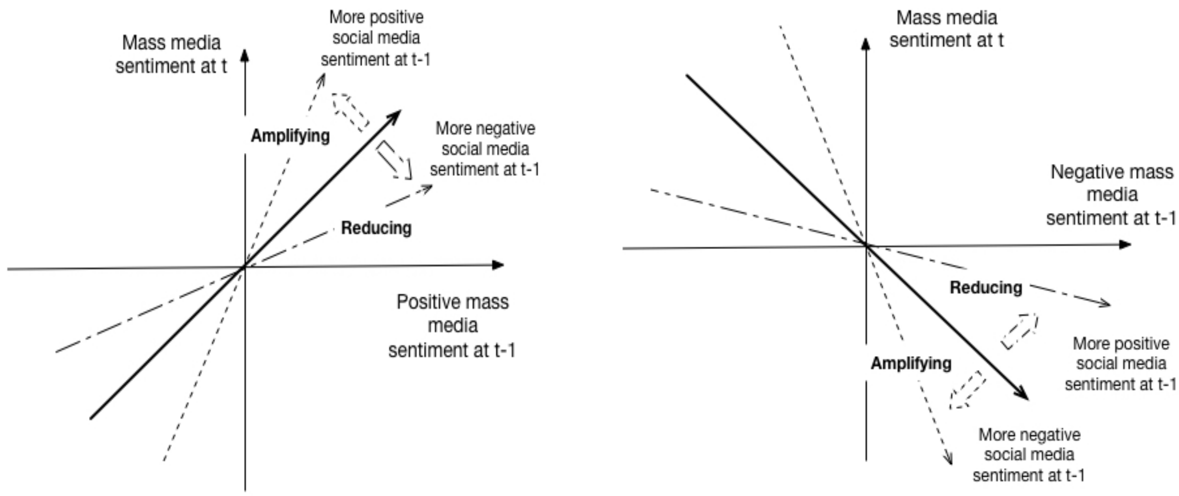


Figure 2. Conceptual Understanding of the Filter Role of Social Media

The competing working mechanism is that mass media may correct social media sentiment. When this mechanism functions, the direction of the coefficient(s) that are explained above is opposite. Specifically, this mechanism predicts that the persistence of mass media sentiment over time decreases if past mass media sentiment and past social media sentiment are consistent, which is the reducing filter role. β_1 is negative and β_4 is positive. The mechanism also suggests that persistence of mass media sentiment over time increases if past mass media sentiment and past social media sentiment are inconsistent, which is the amplifying filter role. That is, β_2 is negative and β_3 is positive.

Table 2 summarizes the explanations of Equation 1.

Table 2. Explanations of Equation 1

β	**Explanation**	**Expectations given the Demand -driven Mechanism (Figure 2)**	**Expectations given the Correcting Mechanism**
Panel A: Effects from Positive Mass Media Sentiment			
β_5	The persistence between mass media sentiment at *t* and positive mass media sentiment at *t-1* (main effect)	$+$	$+$
β_1	The change in the sentiment persistence because of positive social media sentiment at *t-1*	$+$	$-$
β_3	The change in the sentiment persistence because of negative social media sentiment at *t-1*	$-$	$+$
$\beta_5 + \beta_1$	The persistence between mass media sentiment at *t* and positive mass media sentiment at *t-1* after considering positive social media sentiment at *t-1*	$++$ (amplifying)	$+-$ (reducing)
$\beta_5 + \beta_3$	The persistence between mass media sentiment at t and positive mass media sentiment at *t-1* after considering negative social media sentiment at *t-1*	$+-$ (reducing)	$++$ (amplifying)
Panel B: Effects from Negative Mass Media Sentiment			
β_6	The persistence between mass media sentiment at *t* and negative mass media sentiment at *t-1* (main effect)	$-$	$-$
β_2	The change in the sentiment persistence because of positive social media sentiment at *t-1*	$+$	$-$
β_4	The change in the sentiment persistence because of negative social media sentiment at *t-1*	$-$	$+$
$\beta_6 + \beta_2$	The persistence between mass media sentiment at *t* and negative mass media sentiment at *t-1* after considering positive social media sentiment at *t-1*	$-+$ (reducing)	$--$ (amplifying)
$\beta_6 + \beta_4$	The persistence between mass media sentiment at t and negative mass media sentiment at *t-1* after considering negative social media sentiment at *t-1*	$--$ (amplifying)	$-+$ (reducing)

