Does Interaction on Social Media Drive Extremeness or Moderation?*

Devin Shanthikumar
Annie Wang
Shijia Wu†

The Paul Merage School of Business
University of California, Irvine

October 14, 2020

Abstract

Using comment streams on Seeking Alpha articles, we examine whether interacting on social media increases or moderates the extremeness of investors' opinions. Unlike findings from political science, we find that interaction moderates extremeness. Comments become less extreme over the sequence of comments for a given article, within individual comment sub-threads, and over a single user's comments for a given article. Extremeness moderation occurs both following earnings announcements and during non-earnings-announcement windows. Extremeness reduction is stronger when the article users are commenting on is more moderate, and when more commenters are self-identified (i.e., not anonymous). Finally, results suggest that the extremeness reduction triggered by Seeking Alpha articles has market implications. Differences of opinion captured by stock-based measures decrease significantly after Seeking Alpha articles with comments, but not after analyst forecast days or high-news days. Our results provide the first evidence of the influence of social media interaction on the updating of individuals' opinions.

* We thank Lucile Faurel, Thomas Schofi, and workshop participants at the University of San Diego, the AAA Western Regional Doctoral Interchange, and the 2020 American Accounting Association Annual Meeting for helpful comments. We thank the Accounting Area at the Paul Merage School of Business for providing funding for this research. We thank Sahej Kaur Arora for valuable research assistance. We also thank Seeking Alpha for providing article and comment data.
† Devin Shanthikumar, dshanthi@uci.edu (corresponding author); Annie Wang, qiaow4@uci.edu; Shijia Wu, shijiawu@uci.edu.

1. Introduction

Social media is pervasive in modern life. We interact with friends and strangers online, share our thoughts and experiences, and learn from others about everything from investing to exercise to how to install a particular type of lightbulb. The impacts of our widespread social media use are only beginning to be understood. In financial markets, social media serves as an important and relatively new type of information intermediary. Investors obtain information from social media, share information and opinions, and interact with one another online. Research outside of the financial market setting suggests that increasing polarization in our society is in part driven by social media interaction (e.g., Bail et al. 2018). Yet whether social media drives extremeness and polarization in the context of stock market analysis is unknown. It could, conversely, lead to a moderation of extremeness, as individuals with divergent opinions share and exchange ideas.

We address whether social media increases or moderates extremeness of opinions in the context of stock market analysis by exploiting comments on Seeking Alpha (SeekingAlpha.com) articles. One of the unique aspects of social media, as a financial intermediary, is that it allows for direct interaction among market participants. We examine the effects of this unique aspect. This question also has implications for financial markets. Given short-sale constraints and costs, opinion extremeness and disagreement are likely to contribute to overpricing and stock price crash risk (e.g., Diether, Malloy, and Scherbina 2002; Hong and Stein 2003). Thus, if social media use increases or decreases such opinion extremeness, it is likely to have market implications.

Seeking Alpha provides a unique setting with which to examine these questions as we can observe user's self-disclosed opinions over time, anyone posting comments can see the same set of preceding posts regardless of who they follow or are friends with, and individuals' posts form a "thread" with a clear sequence. Seeking Alpha also has the scale and scope to potentially impact financial markets (Chen, De, Hu and Hwang 2014; Farrell, Green, Jame and Markov 2020).

Seeking Alpha provides curated articles about publicly traded stocks, covering over 8,000 stocks, and providing over 7,000 articles a month, from a pool of over 10,000 possible authors. In addition, millions of individuals access Seeking Alpha to research stocks, and to contribute to the site in the form of commenting on articles.[3] We exploit this setting to first examine whether and how the extremeness of opinions expressed by commenters changes with interaction. In particular, we examine whether comments become more or less extreme as a comment stream for a particular article continues, both across users and within-user. Importantly, we employ article fixed-effects, thus controlling for firm-time-specific factors and article-specific factors which may drive extremeness of opinions. If social media discussion drives increasing extremeness, we should find more extreme comments later in a given comment stream. Conversely, if interaction on the Seeking Alpha moderates extremeness, we should see lower extremeness in later comments. Next, we examine specific drivers of an increasing- or moderating-extremeness effect, motivated by prior research on Internet psychology. Finally, we examine the potential effects of Seeking Alpha interaction on financial markets.

In order to gauge extremeness, we make use of a machine-learning based text analysis tool offered by Amazon Web Services, Amazon Comprehend, which allows us to classify both the directional sentiment and the extremeness of that sentiment expressed in each Seeking Alpha comment. This algorithm has been trained on Amazon's vast library, including items such as Amazon product reviews. We find that the type of free-form and colloquial text used in many of the Seeking Alpha comments is effectively analyzed using such a tool. We discuss methods used to evaluate and validate alternative text analysis tools in an appendix. Using the output from Amazon Comprehend, we construct three primary measures of extremeness, capturing slightly

[3] See https://seekingalpha.com/page/about_us for more information. Specific numbers taken from https://seekingalpha.com/page/about_us, accessed 6/13/2019.

different dimensions of comment extremeness: absolute extremeness, relative extremeness, and use of strong/extreme language. These measures allow us to differentiate among three alternative hypotheses for the effect of social media interaction: *convergence,* or *depolarization,* (decreasing absolute and relative extremeness), *group polarization* (increasing absolute extremeness but decreasing relative extremeness), and *belief polarization* (increasing absolute and relative extremeness).

We examine how extremeness evolves at several levels. First, we examine how extremeness evolves at the article level (e.g. with article fixed effects), controlling for user fixed effects. Second, we examine how extremeness evolves for a given user within a single article's comment stream. Third, we examine how extremeness evolves for individual sub-threads, in which users are responding directly to the preceding comment. The results from all of these analyses indicate that interaction on Seeking Alpha *moderates* the extremeness of users' opinions. Comments become less extreme later in a sequence of comments. These results indicate that seeking alpha interaction drives convergence of opinions. While a large number of firm characteristics, such as fundamental uncertainty, performance, and the firm's information environment, might affect the overall extremeness of opinions about the firm, our research design, employing a within-article design, controls for such firm-time characteristics. Our analyses examining within-user within-article variation further control for self-selection of users' commenting behavior – e.g. it is not the case that users with more extreme opinions comment earlier. The *same* user becomes less extreme as commenting progresses, and converges with other users. Together, these results provide strong and consistent evidence of social media interaction driving opinion convergence for stock-related discussions.

We further explore two key differentiators of Seeking Alpha, as compared to other social media platforms, which may drive moderation rather than increasing extremeness. First, the topic of Seeking Alpha – stock analysis – may be inherently less controversial than many topics discussed online, such as politics and social welfare topics. It may be the case that polarization is more likely to occur for inherently controversial topics, while moderation is more likely when the topic lacks that inherent controversy. If this is the case, we should find less of a moderation effect for the subset of Seeking Alpha articles which are more controversial, which we measure as articles which themselves contain more extreme language. Our results are consistent with this hypothesis. Convergence is stronger for less extreme articles. Second, participants on Seeking Alpha often self-identify through voluntary biographies. Prior research suggests that anonymity may be a driving factor behind extremeness online (see, e.g., Scott 2004; Joinson, McKenna, Postmes, and Reips 2007; Qian and Scott 2007). Moreover, anonymity is likely to be a driving factor behind polarization (see, e.g., the discussion in Bail et al.). Thus, we would expect the moderation effect to be stronger when more of the users commenting on a given article have self-identified. Our results are consistent with this prediction as well. Overall, these results indicate both the robustness of extremeness-moderation in the Seeking Alpha setting, and important characteristics of Seeking Alpha which enhance the extremeness-moderation effect. These results can be informative when we consider other social media platforms, and point to important factors to consider – extremeness of the underlying topics and user anonymity.

Finally, we provide initial evidence on whether discussions on Seeking Alpha decrease differences of opinion in capital markets. We examine the evolution of market-based difference of opinion measures subsequent to the release of Seeking Alpha articles with comments. Prior research (e.g., Chen, De, Hu and Hwang 2014; Farrell, Green, Jame and Markov 2020) suggests that users on Seeking Alpha are capital market participants who trade based upon the information they gain, and the sentiment expressed, on Seeking Alpha. Therefore, we would expect the convergence of opinions on Seeking Alpha to be reflected in decreases in difference of opinion in the capital market. We find that market-based differences of opinion decrease significantly upon the publication of a Seeking Alpha article with comment activity, with the effect persisting for at least five trading days. In contrast, there is little evidence of decreases in differences of opinion after analyst earnings forecast days or high-news days. While these results do not establish a causal link between Seeking Alpha and changes in market differences of opinion, they are consistent with such a link, and suggest that the observed pattern is unlikely to be driven by news events in general.

Our study contributes to a broader understanding of social media's effects on society, a question of interest to many fields, including sociology, social psychology, economics, and law. Within finance and accounting, our research contributes to a growing literature on the Internet as an information intermediary (see Lerman 2020 for a description of over 50 papers in this area). Despite this large and quickly growing literature, our paper is the first, to our knowledge, to examine how interaction on social media affects the updating of investors' opinions. Active and wide-spread interaction is one of the aspects that differentiates social media from more traditional information intermediaries such as the press and security analysts. Our study contributes to our understanding of social media as an intermediary.

The remainder of the paper is structured as follows. Section 2 describes relevant background and institutional details, and develops hypotheses. Section 3 describes the data and defines extremeness measures. Section 4 presents results, and Section 5 concludes.

2. Background and Hypothesis Development

In this section, we discuss relevant research and institutional information, and develop our primary hypotheses.

2.1 *Interaction and Extremeness of Opinions*

A long history of research in social psychology has examined the impact that social interaction has on the extremeness of our opinions. Group polarization and attitude polarization are well studied, and show many instances in which interaction in a group results in individual members holding opinions which are more extreme not only than their initial opinions, but that are more extreme than the most extreme individual opinion of any of the members at the start of the interaction (Kelly 2008; Lord, Ross, and Lepper 1979; Moscovici and Zavalloni 1969; Myers and Lamm 1976; Sunstein 2002).

More recently, research has begun to examine whether such polarization effects occur online. However, much of this research focuses on inherently polarizing topics. One of the areas in which this research is best developed is in political science. As Boxell, Gentzkow and Shapiro (2017) summarize, research on political polarization has led researchers to widely differing opinions, with some (e.g., Sunstein 2017) arguing that the Internet plays a fundamental role in increasing political polarization in our society, others (e.g., Boxell, Gentzkow and Shapiro 2017) finding results which suggest polarization is unrelated to Internet use,[4] and yet others (e.g., Barbera 2015) finding results consistent with Internet interaction decreasing polarization.

There are three primary predictions for the effect which interaction, and in particular social media interaction, might have on individuals' opinions. The first is *attitude polarization*, also referred to as belief polarization: individuals' opinions become more extreme and divergent through interaction. The second is *group polarization*, in which the opinions of each individual in the group become more extreme than previously-held opinions, but in the same direction. The third is *convergence,* also referred to as depolarization, in which opinions converge to a less extreme

[4] Boxell, Gentzkow, and Shapiro (2017) find that the growth in political polarization in recent years is largest for those older than 75, of which the demographic groups the least likely to use the internet and social media.

opinion than previously held by the most extreme individuals in the group. These alternate outcomes of group dynamics are explored in a large body of social psychology research.[5] However, this body of research largely deals with small groups which interact in person. Research is just beginning to extend these into the social media setting, and it is still not well understood how group dynamics will affect opinion updating online. In order to differentiate among these alternative predictions, we must examine both how extreme individuals' stated opinions are relative to a neutral opinion (absolute extremism), and how extreme individuals' opinions are relative to other members of the group (relative extremeness).

Focusing on research related to the stock market, a growing body of research examines the Internet in general, and social media in particular, as an information intermediary (see Lerman, 2020, for an overview of the current literature, summarizing over 50 papers in the area). One stream of research examines how firms use social media to communicate to investors. Another stream examines the value relevance of posts on social media, and the predictive ability of social media users, together, for firms' earnings, stock prices, and trading volume. For example, in a concurrent working paper, Farrel, Green, Jame and Markov (2020) find that retail investors trade based upon information they glean from Seeking Alpha articles and comments, and that such trading is more informed than trading prior to the Seeking Alpha articles. However, to the best of our knowledge, no prior literature in these fields has examined whether social media *interaction* influences individuals' opinion updating beyond facilitating access to information. We attempt to bring together this research on social media as a financial market intermediary with research on group dynamics from social psychology.