5.2 Results

5.2.1 Main Analysis

We started our analyses using all the data. We documented evidence on two effects – both amplifying and reducing – which confirms the working mechanism of the demand-driven media bias for the sentiment filter role of social media, suggesting that H1a and H1b were both supported, therefore H2a and H2b were rejected (See Colum 1 of Table 3): the persistence of mass media sentiment (from *t-1* to *t*: both positive or both negative) is stronger when social media sentiment at *t-1* is also consistent with mass media sentiment at *t-1* (both positive or both negative; $\beta_1 = 0.0256, \mathrm{p} < 0.01$; $\beta_4 = -0.0430, \mathrm{p} < 0.01$), which confirms the amplifying filter role of social media; and the persistence of mass media sentiment (from *t-1* to *t*: both positive or both negative) is weaker when social media sentiment at *t-1* is conflicting with mass media sentiment at *t-1* (positive and negative, or negative and positive; $\beta_2 = 0.0141, \mathrm{p} < 0.01$; $\beta_3 = -0.0079, \mathrm{p} < 0.05$), which confirms the reducing filter role of social media. The values of fixed effects of the first three stocks in our sample (Tickers: 000001, 000002, 000004) are 0.4921, 0.9901 and 0.0263. We also reported the mean of Variance Inflation Factors for our regressions to make sure that there is no issue of multicollinearity in our data analysis. All individual VIFs are below 10, the threshold documented in Kennedy (2008, pp 199).

We acknowledge that there is an endogeneity issue in this analysis. One example is that sentiment flows from mass media to social media, a more common scenario. The current analysis does not exclude this possibility. In the following subsections, we aim to relieve this endogeneity issue to the extent possible.

Table 3. Estimation Results

	(1) Section 5.2.1	(2) Section 5.2.2	(3) Section 5.2.3	(4) Section 5.2.4
$SentNewsP_{i,t-1} \times SentWeiboP_{i,t-1}$	0.0257***	0.0197***	0.0167***	0.0199***
	(0.0020)	(0.0021)	(0.0030)	(0.0052)
$SentNewsN_{i,t-1} \times SentWeiboP_{i,t-1}$	0.0131***	0.0106**	0.0108	-0.0131
	(0.0034)	(0.0047)	(0.0074)	(0.0135)
$SentNewsP_{i,t-1} \times SentWeiboN_{i,t-1}$	-0.0079***	-0.0017	-0.0081	-0.0129*
	(0.0024)	(0.0037)	(0.0055)	(0.0075)
$SentNewsN_{i,t-1} \times SentWeiboN_{i,t-1}$	-0.0413***	-0.0281***	-0.0320***	-0.0205*
	(0.0033)	(0.0044)	(0.0102)	(0.0123)
$SentNewsP_{i,t-1}$	0.1104***	-0.0102 ***	-0.0034	-0.0213*
	(0.0054)	(0.0036)	(0.0037)	(0.0113)
$SentNewsN_{i,t-1}$	-0.0609***	0.0047	0.0071	0.0435
	(0.0068)	(0.0070)	(0.0081)	(0.0271)
$SentWeiboP_{i,t-1}$	0.0072***	0.0170***	0.0149***	0.0122***
	(0.0005)	(0.0007)	(0.0006)	(0.0009)
$SentWeiboN_{i,t-1}$	0.0003	-0.0028***	-0.0029***	-0.0062***
	(0.0006)	(0.0009)	(0.0007)	(0.0010)
$SentNews_{i,t-2}$	0.0029*	0.0271***		
	(0.0017)	(0.0019)		
$SentNews_{i,t-3}$	0.0216***	0.0235***		
	(0.0018)	(0.0019)		
$SentWeibo_{i,t-2}$	0.0005	0.0001	-0.0012***	-0.0007
	(0.0003)	(0.0004)	(0.0003)	(0.0006)
$SentWeibo_{i,t-3}$	0.0013***	0.0008**	0.0007**	0.0019***
	(0.0003)	(0.0003)	(0.0003)	(0.0005)
Stock Fixed Effects	Yes	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes	Yes
Obs	1,818,227[3]	1,818,227	1,466,035	340,205
Adj. R^2	0.1413	0.1193	0.0414	0.0425
VIF	3.28	2.2	2.14	2.95

Note: *p<0.10, **p<0.05, ***p<0.01. The standard errors are reported in the parentheses. The standard errors are clustered at the stock and day level. In Columns 2 through 4, $SentNewsP_{i,t-1}$ and $SentNewsN_{i,t-1}$ are calculated based on the announcements in the media at *t-1*.

[3] 1,818,227 observations = 2501 stocks × (365 days per year x 2 years – 3 days), where the missing three days are due to the 3-day lagged control variables.

5.2.2 Analysis with SentNews at t-1 as the Sentiment of Announcements in the Media at t-1

In this subsection, we focused on a specific scenario to relieve this endogeneity issue. In China, by law, companies' announcements should first be published through mass media. The Act 6 of the Administrative Measures on Information Disclosure by Listed Companies (China Securities Regulatory Commission, 2008) explicitly states that "the information discloser cannot publish new information in its company webpage or through other channels before the designated mass media…" In other words, company announcement should not be known by the public before it is released in mass media. Mass media sentiment of announcements at *t-1* when the announcements were published for the first time should be irrelevant to mass media sentiment at *t-2* and social media sentiment at *t-2*. Therefore, announcements can be viewed as external shocks in our setting.

We examined whether the titles of news items include "announcements" or not to distinguish the announcements from the other news items. As required by the regulator, the announcements should include "announcement" in their news titles. We verified this by examining the first 500 news items that had "announcement" in their titles in our dataset, and we found they were all announcements.

We recalculated mass media sentiment at *t-1* with announcements, and we reran the above analysis in Equation 1. Column 2 in Table 3 records our results. We found that as positive (negative) social media sentiment at *t-1* and positive (negative) mass media sentiment reporting announcements at *t-1* increase, mass media sentiment at *t* turns to be more positive or more negative ($\beta_1 = 0.0197, \mathrm{p} < 0.01$; $\beta_4 = -0.0281, \mathrm{p} < 0.01$). This suggests the amplifying filter role of social media as documented in H1a.

Interestingly, we did not find strong evidence for the reducing filter role of social media. The coefficients of the interaction terms between negative (positive) mass media sentiment reporting announcements at *t-1* and positive (negative) social media sentiment at *t-1* are not as significant as the amplifying filter role, although the direction of the coefficients is consistent with our prediction ($\beta_2 = 0.0106, \mathrm{p} < 0.05$; $\beta_3 = -0.0017, \mathrm{p} > 0.10$). Therefore, we confirmed that H1b was partially supported.

One concern of our results in this subsection is whether these results are affected by the diffusion speed of information. For example, social media users at *t-1* may still react to prior firm announcements in the media (at *t-2*). We re-ran the regression using the announcements at *t-2* and social media posts at *t-1*. Consistently, we found evidence for the amplifying filter role but not clear evidence for the reducing filter role.

5.2.3 Analysis with SentNews at t-1 as the Sentiment of Announcements in the Mass Media at t-1 & No Mass Media Coverage at t-2 or at t-3

In subsection 5.2.2, we used the sentiment of announcements in the mass media at *t-1* as a proxy of mass media sentiment at *t-1*. Because announcements are published by companies via the media, past mass media sentiment or past social media sentiment is irrelevant. However, social media sentiment at *t-1* might not react to the sentiment of the announcements; rather it might react to the news articles at *t-2*. Put differently, both mass media sentiment at *t* and social media sentiment at *t-1* are driven by mass media sentiment at *t-2*. To address this concern and to further address the endogeneity issue, we focused on a scenario when there is no mass media coverage at *t-2* and *t-3*. We reran the regression and reported the results in Column 3. We found very similar results as in Subsection 5.2.2: β_1 (β_4) is significant and positive (negative) ($\beta_1 = 0.0167, \mathrm{p} < 0.01$; $\beta_4 = -0.0320, \mathrm{p} < 0.01$), consistent with our prediction of the amplifying filter role of social media driven by the demand-driven media bias argument. Therefore, H1a was supported. β_2 and β_3 are not statistically siginicant. We did not find evidence supporting H1b.

5.2.4 Analysis with SentNews at t-1 as the Sentiment of Announcements in the Mass Media at t-1 & No Mass Media Coverage at t-2 or at t-3 & Social Media Sentiment at t-1 in the Opposite Direction of the Sentiment of the Stock Performance

In this subsection, we aim to further address the endogeneity issue: both social media sentiment at *t-1* and mass media sentiment at *t* may be affected by the stock market at *t-1*. As such, the association in our analysis may be driven by stock return at *t-1*. That is, social media sentiment at *t-1* might not react to the sentiment of the announcements in the media (mass media sentiment at *t-1*); rather it might react to the sentiment of the relevant stock performance on the same day.