[5] See, e.g., "Social Psychology" by Kassin, Fein, and Marcus, or "Social Psychology" by Aronson, Wilson, Akert and Sommers, for overviews of these topics and an introduction to relevant research.

While our primary question is whether interaction on social media affects how we update our opinions about stocks, we are also interested in what drives any effects we observe. In particular, what factors increase the likelihood that a given one of the three predicted outcomes occurs? There are specific aspects of online interaction regarding stocks that are likely to differ from other online interactions. The first, as discussed above, is whether the underlying topic is inherently controversial. Vinoker and Burnstein (1978) find that depolarization is more likely when the topic being discussed is less controversial, whereas attitude polarization is more likely when the underlying topic is more contraversial. A second difference between our setting and that examined in some prior studies of social media is that users often self-identify on platforms such as Seeking Alpha. Within our data, 26.2% of users have a non-blank self-reported user biography. Many users, particularly article writers, participate to improve their reputations, and in some cases to attract business to their analysis or advising practices. Amichai-Hamburger (2007) argues that one of the key differentiators between online and face-to-face communication is anonymity. One of the key effects of anonymity is that it frees people to communicate without the constraints of social norms. This is supported by research in Internet psychology, social psychology, and the literature on computer-mediated communication (e.g., Scott 2004; Joinson, McKenna, Postmes, and Reips 2007; and Qian and Scott 2007). We predict that (a) non-controversial underlying topics, and (b) identified users, are both more likely to lead to *decreases* in opinion extremeness (i.e., convergence), rather than increases in extremeness (i.e., attitude or group polarization).

Finally, we are interested in the impact that these effects on investors' opinion updating have on financial market. Prior literature (e.g., Chen, Hong, Stein 2001; Diether, Malloy, and Sherbina 2002; and Hong and Stein 2003) suggests that differences of opinion are likely to contribute to overpricing and stock price crash risk given short-sales constraints. If interactions on

Seeking Alpha increase the extremeness of users' opinions, particularly if they lead to attitude polarization, then the differences of opinion among users would likely increase on average. On the other hand, if interaction moderates users' extremeness of opinion, it is likely that individuals to some extent are persuaded by opposing views and reduce differences of opinion. The natural question is how this online interaction will impact market differences of opinion. Given that users on Seeking Alpha are likely to be investors (Chen, De, Hu and Hwang, 2014), it is plausible that the changes in their beliefs driven by Seeking Alpha interaction will spill over to the market. To test this, we examine how market-based difference of opinion measures evolve around the release of Seeking Alpha articles, and benchmark events. We predict that differences of opinion expressed in the stock market will *decrease* after the publication of Seeking Alpha articles, consistent with the decrease in opinion extremeness on Seeking Alpha.

2.2 Seeking Alpha and Specific Hypotheses

Seeking Alpha, founded in 2004, has developed into an industry leader in crowd-sourced security analysis. By April 1, 2019, there are a total of 954,343 articles, 17.2 million comments, and 16,400 contributors on Seeking Alpha. There are two primary types of information provided on the site: (1) analysis articles written by contributors, who can be individuals or organizations, and (2) news edited by Seeking Alpha editors, which is mainly news released from other sources, such as Wall Street Journal, and summarized on Seeking Alpha (the source of the piece of news is usually provided as a link in the news).

There are primarily two types of actors on Seeking Alpha: (1) "authors" who provide content (the authors of analysis articles and editors of news articles) and (2) "users" who are consumers of contents. However, Seeking Alpha is a user-generated content platform, in which "contributors" and "users" can overlap. In addition, both contributors and non-contributing users can view, comment, like, and reply to others' comments under each analysis article and news story. Thus, even users who do not "contribute" in the sense of writing articles or editing news stories will often contribute in the sense of writing and responding to comments. Article comments often form a sort of dialogue among users, with many users commenting more than once on a given article, and/or responding directly to other users' comments.

Our focus is on analysis articles and the comment streams which they prompt. Given the discussion above, it is unclear whether interaction on Seeking Alpha will increase or moderate extremeness in users' comments. Seeking Alpha has two characteristics which might be expected to lead to no effect or moderation of extremeness. First, Seeking Alpha users see all comments when they read comments to an article, rather than just those of a particular social network such as Facebook friends or Twitter accounts they follow. However, Bail, et. al. (2018) show in an intriguing experiment that exposing individuals to *opposing* views on social media can further increase political polarization. Similarly, it is not clear whether being exposed to differing opinions about a stock-related article will increase or decrease the extremeness of a user's opinions. Second, Seeking Alpha comments are moderated, and, at least in theory, comments which do not adhere to Seeking Alpha's comment guidelines can be deleted. However, a large set of social media sites, including Facebook, are also moderated (see Appendix C for further discussion and examples), and it is not clear whether moderation has a significant impact on polarization outcomes. Thus we state our primary hypothesis in the null form, under the null that interaction on Seeking Alpha will have no directional impact on users' extremeness. Our primary hypothesis is as follows,

H1. User comments will become no more and no less extreme over the course of users' interaction for a given article.

In order to test H1, we focus on the evolution of extremeness over the comments for individual articles, within parent-reply streams (comment threads) for a given article, and over a single user's comments for a given article.

While we do not make a specific prediction of the direction of the effect for H1, we do make predictions regarding cross-sectional variation in the effect. In particular, given the discussion in Section 2.1, we predict:

H2a. User comments will become more extreme over the course of users' interaction for a given article if the article is more inherently controversial, relative to the change in extremeness which occurs for less inherently controversial articles.

H2b. User comments will become less extreme over the course of users' interaction for a given article if more of the users are identified, via biographies, relative to the change in extremeness which occurs when fewer users are identified.

Finally, we expect the direction of market spillover effect to be consistent with the direction of changing extremeness on Seeking Alpha. If interactions on Seeking Alpha increase the extremeness of users' opinions relative to the mean, we expect market differences of opinion to increase upon the release of a Seeking Alpha article and over the course of the active commenting period. If, in contrast, the extremeness of users' opinions relative to the mean decreases – e.g., Seeking Alpha interaction moderates extremeness – we expect market differences of opinion to decrease. Similarly to H2, we state the market implication hypothesis in the alternative form, however the directional implication depends on the results of testing H1:

H3. Market difference of opinions will change in the same direction as the change in opinion extremeness on Seeking Alpha, subsequent to the publication of a Seeking Alpha article with comments.

3. Data and Sample

3.1 Sample Construction

We begin with the full sample of 74,866 articles on SeekingAlpha.com which are published on the site between April 2018 and April 2019. Opinion articles are written by authors who provide detailed analysis of a stock or a group of stocks.[6] We restrict our analysis to single-ticker opinion articles (similar to Chen, De, Hu and Hwang, 2014, and Gomez, Heflin, Moon, and Warren, 2018) covering common stock (i.e., CRSP share code of 10 and 11) traded on the major exchanges. We focus on slightly higher-visibility stocks which are likely to have a sufficient volume of articles and comments by requiring that the given stock is covered by sell-side analysts in the IBES database. This process leaves us with 18,736 articles. Among the 18,736 single-ticker opinion articles, 1,534 articles have no comments and 314 articles do not have matched Compustat and CRSP data.[7,8] Since our focus is on the comments under articles, we further restrict our sample to articles with at least one comment. This process leaves us with 16,072 opinion articles, 739,057 comments, and 2,026 unique firms in our final sample. Table 1 Panel A details the sample selection process.

Panel B of Table 1 presents the industry distributions of Seeking Alpha articles, comments and the covered firms using the 12 Fama-French industry classification. The industries with the most Seeking Alpha articles are business equipment, healthcare, and wholesale. The industries with the most Seeking Alpha comments are business equipment and consumer durables. Relative

---

[6] We use "opinion article" and "analysis article" interchangeably.

[7] For a randomly selected group of 20% firms, we cross-validated the firms' tickers (i.e., trading symbols) and names on Seeking Alpha and Compustat to make sure the matched Compustat TICKERs and Seeking Alpha TICKERs refer to the same entity. We also manually matched 59 firms for which the Seeking Alpha TICKERs cannot be matched with Compustat TICKERs due to name changes, mergers / acquisitions or spinoffs, which happened after April 9, 2019. Since Compustat keeps the current TICKERs, when we access Compustat data as of March 2020, these TICKERs cannot be matched.

[8] There are 56 IPOs during 2018 in our sample. For some of these firms, the dates of the firms' first Seeking Alpha articles in our sample were before the IPO dates or the CRSP first effective dates. Therefore, these articles are removed from our final sample.

to the overall industry distribution of publicly traded common stock during our sample period, Healthcare and wholesale are relatively over-represented in Seeking Alpha articles and comments, and Finance is heavily under-represented. All other industry groups have close to proportional representation on Seeking Alpha.

General summary statistics on Seeking Alpha articles, users (including authors), and covered firms, are presented in Table 2 Panel A. Our sample includes 16,072 unique articles written by 1,508 authors and 739,057 comments posted by 45,895 unique users (including authors). These articles cover 2,026 unique firms.[9] On average, a unique firm was covered by 7.9 articles with 365 comments. The mean (median) market capitalization of covered firms is $12,954.05 ($1,646.92) million, compared to a mean (median) of $5,655.78 ($375.92) for all publicly traded firms during our sample period. While our firms are, as expected, larger than the typical publicly traded firm, our sample covers a wide range of firm sizes. While there are more articles for larger firms (correlations between number of articles and all three firm size variables are positive and statistically significant with $p<0.001$, ranging from 0.201 to 0.553), there is active commenting even for smaller firms (correlations between total number of comments and firm size are positive and statistically significant, but with smaller magnitudes, ranging from 0.058 to 0.211). Consistent with the evidence in Antweiler and Frank (2004), Nasdaq firms generate more messages than NYSE firms (423 and 310 on average, respectively), despite NYSE firms being larger than Nasdaq firms in terms of market value, total assets, and sales.

The articles on Seeking Alpha are also written by a large set of authors. The average number of articles published by a unique author is 10.7, with more than 25% of the authors only

[9] The total number of comments vary among different tests based on the requirements of model specifications. For example, some tests keep articles published within N days around earnings announcement dates, and some tests focus on comments that have replies.

publishing a single article. A small number of extremely active authors (4.7% of all authors in our sample) publish over 50 articles during the sample period.

On average, each article receives 46 comments within one week of the article publication date. Thirty-three percent of the comments are "parent" comments in which the user is commenting as a direct response to the article, while 67% of the comments are replies to parent comments, creating sub-streams of comments. These comments, occurring within one week of the article, account for 87% of all comments. Panel A of Figure 1 depicts the percentage of comments per day from day zero (article publication date) through day seven. The majority of comments are made within one day of the article's publication. Among those comments made within one week of the article publication date, 50.6% (26.5%) of comments are made on day zero (one). However, comments continue for several days. We include all comments made within the window [0, 5] in our primary analyses, however results are qualitatively similar if we examine a shorter window, such as [0, 1]. For descriptive purposes, in Figure 1 Panels B and C, we plot the distribution of Seeking Alpha comments by the day of the week and by the hour of the day. Commenting is concentrated during weekdays (i.e., Monday to Friday) from 9am to 5pm, Eastern Time. Thus the majority of comments are made during trading hours, and intraday dynamics on the message board could plausibly affect markets in real time.

In Figure 1 Panel D, we plot the distribution of Seeking Alpha articles around earnings announcement dates. The spike around day 0 (i.e., the earnings announcement date) indicates that Seeking Alpha contributors are likely to write more articles regarding their opinions about the stock(s) near the earnings announcement. We also find more articles in the weeks after the earnings announcement than in the weeks before the earnings announcement, excluding day 0.[10] Given the high number of articles after earnings announcements, the issuance of important directional information about the firm on the earnings announcement date, and the high level of investor attention around these events (Da, Engelberg, and Gao 2011; Huang, Nekrasov, and Teoh 2018), we conduct a separate analysis focusing on comments on earnings announcement window articles in Section 4.