Therefore, we focused on the scenario where first social media sentiment on the relevant stock's company mentioned in the announcement at *t-1* was in the opposite direction of the stock performance sentiment at *t-1*; and second there was no news coverage on the relevant stock at *t-2* and at *t-3*. The logic here is to find observations where the stock market and the social media express different sentiments. By doing so, we relieve the endogeneity issue.

We repeated the analysis in Column 3 and presented our results in Column 4 in Table 3. As the positive (negative) social media sentiment at *t-1* and positive (negative) mass media sentiment reporting announcements at *t-1* increases, mass media sentiment at *t* will increase to be more positive or decrease to be more negative ($\beta_1 = 0.0199, \mathrm{p} < 0.01$; $\beta_4 = -0.0205, \mathrm{p} < 0.1$). Therefore, we confirmed that H1a was supported. Similar to the previous tests, we did not find evidence for the reducing filter role of social media. The coefficients of the interaction terms between negative (positive) mass media sentiment reporting announcements at *t-1* and positive (negative) social media sentiment at *t-1* are not statistically significant. Here, H1b was not supported.

To summarize, we found strong evidence for the amplifying filter role of social media driven by the mechanism of "demand-driven media bias". However, the reducing filter role driven by the same mechanism is unclear – there is limited evidence. Therefore, consistently, H1a was supported, and inconsistently, H1b was supported. Accordingly, H2a and H2b were both rejected.

5.3 Robustness Tests

In this section, we ran several robustness tests to make sure that our results are not biased because of our choices. The first concern is that our mass media sentiment is estimated using news articles that only mentioned one stock each, following the standard financial literature (Tetlock, 2011). As such, we dropped more than half of the news articles. To relieve this concern, we ran a robustness test, in which we used all news articles, even if some of them each mentioned several stocks. Column 1 in Table 4 reports the results. Second, in the main regression, we controlled for both social media sentiment and mass media sentiment at *t-2* and *t-3*. They may affect the main results. Therefore, in this robustness test, we did not control for these control variables. The results are reported in column 2 in Table 4. Third, another concern is about how we dealt with sentiment during the weekends, when there is no trading. We ran two robustness tests to relieve this concern. In the first test, we kept the observations on days from Tuesdays to Fridays, which do not involve the weekend sentiment. In the second test, following Deng et al. (2018), we combined the social media sentiment (and the news sentiment) during the weekends with those on Fridays (reported) or on Mondays (unreported). The results of two robustness tests are reported in Columns 3 and 4 in Table 4. The results are consistent to the results in Table 3.

Table 4. Estimation Results for Robustness Tests

	(1) All news	(2) Without sentiment at t-2 and t-3	(3) From Tuesday to Friday	(4) Sentiment on Fridays include that on Fridays and Weekends
$SentNewsP_{i,t-1}$ $\times SentWeiboP_{i,t-1}$	0.0145***	0.0258***	0.0248***	0.0239***
	(0.0021)	(0.0020)	(0.0020)	(0.0018)
$SentNewsN_{i,t-1}$ $\times SentWeiboP_{i,t-1}$	0.0077**	0.0133***	0.0134***	0.0173***
	(0.0030)	(0.0034)	(0.0040)	(0.0036)
$SentNewsP_{i,t-1}$ $\times SentWeiboN_{i,t-1}$	-0.0013	-0.0077***	-0.0086***	-0.0037
	(0.0018)	(0.0024)	(0.0027)	(0.0024)
$SentNewsN_{i,t-1}$ $\times SentWeiboN_{i,t-1}$	-0.0528***	-0.0421***	-0.0431***	-0.0409***
	(0.0033)	(0.0033)	(0.0039)	(0.0037)
$SentNewsP_{i,t-1}$	0.2073***	0.1112 ***	0.1372***	0.1042***
	(0.0066)	(0.0054)	(0.0056)	(0.0053)
$SentNewsN_{i,t-1}$	-0.0696***	-0.0613***	-0.0677***	-0.0549***
	(0.0070)	(0.0069)	(0.0082)	(0.0075)
$SentWeiboP_{i,t-1}$	0.0368***	0.0082***	0.0093***	0.0102***
	(0.0023)	(0.0005)	(0.0006)	(0.0006)
$SentWeiboN_{i,t-1}$	0.0074***	0.0000	-0.0001	0.0010
	(0.0017)	(0.0006)	(0.0007)	(0.0007)
$SentNews_{i,t-2}$	-0.0037*		0.0111	0.0020
	(0.0022)		(0.0023)	(0.0022)
$SentNews_{i,t-3}$	0.0375***		0.0262	0.0063***
	(0.0024)		(0.0022)	(0.0019)
$SentWeibo_{i,t-2}$	0.0051		-0.0005	0.0008*
	(0.0009)		(0.0005)	(0.0005)
$SentWeibo_{i,t-3}$	0.0113***		0.0004	0.0004
	(0.0008)		(0.0004)	(0.0005)
Stock Fixed Effects	Yes	Yes	Yes	Yes
Day Fixed Effects	Yes	Yes	Yes	Yes
Obs	1,818,227	1,818,227	1,466,035	340,205
Adj. R^2	0.1413	0.1193	0.0414	0.0425
VIF	3.28	2.2	2.14	2.95

Note: *p<0.10, **p<0.05, ***p<0.01. The standard errors are reported in the parentheses. The standard errors are clustered at the stock and day level.

6. DISCUSSION

6.1 Implications

Our paper contributes to the literature in numerous important ways. First, we answered an important question: does sentiment emerge in each media type independently and perhaps randomly? Prior studies implicitly assumed that social media and mass media are independent (Dewan & Ramaprasad, 2014; Greenstein & Zhu, 2018). A recent study has shown possible spillover effects from social media into the broader domain (Botting, Dipper, & Hilari, 2017), we present evidence that mass media and social media are associated. In a utilitarian market like the financial market, the likelihood for mass media to follow social media in setting the sentiment of news articles should be low. Therefore, our findings are surprising and counterintuitive.

Second, our findings here potentially also suggest a dark side of social media – mass media news is not always objective and accurate in the financial markets. This is dangerous as it opens up opportunities for social media to "manipulate" the media environment in the financial market and cause readers to have reinforced (biased) opinions, especially for investment decisions. Prior papers focused on the supply-factors of the media bias in this market (Ahern & Sosyura, 2014; Reuter & Zitzewitz, 2006). Our paper complements this literature by providing empirical evidence that demand-driven media bias also exists in the financial market.

Third, our paper has strong contributions to the sentiment literature in the stock market (Deng et al., 2018; Luo et al., 2013; Sul, Dennis, & Yuan, 2017). There is an alternative way how social media sentiment affects the stock market. Social media may affect individual investors, and consequently, is associated with investor behaviors and stock returns. In this paper, we documented a different mechanism, in which social media not only affects its own users, but also affects a broader audience that includes mass media readers, who may later make trading decisions related to the stock market.

Practically speaking, investors in financial markets need to think more critically about where the sentiment of the news is coming from – whether it is 'objective' in a sense, or only capturing investor opinions reflected in social media. Given our findings that reveal "demand-driven media bias" in a way, investors are likely better off focusing on the fundamental information underlying the news rather than on the sentiment presented there, if their objective is to find the fair value of firms.

6.2 Limitations and Future Work

Our paper is made possible because of a unique opportunity that Sina.com provided us with: Our dataset includes data from both platforms, Sina Finance and Sina Weibo, and enables a systematic look at the sentiment flow from one platform to another. This dataset, however, is not very recent – it ranges over 2013 and 2014 – and we did not have access to the content of the microblogs. Future researchers might consider studying this sentiment flow using more recent data, and might also consider analyzing the microblogs using formal attributes such as length as well as semantic attributes such as the speech acts employed.

Another generalizability issue lies in the Chinese stock market that we studied. In themselves, both the Chinese stock market and Chinese social media are important. One important characteristic of China is a tradition of news censorship. However, we do not think censorship affected our findings. Censorship is weaker in the stock market context than in a political setting. Second, censorship is applied more strongly to mass media than to social media. That is, censorship, if any, weakens the association between mass media and social media, meaning it would be more difficult to find significant results about the association.

Some may ask if our findings can be replicated in a political setting. Individuals' opinions on politics are likely to be persistent (for one party and against another), but that is not the case for investors' opinions in the stock market; an investor can change his/her sentiment depending on the ever-changing stock price. Theoretically, in the stock market setting identifying the sentiment flow should be more challenging compared to that in the political setting. Future studies could examine the sentiment flow in other settings.