At the unique user level, each user posts 2.2 comments per article on average. About 35% of users post more than one comment per article, which allows us to study the change of extremeness by user within the same article in our analysis. The average number of total comments posted by a unique user is 16.1 during the sample period.

3.2 Opinion Extremeness Measures

Seeking Alpha comments are not constrained in terms of length. As such, they tend to be longer than tweets posted on Twitter (e.g., Bartov, Faurel and Mohanram 2018) or StockTwits (e.g., Cookson and Niessner 2019) and similar in length to stock messages on Yahoo! Finance and Raging Bull (e.g., Antweiler and Frank 2004).[11] The content of the comments tends to vary dramatically, and the language used by commenters is often casual in nature, including incomplete sentences, typos, slang, and colloquialisms. Below are three comments for an article on Microsoft on Mar 9, 2019:

[10] Drake, Thornock, and Twedt (2017) find a similar spike in Internet coverage of a firm around the firm's earnings announcement, but find more coverage prior to the earnings announcement than following it. It may be that Seeking Alpha articles' focus on analysis leads to more post-announcement articles than is typical on other sites.

[11] Bartov, Faurel and Mohanram (2018) report that the mean (median) length of tweets in their sample is 13 (13) words while Andweiler and Frank (2004) report that the majority of message board messages are between 20 and 50 words. In contrast, the mean (median) length of comments in our Seeking Alpha comment sample is 55 (34). The difference between Seeking Alpha comments and tweets is particularly strong at the high end. Due to the 140-character comment limit on tweets during their sample period, even the 99th percentile of tweets by length have just 27 words. In comparison, the 99th percentile of comments in our sample has 333 words.

> "Wow, u discovered why low growth stocks like MSFT, BA and UNH are trding TWICE their average historicals multiples. Its all buybacks, bro!!! A house of cards it is."

> "Nice article, I appreciate it. Anyway, I would point out two things. The first one is a technical error: it is not correct to compare EV with FCF as the latter is the cash flow available to shareholders as it is calculated after interest expenses whilst the former belongs to the entire firm. Therefore, I think it is not an apple to apple multiple. Anyway, this "error" should be marginal in this case and doens't affect the outcome of your analysis. The second fact that I would like to highlight is that in the analysis of market multiples, investors should take into account not only growth, but also the sustainability and visibility of revenue/earnings/cash flow going forward. Indeed, oftentimes you see multiples de-rating because the business model and hence fundamentals are expected to weaken compared to current forecasts. To conclude, your analysis make a fair point, but in my view it reflects just a part of the MSFT equity story."

> "I will have to keep my largest investment - MSFT - because even though you think it it nearly worthless compared to its price, I think the company is one of only two rated AAA; has great management, and a very dominant position in the software and cloud businesses. As a retiree, that's better than betting on extreme multiple companies like FAANG companies."

As we discuss in Appendix B, we had research assistants hand-code a subset of comments to compare alternate machine-based commenting methods. The two primary categories of algorithms utilized for natural language processing are dictionary-based algorithms and machine-learning algorithms. While dictionary-based algorithms are appropriate for coding the sentiment and extremeness of many types of text, the huge variety in Seeking Alpha comment language and style suggests that a machine-learning approach with large training set may be advantageous.[12]

Within the sub-category of machine-learning algorithms, there is a large variety of technologies and training methods available. We utilize Amazon Web Services' (AWS) Amazon Comprehend tool. Amazon Comprehend employs sophisticated deep-learning algorithms. One of

[12] Two examples of dictionary-based approaches are those developed and used in Loughran and McDonald (2011) and Bochkay, Hales, and Chava (2020). We argue that this approach is better suited to more formal business documents, such as 10-Ks, and more professional business communication, such as press releases or conference calls, rather than the more colloquial and unstructured Seeking Alpha comments. We compare AWS Comprehend output and sentiment scores obtained using word lists from Loughran and McDonald (2011) in more detail in Appendix B.

the key advantages for our purposes is that Amazon Comprehend is trained on Amazon's vast text library, including product reviews on Amazon.[13] These product reviews have a similar colloquial nature as Seeking Alpha comments – providing positive or negative feedback about an underlying product, and often agreeing with or disagreeing with previous reviewers, and sometimes including technical information (e.g., feedback on how a product meets specifications). Unlike Seeking Alpha comments, which do not include any positive/negative, bullish/bearish indicator, Amazon reviews are associated with star ratings, improving their effectiveness as a training dataset to identify sentiment and extremeness.

Appendix B discusses the AWS coding and our validations of the method in more detail. For each comment, AWS categorizes the sentiment of the given comment by giving a likelihood ratio for the comment falling into each of four sentiment categories: positive, negative, neutral, and mixed. The sum of the four likelihood ratios equals 1 per comment.

We examine three specific dimensions of extremeness, constructing measures based on the AWS sentiment outputs. The first measure is the absolute value of the *net_sentiment* of the comment, *Absolute_Extremeness*, where *net_sentiment* is calculated as *pos* – *neg* (the difference in the likelihood ratio between being positive and negative for each comment). This captures how extreme the directional sentiment expressed in the comment is. With either attitude polarization or group polarization, we would expect *Absolute_Extremeness* to increase. The second measure, *Relative_Extremeness*, captures how much the *net_sentiment* of the given comment deviates from the *net_sentiment* of the comments which preceded it. Consider, for example, a situation where

[13] Antweiler and Frank (2004) and Cookson and Niessner (2019) use supervised machine learning models which are trained on a subset of Yahoo! Finance / StockTwits stock messages for which authors attach their opinion (either Bullish or Bearish). They then apply the machine learning model to calculate sentiment likelihood scores for the rest of their sample. However, Seeking Alpha comments contain neither a directional indicator (e.g. bullish/bearish), nor a clear indicator of "extremeness." Thus there is no clear training sample with Seeking Alpha comments as there is with StockTwits messages.

the first several comments for an article are all strongly positive. In this situation, a balanced comment with *net_sentiment* = 0 would actually be quite negative relative to the discussion up to that point. We capture this by subtracting the mean *net_sentiment* of all comments made before comment *i* for the same article from the *net_sentiment* of the given comment.[14] We then take the absolute value of this measure, so that a higher value means higher deviation from preceding sentiment. Thus, a very positive comment following a series of negative comments would have a higher value for *Relative_Extremeness* than a very positive comment following a series of similarly positive comments, even though both of the comments would have a similar value for *Absolute_Extremeness*. As you can see from these examples, *Relative_Extremeness* will decrease with group polarization, in which the entire group moves to a similarly extreme opinion, but will increase with attitude polarization, in which each individual's opinion tends to become more extreme in the pre-discussion direction. Finally, the third measure addresses the gross "strong" sentiment expressed in the comment by taking the sum of the positive and negative sentiment scores. This can be thought of as capturing the extent to which a comment expresses strong language – expressing both positive and negative views, rather than neutral-toned views. We label this variable as *Strong_Language*, and define it as *pos* + *neg*. We would generally expect this to move in the same direction as *Absolute_Extremeness,* but it also captures norms around *how* individuals express their opinions, rather than just what those opinions are. Figure 2 summarizes how each of the three measures is expected to move for each of the three predicted effects of social media interaction: attitude polarization, group polarization, or convergence.

We also examine the evolution of extremeness within parent-reply streams (comment threads). For this analysis, we calculate a slight variation on *Relative_Extremeness.* To calculate

[14] For the first comment within the article, there is no measure for *Relative_Extremeness*.

Stream_Relative_Extremeness, we consider a parent-reply stream of comments as a group when calculating the cumulative means. Thus, *Stream_Relative_Extremeness* captures how the given comment compares to those that preceded it within the given parent-reply stream.

In Table 2 Panel B, we present summary statistics for the main analysis variables. *Absolute_Extremeness* and *Strong_Language* have means of 0.370 and 0.463, respectively. Since *pos* (*neg*) is defined as the likelihood ratio of being positive (negative) for each comment using the AWS machine-learning algorithm for the sentiment analysis, *pos* and *neg* are bounded by zero and one, and the total value for *pos + neg* is also bounded by zero and one. Therefore, by construction, *Absolute_Extremeness* and *Strong_Language* range from zero to one. *Absolute_Extremeness_range* and *Strong_Language_range* capture the range of *Absolute_Extremeness* and *Strong_Language* within the comments for a given article. The summary statistics in Table 2 Panel A show that articles tend to have a wide range of extremeness even within the comment streams for a single article. The calculation of *Relative_Extremeness* requires at least two comments posted within the article and thus the sample size for this variable is slightly smaller.

4. Results

In this section, we present our main findings about the evolution of opinion extremeness over the course of Seeking Alpha comment interaction. First, we look at the change of comment extremeness within comments for the same article, both for the full sample and for articles published immediately after earnings announcements. Second, we exploit instances in which a single user makes multiple comments for a single article to examine the evolution of the extremeness of their comments as the online discussion progresses. Finally, we focus on comment sub-threads – streams of comments in which a "parent" comment generates multiple replies, and study the change of opinion extremeness within parent-reply streams.

4.1 Article Level Analyses

To assess the change of opinion extremeness over the course of discussion within the comments of a single article, we examine how comment extremeness varies with respect to when a comment occurs within the sequence of comments for the given article. We define *Log_Sequence* as the natural logarithms of the order of the comment within the article's comments. Comment timestamps are accurate to the second, allowing us to reconstruct the sequence of comments with a high level of accuracy. We estimate the following regression using all comment-level observations in the sample:

$$\begin{aligned} extremeness\ measure_{i,j,k} &\\ = \alpha + \boldsymbol{\beta}\, Log\_Sequence_{i,j,k} &+ \gamma_1 Log\_Word\_Count_{i,j,k} \\ + \gamma_2\, Parent\_Comment_{i,j,k} &+ \gamma_3\, Volatility\_5\_1_{i,j,k} \\ + \gamma_4\, CAR\_5\_1_{i,j,k} + \gamma_5\, CAR\_30\_6_{i,j,k} &+ \delta_i + \theta_j + \varepsilon_{i,j,k}\,, \end{aligned} \tag{1}$$

where i indicates the user, j indicates the article, and k indicates the comment. *Extremeness measure* is one of *Absolute_Extremeness*, *Relative_Extremeness*, and *Strong_Language*, as defined in Section 3.2. We include user (δ_i) and article (θ_j) fixed effects to control for user- and article-level factors which might affect extremeness of opinions, such as the user's normal expression, e.g. whether this is a user who tends to make comments expressing extreme opinions, and fundamental uncertainty about the stock during this period. We also control for the following comment-specific variables: the word count of the comment, the volatility of the underlying stock daily return during the [-5, -1] day window relative to the comment date, and the cumulative abnormal returns of the underlying stock during [-5, -1] and [-30, -6] windows relative to the comment date.[15] The indicator variable *Parent_Comment* takes the value of one if the comment is a parent comment, commenting directly on the article, and zero if it is a reply to another comment.

Table 3 reports the results of estimating Equation (1). The significantly negative coefficients on *Log_Sequence* indicate a *decrease* of opinion extremeness over the course of discussion within an article, which is consistent with social media interaction leading to convergence, or depolarization, a moderating of the extremeness of users' opinions. The results are consistent across all three opinion extremity measures, implying that extremeness decreases both in directional magnitude (*Absolute_Extremeness*), relative to preceding comments (*Relative_Extremeness*), and in terms of total strong language used (*Strong_Language*). The economic magnitudes of the effects are also meaningful. A one standard-deviation change in *Log_Sequence* is associated with a decrease in *Absolute_Extremeness* (*Relative_Extremeness, Strong_Language*) of 3.9% (10.0%, 4.7%) of a standard deviation.