7. CONCLUSION

Mass media sentiment in financial markets can affect investor decisions and hence deserves careful consideration. However, there has been little investigation into exactly *how* mass media sentiment emerges for stock-related news. This paper specifically examined a question: how does social media sentiment affect mass media sentiment in financial markets? We argued for, and tested, the sentiment filter role of social media: social media may modify the persistence of mass media sentiment over time, either amplifying or reducing it, driven possibly by two competing mechanisms – (1) the mechanism of the demand-driven media bias and (2) the mechanism of mass media correcting social media's noises (if any). We found evidence for the sentiment amplifying filter role of social media driven by the demand-driven media bias. We found limited evidence for the sentiment reducing filter role that is driven by the same mechanism.

REFERENCES

Ahern, K. R., & Sosyura, D. (2014). Who Writes the News? Corporate Press Releases during Merger Negotiations. *Journal of Finance*, *69*(1), 241–291. https://doi.org/10.1111/jofi.12109

Ahern, K. R., & Sosyura, D. (2015). Rumor has it: Sensationalism in financial media. *The Review of Financial Studies*, *28*(7), 2050–2093.

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236.

Ansolabehere, S. (2006). The Orientation of Newspaper Endorsements in U.S. Elections, 1940--2002. *Quarterly Journal of Political Science*, *1*(4), 393–404. https://doi.org/10.1561/100.00000009

Antweiler, W., & Frank, Z. M. (2004). Is all that talk just noise? the information content of Internet stock message boards. *Journal of Finance*, *59*(3), 1259–1294. https://doi.org/10.1017/CBO9781107415324.004

Botting, N., Dipper, L., & Hilari, K. (2017). The effect of social media promotion on academic article uptake. *Journal of the Association for Information Science and Technology*, *68*(3), 795–800.

China Securities Regulatory Commission. (2008). the Administrative Measures on Information Disclosure by Listed Companies.

Da, Z., Engelberg, J., & Gao, P. (2011). In Search of Attention. *Journal of Finance*, *66*(5), 1461–1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x

Deng, S., Huang, Z. (James), Sinha, A. P., & Zhao, H. (2018). The interaction between microblog sentiment and stock returns: An empirical examination. *MIS Quarterly*, *42*(3), 895–918. https://doi.org/10.25300/MISQ/2018/14268

Dewan, S., & Ramaprasad, J. (2014). Social media, traditional media, and music sales. *Mis Quarterly*, *38*(1), 101–121.

Dong, H., & Gil-Bazo, J. (2020). Sentiment stocks. *International Review of Financial Analysis*, *72*, 101573. https://doi.org/https://doi.org/10.1016/j.irfa.2020.101573

Dougal, C., Engelberg, J., Garcia, D., & Parsons, C. A. (2012). Journalists and the stock market. *The Review of Financial Studies*, *25*(3), 639–679.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. California: Stanford University Press.

García, D. (2013). Sentiment during Recessions. *Journal of Finance*, *68*(3), 1267–1300. https://doi.org/10.1111/jofi.12027

Gentzkow, M., & Shapiro, J. M. (2008). Competition and Truth in the Market for News. *Journal of Economic Perspectives*, *22*(2), 133–154.

Gentzkow, M., & Shapiro, J. M. (2010). What drives media slant? Evidence from US daily newspapers. *Econometrica*, *78*(1), 35–71.

Gentzkow, M., Shapiro, J. M., & Stone, D. F. (2015). Media Bias in the Marketplace. In *Handbook of*

Media Economics (Vol. 1, pp. 623–645). Elsevier B.V. https://doi.org/10.1016/B978-0-444-63685-0.00014-0

Greenstein, S., & Zhu, F. (2018). Do experts or crowd-based models produce more bias? Evidence from encyclopedia britannica and wikipedia. *MIS Quarterly*, *42*(3), 945–959. https://doi.org/10.25300/MISQ/2018/14084

Harmon-Jones, E. E., & Mills, J. E. (1999). Cognitive dissonance: Progress on a pivotal theory in social psychology. In *Scientific Conferences Program, 1997, U Texas, Arlington, TX, US; This volume is based on papers presented at a 2-day conference at the University of Texas at Arlington, winter 1997*. American Psychological Association.

Harwit, E. (2014). The Rise and Influence of Weibo (Microblogs) in China. *Asian Survey*, *54*(6), 1059–1087. https://doi.org/10.1525/as.2014.54.6.1059

Hu, Z., Liu, W., Bian, J., Liu, X., & Liu, T.-Y. (2018). Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (pp. 261–269). ACM.

Huang, D., Jiang, F., Tu, J., & Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *Review of Financial Studies*, *28*(3), 791–837. https://doi.org/10.1093/rfs/hhu080

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

Larcinese, V., Puglisi, R., & Snyder, J. M. (2011). Partisan bias in economic news: Evidence on the agenda-setting behavior of U.S. newspapers. *Journal of Public Economics*, *95*(9–10), 1178–1189. https://doi.org/10.1016/J.JPUBECO.2011.04.006

Li, S., Song, X., & Wu, H. (2015). Political Connection, Ownership Structure, and Corporate Philanthropy in China: A Strategic-Political Perspective. *Journal of Business Ethics*, *129*(2), 399–411. https://doi.org/10.1007/s10551-014-2167-y

Liang, Y., & Shen, W. (2016). Fan economy in the Chinese media and entertainment industry: How feedback from super fans can propel creative industries' revenue. *Global Media and China*, *1*(4), 331–349.

Luo, X., Zhang, J., & Duan, W. (2013). Social media and firm equity value. *Information Systems Research*, *24*(1), 146–163.

Mai, F., Shan, Z., Bai, Q., Wang, X., & Chiang, R. H. L. (2018). How does social media impact Bitcoin value? A test of the silent majority hypothesis. *Journal of Management Information Systems*, *35*(1), 19–52.

Mullainathan, S., & Shleifer, A. (2005). The market for news. *American Economic Review*, *95*(4), 1031–1053.

Park, J., Konana, P., Gu, B., Kumar, A., & Raghunathan, R. (2013). Information valuation and confirmation bias in virtual communities: Evidence from stock message boards. *Information Systems Research*, *24*(4), 1050–1067. https://doi.org/10.1287/isre.2013.0492

Puglisi, R., & Snyder Jr, J. M. (2015). Empirical studies of media bias. In *Handbook of media economics* (Vol. 1, pp. 647–667). Elsevier.

Renault, T. (2017). Intraday online investor sentiment and return patterns in the U.S. stock market. *Journal of Banking and Finance*, *84*, 25–40. https://doi.org/10.1016/j.jbankfin.2017.07.002

Reuter, J., & Zitzewitz, E. (2006). Do ads influence editors? Advertising and bias in the financial media. *The Quarterly Journal of Economics*, *121*(1), 197–227.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, *19*(1), 22–36.

Simon, H. a. (1982). *Models of bounded rationality. History* (Vol. 2). MIT Press. https://doi.org/citeulike-article-id:1204470

Sul, H. K., Dennis, A. R., & Yuan, L. (2017). Trading on twitter: Using social media sentiment to predict stock returns. *Decision Sciences*, *48*(3), 454–488.

Tetlock, P. C. (2007). Giving content to investor sentiment : The role of media in the stock market. *Journal of Finance*, *62*(3), 1139–1168.

Tetlock, P. C. (2011). All the news that's fit to reprint: Do investors react to stale information? *Review of Financial Studies*, *24*(5), 1481–1512. https://doi.org/10.1093/rfs/hhq141

Tetlock, P. C. (2014). Information transmission in finance. *Annual Review of Financial Economics*, *6*(1), 365–384. https://doi.org/10.1146/annurev-financial-110613-034449

Tetlock, P. C. (2015). The role of media in finance. In *Handbook of media Economics* (Vol. 1, pp. 701–721). Elsevier.

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, *63*(3), 1437–1467.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Zheng, S., Wang, J., Sun, C., Zhang, X., & Kahn, M. E. (2019). Air pollution lowers Chinese urbanites' expressed happiness on social media. *Nature Human Behaviour*, 1.

APPENDIX

Tables A.1, A.2, and A.3, respectively, list the definitions of all variables, summary statistics, and correlations among all variables. Table A.4 shows the distributions of Sina Weibo posts and news articles. Table A.5 reports top 10 mentioned stocks in Weibo and in news.