Next, we examine whether the same relation occurs following earnings announcements. An earnings announcement provides a salient information signal which is typically directional in nature (e.g. positive or negative earnings surprise). This may give users a stronger directional ex ante opinion about the firm which could prompt polarization rather than moderation. We use the same specification but restrict to the subsample of articles published within the [+1, +5] window relative to the earnings announcement date of the firm covered by the article. We exclude articles published before the earnings announcement or on day 0 to ensure that all comments are made after the earnings announcement. This subsample includes roughly 12% of the comments in our full sample. Table 4 reports the results. We find that the coefficient on *Log_Sequence* is

[15] The comment level measures are calculated based on trading days with a 4:00pm cutoff in the Eastern Time Zone.

significantly negative, which is consistent with the full sample results that opinion extremeness decreases over the discussion within each article.[16, 17] Thus, it appears that interaction on Seeking Alpha moderates, rather than increases, extremeness. In particular, given that both absolute and relative extremeness decrease, interaction on Seeking Alpha leads to a convergence in opinions towards the middle.

4.2 Within-User Article Level Analyses

While our primary specification in Equation (1) includes article and user fixed effects, it is possible that a given user chooses to comment earlier when they have a strong opinion about an article, and takes longer to comment when their opinion is less extreme. To address this alternative explanation, we exploit instances in which a single user comments more than once for a given article. If the same individual becomes less (more) extreme in their opinion as the online interaction progresses, it suggests that it is the interaction itself which moderates (increases) the individual user's opinion extremeness. To understand whether the increased interaction on social media influences an individual user's opinion extremeness, we construct a new sequence variable called *Log_Sequence_User*, which equals the natural logarithms of the order of the comment *for a particular user* within an article's comments. For example, if the user posts three comments in the same article at different times, *Log_Sequence_User* will have values of 0, 0.69, and 1.10 for their first, second, and third comments, respectively. For the user level test, we require that each user has at least five comments within the same article. A total number of 296,918 comments satisfy

[16] In a robustness test, we require the comments to be within two days of the article publication date and the results remain similar.

[17] We also run regressions on non-earnings announcement window articles and comments (defined as articles published after five days of the most recent earnings releases, or after five days of the most recent earnings releases *and* five days before the most recent earnings releases), to ensure that earnings-announcement-window articles are not driving our overall results. Results are nearly identical to those reported in Table 3.

this condition.[18] We modify Equation 1 by substituting the new sequence variable *Log_Sequence_User*, and estimate Equation 2, below,

$$\begin{aligned} &extremeness\ measure_{i,j,k} \\ &= \alpha + \boldsymbol{\beta}\ Log\_Sequence\_User_{i,j,k} + \gamma_1 Log\_Word\_Count_{i,j,k} \\ &+ \gamma_2\ Parent\_Comment_{i,j,k} + \gamma_3\ Volatility\_5\_1_{i,j,k} \\ &+ \gamma_4\ CAR\_5\_1_{i,j,k} + \gamma_5\ CAR\_30\_6_{i,j,k} + \delta_i + \theta_j + \varepsilon_{i,j,k}\ . \end{aligned} \tag{2}$$

Table 5, Panel A, presents the results. The coefficients on our variable of interest, *Log_Sequence_User*, are significantly negative across all three models. Thus the same individual is less extreme for later comments than for earlier comments for the same set of articles. In Panel B, we further supplement Equation (2) with article-user fixed effects. Thus the results presented in Panel B are within-article-user. Once again, the coefficients on *Log_Sequence_User* are all significantly negative, with $p<0.01$. These results increase our confidence in attributing the decrease of opinion extremeness, documented in Section 4.1, to users' interaction with other participants on Seeking Alpha. Thus, the results consistently lead to rejecting H1, that Seeking Alpha interaction has no effect on users' extremeness of opinions about a given firm/article. Instead, we consistently find that interaction moderates extremeness of opinions.

4.3 Comment Parent-Reply Sub-Stream Analysis

In this section, we look at the change of opinion extremeness within another interaction circle—the parent-reply stream. Within each article, users may voluntarily group into different discussion circles depending on the sub-topics they have common interest in. This is observable as the parent-reply streams, in which a "parent" comment replies directly to the article, and "reply"

[18] We also examined robustness of these results to alternate cutoffs. A lower cutoff increases the sample size, but reduces power to detect an effect. Requiring that each user has at least three (four) comments within the same article, the sample size increases to 281,775 (233,328) and the coefficients on *Log_Sequence_User* remain significantly negative at the 1% level.

comments reply to either the parent comment or a subsequent reply comment. For example, some users with interest about the fundamentals of the underlying company might discuss these within one parent-reply stream, whereas another set of users might discuss details about trade execution in another parent-reply stream for the same underlying article. If it is truly online interaction that moderates users' opinion extremeness, then interaction within a parent-reply thread should decrease comment extremeness. Conversely, if users self-select into parent-reply streams with like-minded users, the parent-reply stream can act as an "echo chamber" of like-minded individuals, resulting in increasing extremeness.

In untabulated summary statistics, we find that on average there are 15 unique parent-reply streams for a single article. About 48.62% of parent-reply "streams," however, are just a parent, with no reply. We exclude these observations from our analysis. For parent-reply streams with at least two comments (including the parent comment), we estimate the following regression:

$$\begin{aligned} extremeness\ measure_{i,j,k} &= \alpha + \boldsymbol{\beta}\ Log\_Sequence\_Stream_{i,j,k} + \gamma_1 Log\_Word\_Count_{i,j,k} \\ &+ \gamma_2\ Volatility\_5\_1_{i,j,k} + \gamma_3\ CAR\_5\_1_{i,j,k} + \gamma_4\ CAR\_30\_6_{i,j,k} + \delta_i \\ &+ \theta_p + \varepsilon_{i,j,k}\ , \end{aligned} \tag{3}$$

where the key variable of interest is *Log_Sequence_Stream*. *Log_Sequence_Stream* is defined similarly to *Log_Sequence*, but uses the order of comments within a parent-reply stream rather than the order of comments for the article. We also modify *Relative_Extremeness.* We define *Stream_Relative_Extremeness* as the absolute value of *net* minus the mean *net* sentiment of all comments made before comment *i* within the same parent-reply stream. Regression (3) is estimated with user (δ_i) and parent (θ_p) fixed effect.

Table 6 Panel B displays the results of estimating Equation (3) using all comment observations in parent-reply streams with at least two comments. We find significantly decreasing extremeness for all three measures, *Absolute_Extremeness, Stream_Relative_Extremeness* and *Strong_Language.*[19]

Because parent-reply streams are by definition shorter (or at most the same length) as the overall comment stream for an article, a large percentage of the sample has a very small number of comments within a single stream. Over 50% of the sample falls within parent-reply streams with five or fewer comments. In untabulated analyses, we estimate the model for three subsets of parent-reply streams, with number of comments falling in the following ranges: (1,5], [6, 31], and [32, 300] comments (below 50th percentile, 50th-90th percentile, and above 90th percentile). Results are robust.

Overall, the results for parent-reply streams reinforce the main results, indicating that discussion on Seeking Alpha decreases the extremeness of users' opinions.

4.4 Cross-Sectional Variation: Source Article Extremeness and Commenter Identity

An important difference between Seeking Alpha and certain previously examined online settings, such as the political science settings discussed in Section 2.1, is that companies and stocks are less likely to be inherently polarizing and controversial. However, there are times when a given firm may be more inherently controversial. A high-profile example for a private company would be when Dick's Sporting Goods chose to stop selling firearms. This was a highly controversial decision, both politically, and in terms of how it would likely affect the company financially. As discussed in Section 2.1, we expect that convergence, the decrease of opinion extremeness, is less likely to happen when the underlying article being discussed is inherently more controversial. We

[19] Note that *Stream_Relative_extremeness* is constructed based on the mean extremeness *before* the comment *i* within a parent-reply stream, so the parent comment has a missing value for *Stream_Relative_extremeness*. Thus the sample size is smaller for *Stream_Relative_extremeness* than for *Absolute_extremeness* and *Strong_Language*. Alternatively, we define the benchmark for the parent comment as zero, neutral sentiment, and define *Stream_Relative_extremeness* as the absolute value of (*net* – 0) for the parent. Using this alternate variable, the coefficient on *Log_Sequence_Stream_*remains significantly negative at the 1% level.

proxy for this by using the extremeness of the language used in the article itself. We define *Art_Strong_Language_High* equal to 1 if the level of the extremeness of the language used in the article is in the top decile. Article level language extremeness is defined as the sum of the number of positive words plus the number of negative words divided by the total words count per article, following the Loughran and McDonald (2011) financial dictionary, as articles are more likely to be consistent with financial language analyzed in their study than plain English captured by AWS' Comprehend Sentiment Scores. To test our prediction, we re-estimate equation (1) adding the interaction of *Log_Sequence* and *Art_Strong_Language_High*. The results are presented in Table 7 Panel A. We find significantly positive coefficients on the interaction term in all three models. The significantly positive coefficients on the interaction terms are consistent with our prediction. In other words, the extremeness-decrease we document is concentrated among articles which are *not* inherently extreme. This suggests that an important difference between our setting and those studied previously is the underlying discussion topic: Investing is likely to be less inherently controversial than certain other topics discussed online.

As discussed in Section 2.1, we predict that commenters are more likely to moderate extremeness if they are identified. Online anonymity is argued to be one of the factors that increases extremeness online (Dyer and Kim, 2020). We collect data on Seeking Alpha commenters' self-reported biographical descriptions. Table 7 Panel B reports descriptive statistics. The vast majority of commenters do not have a biography – they leave this portion of their profile blank. However, authors are more likely to have biographies, and to have longer and more informative biographies, than non-author users. The mean (median) count of word for non-author users is 9 (20), and 92 (56) for authors. This difference in biographical length suggests that authors may be more subject to reputational discipline, and are thus the least likely to engage in polarizing discourse. Therefore, in the subsequent analysis, we analyze non-author user comments and author comments separately.

For user-type commenters, we first define an indicator variable, *User_Detail*, equal to 1 if the user has a non-blank biographical description, and zero otherwise.[20] We expect these individuals to be those who are most likely to moderate extremeness, both in their own comments, and in terms of the overall effect they have on the discussion. We then construct an article level measure *%User_Detail* that equals the percentage of non-author user comments made by users with *User_Detail*=1 for that article, and define *%User_Detail_High* equal to 1 if *%User_Detail* is in the top decile across articles. *%Author_Detail_High* is defined similarly but is constructed for author-type users. We predict that articles with a high percentage of non-anonymous commenters are more likely to experience a decrease in opinion extremeness over the course of discussion within an article. We re-estimate equation (1) adding an interaction term between *Log_Sequence* and (a) *%User_Detail_High*, and (b) *%Author_Detail_High*, separately.

The results are presented in Table 7, Panels C and D. Consistent with our prediction, the coefficients on the interaction terms in all three models are significantly negative, indicating a greater decrease in opinion extremeness among articles with a higher percentage of non-anonymous commenters, for both non-author users and for authors. To further test whether user- and author-type users have incremental effects on convergence of opinions, we include both sets of variables in one regression. The results are presented in Table 7 Panel E. The coefficients on both interaction terms are significant and negative in all three models. Overall, these results suggest that an important differentiator between Seeking Alpha and other forums where opinion-extremeness may increase is that a significantly amount of commenters engage in self-

[20] The results are robust when we define *User_Detail* using a minimum length biography of five, ten, or fifteen words.

identification and reputation-building. This is likely to contribute to moderation of, rather than increasing, extremeness.