Table A.1. Definitions of Variables

Acronym	Variable	Measure
SentNews	Mass media sentiment	The logarithm value of one plus the number of positive news articles over one plus the number of negative news articles on stock *i* at day *t*.
SentNewsP (Column 1, Table 3; Table 4)	Positive mass media sentiment	The logarithm value of one plus the number of positive news articles on stock *i* at day *t*.
SentNewsN (Column 1, Table 3; Table 4)	Negative mass media sentiment	The logarithm value of one plus the number of negative news articles on stock *i* at day *t*.
SentNewsP (Columns 2-4, Table 3)	Positive mass media sentiment of announcements	The logarithm value of one plus the number of positive announcements in the media on stock *i* at day *t*.
SentNewsN (Columns 2-4, Table 3)	Negative mass media sentiment of announcements	The logarithm value of one plus the number of negative announcements in the media on stock *i* at day *t*.
SentWeibo	Social media sentiment	The natural logarithm value of one plus the number of positive weibos over one plus the number of negative weibos on stock *i* at day *t*.
SentWeiboP	Positive social media sentiment	The logarithm value of one plus the number of positive weibos on stock *i* at day *t*.
SentWeiboN	Negative social media sentiment	The logarithm value of one plus the number of negative weibos on stock *i* at day *t*.

Table A.2. Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	Max.
SentNews	0.05911	0.29028	-3.33221	3.89182	3.89182
SentNewsP (that measures positive mass media sentiment)	0.08396	0.29652	0.00000	4.82831	4.82831
SentNewsN (that measures negative mass media sentiment)	0.02498	0.15727	0.00000	4.80402	4.80402
SentNewsP (that measures positive mass media sentiment of announcements)	0.01794	0.12434	0.00000	2.56495	2.56495
SentNewsN (that measures negative mass media sentiment of announcements)	0.00382	0.05770	0.00000	1.79176	1.79176
SentWeibo	0.87459	1.10245	-7.41391	11.16886	11.16886
SentWeiboP	1.25532	1.28830	0.00000	12.21772	12.21772
SentWeiboN	0.38073	0.75409	0.00000	11.79473	11.79473

Note: The number of observations are 1,818,227.

Table A.3. Correlation Matrix

	1	2	3	4	5	6	7	8
1. SentNews	1							
2. SentNewsP (that measures positive mass media sentiment)	0.2843	1						
3. SentNewsN (that measures negative mass media sentiment)	0.0781	0.3061	1					
4. SentNewsP (that measures positive mass media sentiment of announcements)	0.0403	0.4633	0.0344	1				
5. SentNewsN (that measures negative mass media sentiment of announcements)	0.0038	0.0339	0.3991	-0.0096	1			
6. SentWeibo	0.0908	0.1284	-0.03	0.0345	-0.0346	1		
7. SentWeiboP	0.1405	0.253	0.1203	0.0548	0.0134	0.812	1	
8. SentWeiboN	0.1074	0.2446	0.2494	0.0432	0.0735	-0.0748	0.5214	1

Note: All correlations are significant at the level of 0.01.

Table A.4. Distributions of Sina Weibo Posts and News Articles

Panel A. Distribution of Sina Weibo Posts from 2013 to 2014

Min	10%	25%	50%	75%	90%	Max
658	4551	6565	10011	16642	30974	3972731

Panel B. Distribution of News Articles from 2013 to 2014

Min	10%	25%	50%	75%	90%	Max
2	39	62	98	156	250	16203

Table A.5. Top 10 Mentioned Stocks in Weibo and in News

Panel A. Top 10 mentioned Stocks in Weibo

Ticker	Num. of Weibo	Num. of News	Percentage (Weibo)	Percentage (News)
300111	3972731	319	6.7673%	0.0705%
300146	1352237	280	2.3034%	0.0619%
002594	938629	2067	1.5989%	0.4566%
002029	877804	183	1.4953%	0.0404%
600332	733050	708	1.2487%	0.1564%
600663	660167	2237	1.1246%	0.4942%
600016	586337	3034	0.9988%	0.6703%
600859	582708	667	0.9926%	0.1474%
601166	566510	2063	0.9650%	0.4558%

Panel B. Top 10 mentioned Stocks in News

Ticker	Num. of Weibo	Num. of News	Percentage (Weibo)	Percentage (News)
000061	85167	16203	0.1451%	3.5795%
601099	45840	9647	0.0781%	2.1312%
601988	226378	8755	0.3856%	1.9341%
601857	84191	5596	0.1434%	1.2362%
601519	46491	5388	0.0792%	1.1903%
000002	215672	5361	0.3674%	1.1843%
601398	197072	5079	0.3357%	1.1220%
600036	408187	4392	0.6953%	0.9703%
600837	92281	3131	0.1572%	0.6917%
600016	586337	3034	0.9988%	0.6703%