4.5 Financial Market Effects: Market Differences of Opinion Following Seeking Alpha Articles

The results reported in Sections 4.1 through 4.4 show consistent evidence that online interaction moderates opinion extremeness for Seeking Alpha users, who are likely capital market participants (Chen, De, Hu and Hwang, 2014). If these users trade based upon their evolving beliefs, we should find that market-based differences of opinion decrease upon the publication of Seeking Alpha articles. To better understand if and how Seeking Alpha interaction impacts financial markets, we examine the evolution of three market-based differences of opinion (*DIVOP_Market*) measures after Seeking Alpha articles. We estimate the following model:

$$\begin{aligned} DIVOP\_Market_{i,t} &= \alpha + \beta_1 Day\_m1^e_{i,t} + \beta_2 Day\_0^e_{i,t} + \beta_3 Day\_1^e_{i,t} + \beta_4 Day\_2^e_{i,t} \\ &+ \beta_5 Day\_3^e_{i,t} + \beta_6 Day\_4^e_{i,t} + \beta_7 Day\_5^e_{i,t} + \gamma_1\, Volatility\_5\_1_{i,t} \\ &+ \gamma_2\, CAR\_5\_1_{i,t} + \gamma_3\, CAR\_30\_6_{i,t} + \gamma_4 DIVOP\_Market_{i,t-1} + \delta_i \\ &+ \theta_t + \varepsilon_{i,t}\,, \end{aligned} \tag{4}$$

where *i* indicates firm, *t* indicates trading day, and the superscript "e" represents the event of interest. *DIVOP_Market* is one of the three market-based divergence of opinion measures for stock *i* on day *t* following Cookson and Niessner (2019) and Garfinkel (2009). Specifically, the three *DIVOP_Market* measures are (1) standardized unexplained volume (*SUV*), (2) log abnormal volume (*AbLogVol*), and (3) abnormal market-adjusted turnover (*Ab_AdjTurnover*). Detailed variable definitions are in Appendix A. The main variables of interest are *Day_N* (N = 1 to 5), which is an indicator variable that equals one if day *d* is the *N*th trading day subsequent to the event date of interest.[21] The coefficients on these variables capture the difference in *DIVOP_Market* from two days before the Seeking Alpha article, to N days after the article. We use two days prior to the publication of the article as our benchmark to capture the net effect of the article and comment stream – if a post-article decrease is simply mean reversion of an article-induced spike, this model will show no net effect. If, in contrast, the comment interaction decreases differences of opinion, relative to what they would be without the article and comments, we should find negative coefficients on *Day_N*. We include stock fixed effects (δ_i) and date fixed effects (θ_t). *Day_m1* is an indicator variable that equals to one if day *d* is one day preceding the event day of interest. If the online discussion following a Seeking Alpha article publication reduces opinion extremeness, which in turn mitigates the divergence of opinions in the market, we expect $\beta_n < 0$, for n = 0 to 5.

The primary alternative explanation for negative coefficients on the *Day_N* variables is that it is not interaction in Seeking Alpha comment streams that is affecting market differences of opinion, but rather that any article such as the Seeking Alpha article will lead to a decrease in differences of opinion. Such articles provide information, analysis, and draw investor attention, all of which might have an impact on differences of opinion. To address these alternate explanations, we examine two sets of alternative event dates. First, we examine analyst earnings forecast dates. If information events similar to the articles – where an individual is sharing their opinion about the firm – drive a decrease, we should see a decrease after these forecasts. Second, we examine days with top tercile news coverage for the given firm-year, which we call high-news-count days. If news events draw investor attention which decreases *DIVOP_Market*, we should observe a

[21] Articles published after trading hour are adjusted to the next day, and analyst forecast announcement after trading hours are also adjusted to the next day.

decrease following these high-news-count days. However, none of these events is associated with online discussion in a similar manner as Seeking Alpha articles.

Table 8 describes the sample construction for each of the four samples we examine. The primary restriction we make is that for each event type, we exclude days with another event of the same type, a Seeking Alpha article, or an earnings announcement, in the [-7, +7] window surrounding the given event, to focus on non-overlapping event windows in which it is more likely that changes in *DIVOP_Market* are associated with the event of interest.[22] For Seeking Alpha articles, we include a sample which further excludes both analyst-forecast and high-news-count overlaps. Finally, we require that there are no missing *DIVOP_Market* measures for the eight-day window (i.e., [-2, +5] days around the information event day). After these procedures, we obtain a sample of 4,217 firm-article days for Seeking Alpha articles, including 1,812 unique firms. The sample sizes for the other two event types, analyst forecast days and high-news-count days, are similar.

Results of estimating Equation (4) are displayed in Table 9. Panel A shows the results with standardized unexpected volume, *SUV*, as the dependent variable. Column (1) reports results for the sample of Seeking Alpha article observations. The coefficients on indicators *Day_1* to *Day_5* are all negative and significant ($p<0.01$). In terms of the economic significance, on day t+1, *SUV* is -0.065 smaller than day t-2 *SUV*. This is a significant magnitude given that the mean (median) of *SUV* on all days in our sample period is 0.052 (-0.203) and the mean (median) of *SUV* on all day t-2 is 0.360 (-0.085). This smaller *DIVOP_Market* on day t+1 compared to day t-2 is consistent

[22] In additional analyses, we estimate Equation 4 on samples allowing for overlapping event windows. As expected, the results are more similar across the three event types, Seeking Alpha article days, analyst forecast days, and high-news-count days, when windows with two or three of the given event types are included in the corresponding samples. However, the coefficient estimates remain more negative and statistically significant for the Seeking Alpha sample than the other two.

with the notion that market divergence of opinion decreases after the release of Seeking Alpha article, compared to the pre-article level. The results also show that the decrease in divergence persists for five days after the article is published, which corresponds to the typical duration of active discussion following Seeking Alpha articles. In Column (2), we exclude both analyst-forecast days and high-news-count days from Seeking Alpha article days, and the results are similar to Column (1). Columns (3) and (4) present the results for analyst-forecast days and high-news-count days, respectively. In both columns, we do not see a similar decreasing pattern in *DIVOP_Market* after the corresponding event dates. In Panels B and C, we find similar results when using *AbLogVol* and *Ab_AdjTurnover* as dependent variables. Overall, the results are consistent with online interaction on Seeking Alpha being associated with reduced divergence of opinions in the stock market.[23]

5. Conclusion

We utilize a state-of-the-art text analysis service offered by Amazon Web Services to code not just the direction of sentiment expressed in an individual's online interaction, but their extremeness. Examining absolute and relative extremeness and extremeness of language used, we find strong and consistent results that interacting on social media *decreases* extremeness when it comes to stock-market analysis. Individuals become less extreme when commenting on Seeking Alpha analysis articles as comment discussions progress. These results hold when including article fixed effects, which control for firm-time-specific factors such as fundamental uncertainty and differences of opinion about the firm, and when including user fixed effects, which control for the

[23] In a robustness test, we partition our article days sample into days with Seeking Alpha articles but without news articles, with positive but low news count days (i.e., total news count < 8), and with high news count days (i.e., total news count >=8). The decreasing pattern of *DIVOP* is consistent across all three subsamples. Thus these results are not unique to either "isolated" Seeking Alpha articles, or to Seeking Alpha articles which occur when a firm has high news coverage overall. This suggests that decreases in *DIVOP* are not driven by other concurrent public news events.

user-specific tendencies towards extremism (e.g. whether certain types of users select into commenting at certain times in discussions). Moreover, our results hold when examining within-user variation for a given article. The *same* user becomes more moderate in their views as the online discussion progresses. Thus, our results provide strong evidence that interacting on social media moderates extremism, leading to convergence in investors' opinions.

We also examine how extremism evolves over "parent-reply" comment sub-streams. On the one hand, these sub-streams should have the strongest impact on a user, as they are the subset of comments which the user has found most relevant to themselves. On the other hand, sub-streams can allow for "echo chambers" of like-minded individuals to form, which can lead to increasing extremeness within the stream. We find that even within parent-reply comment sub-streams, comments become more moderate over the course of the stream, converging.

To better understand how social media interaction affects extremeness, we examine two factors which might drive the decreasing extremeness effect: the content of Seeking Alpha material being less inherently controversial than some other online topics, and the commenters of Seeking Alpha often being self-identified rather than anonymous. Our results support both of these conjectures. The moderation effect is lower for articles which are more extreme themselves, suggesting that Seeking Alpha is likely to have a weaker moderation effect in times when a given stock is more controversial. The moderation effect is higher if more of the individuals commenting on an article are identified, suggesting an important role for self-identification and reputation building in driving the moderating effect of Seeking Alpha comment boards. Collectively, these results provide insight specifically on Seeking Alpha's effects, but more generally on when we might expect online interaction to have more of a moderating, versus extremeness-increasing, effect on user opinions.

Finally, we examine whether the extremeness-moderation we observe may impact financial markets by examining the evolution of market differences of opinion after Seeking Alpha articles and comparable events which differ on a key dimension: they lack the Seeking Alpha comment forum. While market-based differences of opinion measures decrease during the week following a Seeking Alpha article, they do not decrease following analyst forecast days or high-news-count days. These results are consistent with the Seeking Alpha extremeness-moderation effect impacting the market. While drawing a causal link between Seeking Alpha and financial market effects of extremeness-reduction is beyond the scope of this study, our results suggest that this is an important area for future research.

Our results provide insight into investors' opinion updating in the Internet era. They suggest that research into the effects of social media interaction on stock price dynamics should consider extremeness. But our results are also more broadly relevant given the growing societal debate about the effects of social media on polarization and extremism. Rather than increasing extremism, interaction on Seeking Alpha moderates it. Given the vast role that social media in particular, and the Internet more generally, plays in our world today, there are many questions which remain for future research to explore. Our results suggest that online interaction, and the resulting dynamics, are important factors to consider.

References

Amichai-Hamburger, Y. 2007. Personality, individual differences and Internet use. *The Oxford Handbook of Internet Psychology* pp. 187-204..

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

Bail, C. A., L. P. Argyle, T. W. Brown, J. P. Bumpus, H. Chen, M. B. F. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky, 2018. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences, 115*(37), pp. 9216-9221.

Barbera, P., 2015. How social media reduces mass political polarization. Evidence from Germany, Spain, and the U.S. *Working paper, University of Southern California.*

Bartov, E., Faurel, L. and Mohanram, P.S., 2018. Can Twitter help predict firm-level earnings and stock returns?. *The Accounting Review*, *93*(3), pp.25-57.

Beaver, W.H., 1968. The information content of annual earnings announcements. *Journal of Accounting Research*, pp.67-92.

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

Boxell, L., M. Gentzkow, and J. M. Shapiro, 2017. Is the internet causing political polarization? Evidence from demographics. *NBER Working Paper No. w23258.*

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

Cookson, J.A. and Niessner, M., 2019. Why Don't We Agree? Evidence from a Social Network of Investors. *The Journal of Finance, forthcoming*.

Da, Z., Engelberg, J. and Gao, P., 2011, March. In search of Attention. *The Journal of Finance*, *66*(5), pp.1461-1499.

Diether, K.B., Malloy, C.J. and Scherbina, A., 2002. Differences of opinion and the cross section of stock returns. *The Journal of Finance*, *57*(5), pp.2113-2141.

Drake, M.S., Roulstone, D.T. and Thornock, J.R., 2012. Investor information demand: Evidence from Google searches around earnings announcements. *Journal of Accounting Research*, *50*(4), pp.1001-1040.

Dyer, Travis and Kim, Eunjee, Anonymous Equity Research (September 4, 2020). Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3686739Farrell, M., Green, T.C., Jame, R., and Markov, S., 2020. The Democratization of Investment Research and the Informativeness of Retail Investor Trading. Working paper, Available at SSRN: https://ssrn.com/abstract=3222841

Garfinkel, J.A., 2009. Measuring investors' opinion divergence. *Journal of Accounting Research*, *47*(5), pp.1317-1348.

Gomez, E., Heflin, F., Moon, J. and Warren, J., 2018. Crowdsourced Financial Analysis and Information Asymmetry at Earnings Announcements. *Georgia Tech Scheller College of Business Research Paper*, (18-45).

Huang, X., Nekrasov, A. and Teoh, S.H., 2018. Headline salience, managerial opportunism, and over-and underreactions to Earnings. *The Accounting Review*, *93*(6), pp.231-255.

Hudson, D. L., 2019. In the age of social media, expand the reach of the first amendment. *Human Rights Magazine* 43 (4).

Hutto, C.J. and Gilbert, E., 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth International AAAI Conference on Weblogs and Social Media*.

Jackson, B. F., 2014. Censorship and freedom of expression in the age of Facebook. *New Mexico Law Review* 44: 121-167.

Joinson, A., K. McKenna, T. Postmes, and U. Reips, editors., 2007. *The Oxford Handbook of Internet Psychology*. Oxford University Press.

Kandel, E. and Pearson, N.D., 1995. Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, *103*(4), pp.831-872.

Kelly, T., 2008. Disagreement, dogmatism, and belief polarization. *The Journal of Philosophy*, *105*(10), pp.611-633.

Lerman, A., 2020. Individual investors' attention to accounting information: evidence from online financial communities. *Contemporary Accounting Research. Forthcoming.*

Lord, C.G., Ross, L. and Lepper, M.R., 1979. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), p.2098.

Loughran, T. and McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), pp.35-65.

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

Moscovici, S. and Zavalloni, M., 1969. The group as a polarizer of attitudes. *Journal of Personality and Social Psychology*, *12*(2), p.125.

Myers, D.G. and Lamm, H., 1976. The group polarization phenomenon. *Psychological Bulletin*, *83*(4), p.602.

Qian, H. and Scott, C.R., 2007. Anonymity and self-disclosure on weblogs. *Journal of Computer-Mediated Communication*, *12*(4), pp.1428-1451.

Scott, C.R., 2004. Benefits and drawbacks of anonymous online communication: Legal challenges and communicative recommendations. *Free speech yearbook*, *41*(1), pp.127-141.

Shore, J., Baek, J. and Dellarocas, C., 2019. Network structure and patterns of information diversity on Twitter. *MIS Quarterly*, *42*(3), pp.947-872.

Sohangir, S., Petty, N. and Wang, D., 2018, January. Financial Sentiment Lexicon Analysis. In *2018 IEEE 12th International Conference on Semantic Computing (ICSC)* (pp. 286-289). IEEE.

Sunstein, C.R., 2002. The law of group polarization. *Journal of Political Philosophy*, *10*(2), pp.175-195.

Sunstein, C. R., 2017. *#Republic: Divided democracy in the age of social media*. Princeton University Press.

Appendix A. Variable Definitions

Variable	Definition
DEPENDENT VARIABLES	
Absolute_Extremeness	abs(*net_sentiment*). *net_sentiment* equals *pos* - *neg*, where *pos* (*neg*) is the likelihood ratio of being positive (negative) for each comment, using AWS machine-learning algorithm for sentiment analysis.
Strong_Language	*pos* + *neg*.
Relative_Extremeness	abs(*relative_sentiment*). *relative_sentiment* is calculated as *net_sentiment* - mean(*net_sentiment*), where mean(*net_sentiment*) is the mean of *net_sentiment* of all comments before comment *i* within the article.
Stream_Relative_Extremeness	abs(*stream_relative_sentiment*). *stream_relative_sentiment* is calculated as *net_sentiment* - mean(*net_sentiment*), where mean(*net_sentiment*) is the mean of *net_sentiment* of all comments before comment *i* for a particular parent-reply stream.
AbLogVol	The difference between log volume on day t and the average log volume from t-140 to t-20 trading days for stock s. *Lag_AbLogVol* is the day t-1 *AbLogVol*.
SUV	Standardized Unexplained Volume as defined in Garfinkel (2009) (p. 1326). Unexplained Volume (*UV*) is the residual from a regression of trading volume on absolute value of positive returns and the absolute value of negative returns. *UV* is standardized by the standard deviation of the residuals from the above regression, calculated over the model's estimation period, to obtain *SUV*. *Lag_SUV* is the day t-1 *SUV*.
Ab_AdjTurnover	The difference between market-adjusted turnover for stock i on day t and the average of market-adjusted turnover from t-140 to t-20 trading days, multiplied by 100. MMarket-adjusted turnover is calculated as the stock turnover for stock i on day t minus market turnover on day t. MMarket turnover is the total trading volumes on day t divided by total stocks outstanding on day t for all stocks on NYSE and AMEX. Volume is adjusted for NASDAQ stocks following Anderson and Dyl (2005), i.e., 0.62 times stock i's turnover. *Lag_Ab_AdjTurnover* is the day t-1 *Ab_AdjTurnover*.
INDEPENDENT VARIABLES OF INTEREST	
Log_Sequence	The natural logarithm of *Sequence*, where *Sequence* is the position of the comment for the given article when all comments for the article are ordered chronologically.
Log_Sequence_User	The natural logarithm of *Sequence_User*, where *Sequence_User* is the position of the comment when all comments for the given user and article are ordered. chronologically
Log_Sequence_Stream	The natural logarithm of *Sequence_Stream*, where *Sequence_Stream* is the position of the comment when all comments within the given parent-reply

	stream are ordered chronologically.
Day_m1	An indicator variable that equals to one if trading day *d* is one day preceding the event date, and zero otherwise.
Day_N (N = 0 to 5)	An indicator variable that equals to one if trading day *d* is *N*th trading day subsequent to the event date, and zero otherwise.
OTHER VARIABLES	
USER LEVEL VARIABLES	
Log_ Word_Count	The natural logarithm of (1 + word count of the comment).
Parent_Comment	An indicator variable which takes the value 1 if the comment is a parent comment.
Stream_Comments	The number of comments in a parent-reply stream, including the parent comment
Volatility_5_1	Volatility of the underlying daily stock return during [-5, -1] days of the comment date.
CAR_5_1	Cumulative abnormal return (market-adjusted) during [-5, -1] days of the comment date.
CAR_30_6	Cumulative abnormal return (market-adjusted) during [-30, -6] days of the comment date.
ARTICLE LEVEL VARIABLES	
Absolute_Extremeness_range	The difference of highest and lowest *Absolute_Extremeness* within the same article.
Strong_Language_range	The difference of highest and lowest *Strong_Language* within the same article.
Absolute_Extremeness_std	The standard deviation of *Absolute_Extremeness* among all comment within the same article.
Strong_Language_std	The standard deviation of *Strong_Language* among all comment within the same article.
Art_Strong_Language_high	Seeking Alpha article level language extremeness, defined as (# of positive words + # of negative words) / Total words count. Positive and negative words list are from Loughran and McDonald (2011), and we incorporate double negation when calculate # of positive and negative words.
%User_Detail_High	Articles with high percentage of comments made by user-type commenters with non-blank biographies. First, we define *User_Detail* = 1 if the user-type commenter has a non-blank biographical description. *%User_Detail* is constructed at article level and is the percentage of comments made by user-type commenters with *User_Detail* = 1 (if one commenter made more than one comment, her comment will be counted more than once in the numerator). We then rank *%User_Detail* into deciles and *%User_Detail_High* takes value of 1 if the article is in the top decile, and 0 otherwise.
%Author_Detail_High	Articles with high percentage of comments made by author-type commentors with non-blank biographies. The variable is defined similar to *%User_Detail_High*, but based on author-type commenters.
%Author_High	Articles with high percentage of comments made by article author. First, we calculate *%Author* as the percentage of comments made by that article's author for each article. Then we define *%Author_High* equal to 1 if the article's *%Author* is in the top decile across all articles, and 0 otherwise.

FIRM LEVEL VARIABLES	
Log_MVE	The natural logarithm of market value of equity at the most recent fiscal year end (in millions).
Log_Sales	The natural logarithm of total sales at the most recent fiscal year end (in millions).
Log_Assets	The natural logarithm of total assets at the most recent fiscal year end (in millions).

Appendix B. Amazon Web Services (AWS) Sentiment Scores

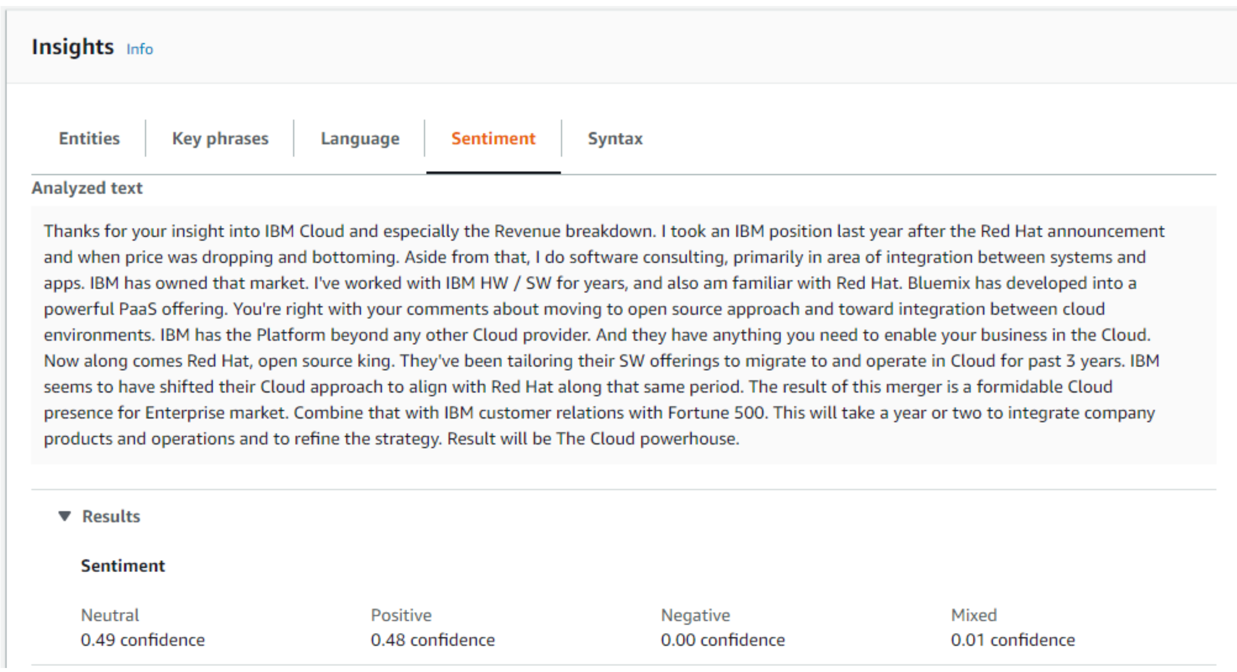
Our study is the first, to our knowledge, to use the Amazon Comprehend tool from Amazon Web Services (AWS) to analyze Seeking Alpha comments or similar financially focused social media. It is unclear ex-ante which text analysis tool is most appropriate to gauge extremeness in free-form written text such as Seeking Alpha comments, particularly given the continuous and rapid evolution in text analysis technology. We conducted several steps to evaluate alternative tools and choose the one most appropriate for our setting. In this appendix, we provide additional information about Amazon Comprehend sentiment analysis and our evaluation and validation methods. In particular, we conducted a hand-coding of comments and compared the resulting coding to both the AWS outputs and alternative text analysis options we considered. Second, we examine whether disagreement captured by our data is reflective of market disagreement. We associate disagreement measures, capturing divergence of opinion, which are calculated based on AWS sentiment scoring of Seeking Alpha comments, with market-based disagreement measures. Both methods suggest that AWS provides sentiment scores which compare well with alternatives for our purposes.

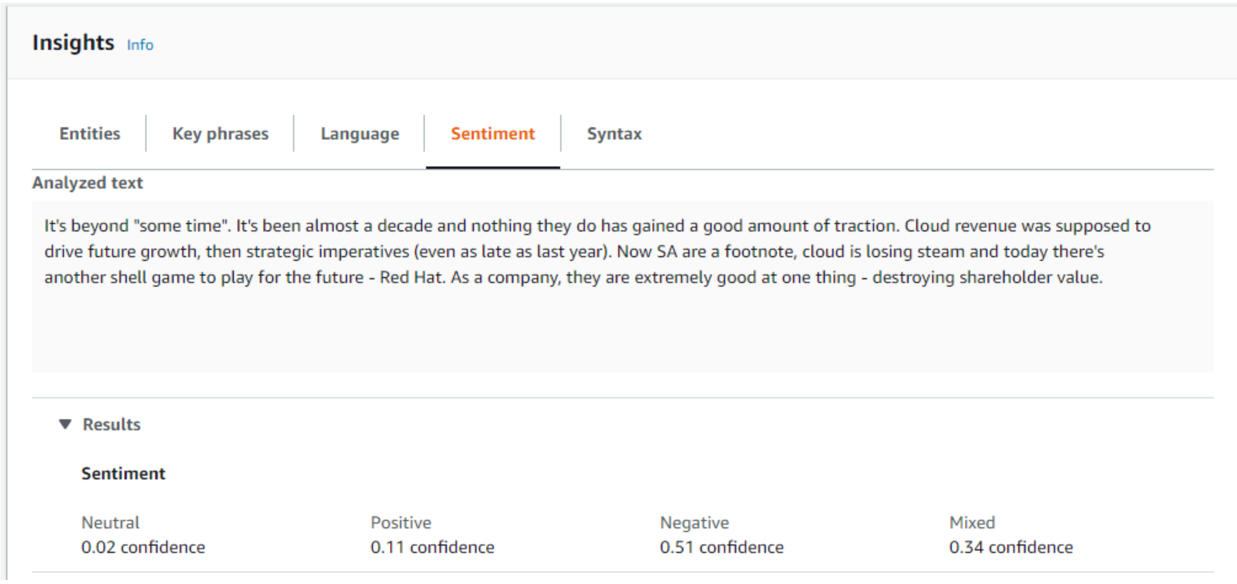
B.1 Amazon Comprehend Sentiment Analysis

Amazon's AWS Comprehend is a machine-learning-based service that uses natural language processing (NLP) to extract the emotional sentiment of a document. Unlike many other machine-learning based approaches, AWS employs deep learning algorithms, trained using Amazon's vast library of unstructured text, to extract high-level features from text. It then uses this pre-trained model to examine and analyze documents so that we do not need to provide training data.

Figure B1 shows two examples of what AWS returns from the sentiment analysis function. The figure shows two comments for an article analyzing IBM. The sentiment analysis operation returns scores in four sentiment dimensions: neutral, positive, negative and mixed. The scores represent the likelihood that the sentiment was correctly detected as being in the relevant category. For example, in the first example below it is 48 percent likely that the text has a *positive* sentiment. There is a less than 1 percent likelihood that the text has a *negative* sentiment.

Figure B1. Example of AWS Amazon Comprehend Output for Seeking Alpha Comments





B.2 Evaluation and Validation of Machine-Learning Algorithm, Method 1: Human coding

We randomly selected a list of 100 comments from our sample. Three individuals[24] were assigned to hand-code the sentiment of each comment using two methods. The first method (i.e., category coding) was to classify each comment into one of four categories: positive, negative, neutral, and mixed. Based on that, the second method (i.e., score coding) was to rate each comment for how positive or negative it is and assign one of the scores{-1,-0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1}, where more positive (negative) number indicates more positive (negative) sentiment in the comment and the score of 0 is equivalent to "neutral". Instructions were to code each comment based upon the reading of the comment itself, and not to the bullish or bearish opinion about the stock expressed in the comment.[25] For example, an extremely bullish comment would be negative if it is largely criticizing a bearish article. Alternatively, a bullish comment that focuses on positive

[24] Among the three human coders, two are graduate students and one is a senior undergraduate student. All three have accounting and finance specializations.

[25] For example, in reading of the comment "*LGCY soared 3.45% today. If an investor bought in with $1m when the market opens and sell when market close today, they will make $34.5k in one day. Easy money...and people complain that it's hard to make money in stock market. Geez*", two out of three human coders think the sentiment of this comment is positive. However, it does not mean the two human coders have positive attitude towards the stock LGCY.

forecasts about a given stock would be positive. All the tasks are performed independently across the three human coders.

All three human coders agreed on the sentiment category for 49 out of 100 comments, and for 90% of the comments at least two out of the three coders agreed. Based on self-reports, it took roughly 2.5 hours for each human coder to complete the task. Since our original sample includes more than 700,000 comments, it would be extremely time-consuming to code the sentiment using human coders. In a follow-up meeting with all human coders, they reflected that it was hard to maintain coding consistency across different comments.

Though human coding is not ideal for the large sample setting, it is valuable to compare human coding results with the sentiment measures obtained using machine-learning algorithms. This serves both as a validation test of specific machine-learning algorithms for the data in our study and allows us to compare the appropriateness of alternative automated coding methods. We compared the human score coding with two alternative machine-based methods for coding the comments: AWS and VADER.[26] AWS measures outperformed VADER measures in the sense that AWS measures provide more human-like ratings, with a correlation between the human rating and AWS (Vader) of 0.552 (0.264), with p-value < 0.001 (p-value = 0.0126). In addition to capturing the direction of sentiment more consistently to human coders, AWS provided more informative ratings about the *extremeness* of the sentiment. Vader much more often rated comments as neutral or having weak sentiment. After comparing the three methods along the

[26] VADER (*Valence Aware Dictionary and sEntiment Reasoner*), developed, evaluated and validated by Hutto and Gilbert (2014), is a lexicon and rule-based sentiment analysis tool that incorporates the impact of grammatical and syntactical rules including punctuation, capitalization and contrastive conjunction. Hutto and Gilbert (2014) find that in the setting of microblog content on social media VADER performs better than other lexicons, such as Linguistic Inquiry Word Count (LIWC), General Inquirer (GI), Affective Norms for English Words, (ANEW), SentiWordNet (SWN), SenticNet (SCN), and Word-Sense Disambiguation (WSD). Sohangir, Petty, and Wang (2018) apply VADER to StockTwits and find that VADER outperforms SentiWordNet and TextBlob in classifying Bullish and Bearish. For VADER sentiment analysis, see https://github.com/cjhutto/vaderSentiment.

dimensions of time efficiency, consistency, accuracy in capturing directional sentiment, and ability to pick up extremeness, we concluded that the AWS machine-learning-based algorithm is an effective tool for analyzing extremeness of the Seeking Alpha comments, a form of social media posting, at scale.

B.3 Evaluation and Validation of Machine-Learning Algorithm, Method 2: Association Between Comment-based Disagreement and Market Disagreement Measures

Based upon the results of Evaluation and Validation Method 1, described above, we used AWS coding for the full sample of comments. For our second evaluation and validation of the AWS coding, we took the resulting AWS coding and calculated divergence of investor opinion (*DIVOP*) measures, adapted from Antweiler and Frank (2004) and similar to those used in Cookson and Niessner (2019). We then related these measures to market-based measures of *DIVOP*. If Seeking Alpha comments are capturing investor viewpoints which are reflected in the market, and if AWS is correctly coding the comments, we should find significant positive relations between the AWS- and market-based measures.

In order to measure the divergence of opinions in Seeking Alpha, we use the following formula, following Antweiler and Frank (2004), in which *AggregateSentiment* is one of several alternate measures for sentiment aggregated across all comments for stock *s* on day *t*.

$$DIVOP_{s,t} = \sqrt{1 - AggregateSentiment_{s,t}{}^{2}}\,. \tag{B1}$$

To calculate *AggregateSentiment*, we use two alternative weighting schemes and use zero as the cutoff to identify a comment as “bullish” or “bearish.” In total, we construct three measures for divergence of opinion. Our first measure, *DIVOP_Category* is defined as follows, where a comment *k* is categorized as bullish, bearish, or hold, based upon the AWS category which has the highest value. A comment which is highest for pos (neg, neutral, or mixed) is coded as bullish (bearish, hold, and hold, respectively). For this categorical measure, *AggregateSentiment* is calculated as follows,

$$AggregateSentiment = \frac{N_{bull} - N_{bear}}{N_{bull} + N_{bear}}. \tag{B2}$$

This first method is equivalent to using equal weighting for each comment. Our second measure, *DIVOP_weight,* takes into account the magnitude, or extremeness, of the sentiment expressed in the comments. Specifically,

$$AggregateSentiment = \frac{SUM_{bull} + SUM_{bear}}{SUM_{bull} - SUM_{bear}}, \tag{B3}$$

where SUM_{bull} (SUM_{bear}) is the summation of *net* of all comments, and where a comment is categorized as bullish (bearish) if *net* > 0 (*net* < 0). Note that SUM_{bear} is a negative number, thus *AggregateSentiment* is still capturing a directional sentiment measure. This second method is equivalent to use a weighting scheme, where the weight is the magnitude/extremeness of each comment (see Antweiler and Frank 2004 for a discussion of alternative weighting schemes). For both measures, we calculate *AggregateSentiment*, for stock *s* on day *t* from comments posted between the market close of day *t*-1 to the market close of day *t*. We set *DIVOP_category* and *DIVOP_weight* to zero if none of the comments for the given stock-day are identified as bullish or bearish. Finally, our third measure, *DIVOP_stddev*, is simply the standard deviation of all comments' *net* sentiment scores for a particular stock *s* at day *t* and is set to 0 if there is only one comment.

All of these measures capture the extent of the disagreement across investors, however each measures disagreement in a slightly different way. For example, measure *DIVOP_category* does not factor in the strength of the bullish or bearish sentiment, other than in the identification of a comment as bullish/bearish, whereas the other measures factor in the extremeness of the bullish or bearish sentiment. Table B1 provides descriptive statistics for all three variables. Comparing *DIVOP_category* with *DIVOP* measure in Cookson and Niessner (2019) Table 4, the standard deviation is similar (0.455 vs 0.446), while our mean (0.380) is slightly smaller than theirs (0.467).[27] The correlations across all three variables are positive and statistically significant, as we would expect given that all three measure divergence of opinion. However, additional analyses suggest that multicollinearity is not a serious problem when all three are included together in our tests. For completeness, we measure each variable in turn as well as examining combinations.

We relate each of our three *DIVOP* measures to each of three market-based disagreement measures. The first market measure is the log abnormal volume (*AbLogVol*) used in Cookson and Niessner (2019), which is the difference between log volume of stock *s* on trading day *t* and the average log volume from *t*-140 to *t*-20 trading days for stock *s*; the next two variables are the standardized unexplained volume (*SUV*) and abnormal market-adjusted turnover (*Ab_AdjTurnover*) developed in Garfinkel (2009).[28] We use a similar specification to that used in Cookson and Niessner (2019) Table 7, in which they regress abnormal log volumes on their disagreement measures constructed from StockTwits. Specifically, we run the following regression:

$$DIVOP\_Market_{s,t} = \alpha + \beta\, DIVOP\_Comments_{s,t} + \lambda_1\, DIVOP^{m}_{s,t-1} + \lambda_2\, Controls_{s,t} + \gamma_s + \xi_t + \varepsilon_{s,t}\ , \tag{B4}$$

where $DIVOP\_Market_{s,t}$ is one of the market-based *DIVOP* measures of stock *s* on day *t* and $DIVOP\_Comments_{s,t}$ is one of the three comment-based *DIVOP* variables, *DIVOP_category*,

[27] Cookson and Niesner (2019) use categorical classification of messages on StockTwits and their average disagreement is 0.467.

[28] Garfinkel (2009) evaluates six commonly used market-based *DIVOP* measures and finds that unexplained volume measures, specifically abnormal market-adjusted turnover (*Ab_AdjTurnover*) and *SUV*, are better proxies of true divergence of opinion than all other measures examined.

DIVOP_weight, or *DIVOP_stddev*, for stock *s* on day *t*. We include stock and day fixed effects. We also control for $DIVOP\_Market_{i,t-1}$, the one-day lagged market-based *DIVOP* measure, *Volatility_5_1*, volatility of stock *s*'s return during [-5, -1] days, *CAR_5_1*, cumulative abnormal returns during [-5, -1] days, and *CAR_30_6*, cumulative abnormal returns during [-30, -6] days.

Table B2 presents the results. All three measures of disagreements are significantly and positively associated with all three market-based *DIVOP* measures. Thus we find that divergence of opinion within Seeking Alpha comments is associated with concurrent market divergence of opinion, consistent with Seeking Alpha comments capturing opinions held by market participants and with AWS providing meaningful coding of the sentiment conveyed by the comments.

While the results presented in Table B2 suggest that the AWS-sentiment-coded comments are capturing meaningful market-relevant differences of opinion, our study is focusing on *extremeness* of opinion. The *DIVOP_weight* and *DIVOP_stddev* measures factors in extremeness of opinion, while *DIVOP_category* does not. As a second analysis, we examine whether *DDIVOP_weight* and *DIVOP_stddev* have incremental explanatory power for market divergence of opinion relative to *DIVOP_category*. This is informative as to whether AWS is providing a meaningful and market-relevant measure of the "extremeness" of opinions, beyond the information conveyed in the categorical classification.

The results of these regressions are presented in Table B3. In all three models where *DIVOP_weight* and *DIVOP_category* are considered together, *DIVOP_weight* has incremental explanatory power for market divergence of opinion. The coefficients are positive and statistically significant. This is particularly striking given that the two variables are calculated in virtually identical ways, the key difference being whether extremeness is factored into the calculation. The coefficients on *DIVOP_stddev* are also statistically significant, suggesting that standard deviation also provides incremental information above the categorical variable. The two variables are slightly less comparable given that the calculation methods are different. This does suggest, however, that the way in which extremeness is examined is important. Keeping this in mind, in our paper we define three dimensions of extremeness which are potentially important, and report results for all three throughout.

Although some users who post comments on Seeking Alpha may not have any position in the underlying stocks and Seeking Alpha users only make up a small fraction of all investors in the stock market, the test results suggest that disagreement measures based on AWS sentiment scores can proxy for overall disagreement in the market, and validate the use of AWS sentiment scores as a measure of comment sentiment and the extremeness of that sentiment.

Table B1: Disagreement Measures, Summary Statistics

Panel A: Descriptive statistics

Variable	N	Mean	SD	1%	10%	25%	50%	75%	90%	99%
DIVOP_category	41,860	0.38	0.455	0	0	0	0	0.943	1	1
DIVOP_weight	41,860	0.515	0.429	0	0	0	0.68	0.94	1	1
DIVOP_stddev	41,860	0.294	0.211	0	0	0.049	0.345	0.458	0.715	0.996
Volatility_5_1	41,860	0.029	0.053	0	0.003	0.01	0.018	0.032	0.189	2.071
CAR_5_1	41,860	0	0.139	-2.785	-0.317	-0.031	0	0.028	0.321	4.543
CAR_30_6	41,860	-0.005	0.211	-5.21	-0.465	-0.069	-0.002	0.06	0.461	4.141
AbLogVol	41,300	0.184	0.692	-3.365	-1.053	-0.24	0.085	0.48	2.493	7.983
SUV	41,856	0.431	1.826	-1.955	-1.955	-0.562	-0.03	0.788	10.418	10.418
Ab_AdjTurnover	41,302	0.002	0.015	-0.03	-0.03	-0.003	0	0.003	0.096	0.096

Panel B: Pairwise correlation coefficients (Pearson (lower) & Spearman (upper)) among DIVOP measures

	DIVOP_ category	*DIVOP_ weight*	*DIVOP_ stddev*	*AbLogVol*	*SUV*	*Ab_ AdjTurnover*
DIVOP_category	1	0.790***	0.648***	0.038***	0.029***	0.050***
DIVOP_weight	0.772***	1	0.704***	0.033***	0.026***	0.046***
DIVOP_stddev	0.643***	0.728***	1	0.035***	0.027***	0.039***
AbLogVol	0.043***	0.037***	0.038***	1	0.724***	0.758***
SUV	0.047***	0.043***	0.039***	0.712***	1	0.575***
Ab_AdjTurnover	0.056***	0.048***	0.041***	0.716***	0.614***	1

This table presents summary statistics for *DIVOP* validation variables. Panel A shows the descriptive statistics of *DIVOP* variables and control variables in the regression analysis. Panel B displays the correlation coefficients among AWS-based and market-based *DIVOP* measures. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively, based on two-tailed tests. *SUV* and *Ab_AdjTurnover* are winsorized at the 1st and 99th percentiles.

Table B2: Association Among Disagreement Measures

VARIABLES	(1)	(2) *AbLogVol*	(3)	(4)	(5) *SUV*	(6)	(7)	(8) *Ab_AdjTurnover*	(9)
DIVOP_category	0.059***			0.195***			0.001***		
	(0.006)			(0.019)			(0.000)		
DIVOP_weight		0.060***			0.195***			0.001***	
		(0.006)			(0.020)			(0.000)	
DIVOP_stddev			0.107***			0.351***			0.002***
			(0.012)			(0.040)			(0.000)
Volatility_5_1	0.398**	0.395**	0.398**	4.781***	4.778***	4.786***	0.019***	0.019***	0.019***
	(0.157)	(0.156)	(0.157)	(0.592)	(0.591)	(0.594)	(0.006)	(0.006)	(0.006)
CAR_5_1	-0.033	-0.034	-0.034	-0.454**	-0.456**	-0.456**	-0.001	-0.001	-0.001
	(0.063)	(0.063)	(0.063)	(0.221)	(0.220)	(0.220)	(0.003)	(0.003)	(0.003)
CAR_30_6	0.049**	0.050**	0.050**	0.040	0.041	0.041	0.001	0.001	0.001
	(0.021)	(0.021)	(0.021)	(0.072)	(0.073)	(0.074)	(0.001)	(0.001)	(0.001)
Lag_AbLogVol	0.607***	0.607***	0.608***						
	(0.013)	(0.013)	(0.013)						
Lag_SUV				0.466***	0.467***	0.467***			
				(0.012)	(0.012)	(0.012)			
Lag_Ab_AdjTurnover							0.490***	0.491***	0.491***
							(0.019)	(0.019)	(0.019)
Constant	0.031***	0.023***	0.022***	-0.001	-0.028	-0.031	-0.000	-0.000**	-0.000**
	(0.005)	(0.005)	(0.005)	(0.018)	(0.019)	(0.020)	(0.000)	(0.000)	(0.000)
Observations	41,153	41,153	41,153	41,713	41,713	41,713	41,157	41,157	41,157
R-squared	0.635	0.635	0.635	0.464	0.464	0.464	0.465	0.464	0.464
Date FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Cluster by date and firm	YES	YES	YES	YES	YES	YES	YES	YES	YES

This table presents results of regressions of each of market-based *DIVOP* measures (*AbLogVol*, *SUV*, and *Ab_AdjTurnover*) on the *DIVOP* measures (*DIVOP _category*, *DIVOP _weight*, and *DIVOP _stddev*) and a set of controls, including the volatility of stock *s*'s returns during [-5, -1] days (*Volatility_5_1*), cumulative abnormal returns during [-5, -1] days (*CAR_5_1*), and cumulative abnormal returns during [-30, -6] days (*CAR_30_6*). Column (1) to (3) use *AbLogVol* as the dependent variable, Column (4) to (6) use *SUV* (winsorized at 1st and 99th percentiles), and Column (7) and (9) use *Ab_AdjTurnover* (winsorized at 1st and 99th percentiles) as the dependent variable. All regressions include firm and day fixed effects. Standard errors are clustered by firm and day. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively, based on two-tailed tests. Clustered standard errors are in the parentheses. The definitions of variables in the table are detailed in Appendices A and B.

Table B3: Incremental Explanatory Power of Disagreement Measures

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	*AbLogVol*		*SUV*		*Ab_AdjTurnover*	
DIVOP_category	0.038***	0.046***	0.129***	0.152***	0.001***	0.001***
	(0.008)	(0.008)	(0.025)	(0.022)	(0.000)	(0.000)
DIVOP_weight	0.030***		0.096***		0.001**	
	(0.008)		(0.026)		(0.000)	
DIVOP_stddev		0.040***		0.132***		0.001*
		(0.015)		(0.045)		(0.000)
Volatility_5_1	0.395**	0.396**	4.772***	4.776***	0.019***	0.019***
	(0.156)	(0.157)	(0.591)	(0.592)	(0.006)	(0.006)
CAR_5_1	-0.033	-0.033	-0.453**	-0.453**	-0.001	-0.001
	(0.063)	(0.063)	(0.221)	(0.221)	(0.003)	(0.003)
CAR_30_6	0.049**	0.049**	0.039	0.039	0.001	0.001
	(0.021)	(0.021)	(0.072)	(0.073)	(0.001)	(0.001)
Lag_AbLogVol	0.606***	0.606***				
	(0.013)	(0.013)				
Lag_SUV			0.466***	0.466***		
			(0.012)	(0.012)		
Lag_Ab_AdjTurnover					0.490***	0.490***
					(0.019)	(0.019)
Constant	0.024***	0.025***	-0.025	-0.023	-0.000**	-0.000**
	(0.005)	(0.005)	(0.019)	(0.020)	(0.000)	(0.000)
Observations	41,153	41,153	41,713	41,713	41,157	41,157
R-squared	0.635	0.635	0.464	0.464	0.465	0.465
Date FE	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES
Cluster by date and firm	YES	YES	YES	YES	YES	YES

This table presents results of regressions of each of market-based *DIVOP* measures (*AbLogVol*, *SUV*, and *Ab_AdjTurnover*) on the *DIVOP* measures (*DIVOP _category*, *DIVOP _weight*, and *DIVOP _stddev*) and a set of controls, including the volatility of stock *s*'s returns during [-5, -1] days (*Volatility_5_1*), cumulative abnormal returns during [-5, -1] days (*CAR_5_1*), and cumulative abnormal returns during [-30, -6] days (*CAR_30_6*). Columns (1) and (2) use *AbLogVol* as the dependent variable, Columns (3) and (4) use *SUV* (winsorized at 1st and 99th percentiles), and Columns (5) and (6) use *Ab_AdjTurnover* (winsorized at 1st and 99th percentiles) as the dependent variable. All regressions include firm and day fixed effects. Standard errors are clustered by firm and day. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively, based on two-tailed tests. Clustered standard errors are in the parentheses.

Appendix C. Examples of Content Deletion on Social Media Websites

Like most social media platforms involving user-generated content, Seeking Alpha has comment guidelines for users and explicitly states that they will delete certain types of comments which violate these guidelines. A full list of the types of comments that they would delete is outlined on their website, but include posts with obscenities, personal attacks, or promotion of the commenter's own business or website.[29] While this may seem like a unique feature of Seeking Alpha, other platforms that are widely known to moderate content include Facebook, StockTwits, and others. We provide additional detail on these examples below. Unfortunately, it is not clear how frequently any of these sites delete user-generated material, and thus it is difficult to compare Seeking Alpha's use of deletion with other social media platforms.

One social media example that particularly focuses on investors, similarly to Seeking Alpha, is StockTwits. StockTwits uses both algorithmic and human curation to detect and delete spam, abusive or offensive messages, and cheerleading/bashing messages.[30] Overall, this list appears similar to the list of content which is not allowed on Seeking Alpha. Facebook provides several ways that users can interact with each other. For Facebook's group pages, where members of a group can interact by posting messages and commenting on each other's posts. To manage the group (public or private), group administrators (admins) have the right to deny posts, remove and delete comments on a post, and block member activities.[31] Because restrictions are defined and enforced by each group individually, restrictions vary across groups. More broadly, Facebook implements community standards[32] and utilizes algorithms to censor and remove certain types of

[29] See https://seekingalpha.com/page/comment_guidelines.
[30] See https://blog.stocktwits.com/a-users-guide-for-best-practices-on-stocktwits-f0f55e2e8603.
[31] See https://www.facebook.com/help/1686671141596230/?helpref=hc_fnav.
[32] See https://www.facebook.com/communitystandards/introduction.

contents, such as objectionable content (e.g., hate speech) and violent or criminal content. Individual countries can also order Facebook to take down posts globally if the content is determined to be defamatory or otherwise illegal.[33] Other social media platforms that also apply algorithmic and/or human curation to detect and delete certain types of user-generated content include TripAdvisor[34] and Yelp.[35] There is an active discussion among legal scholars and practitioners regarding this type of censorship,[36] however it is currently legal, and, perhaps surprisingly, common.

[33] See https://www.nytimes.com/2019/10/03/technology/facebook-europe.html.

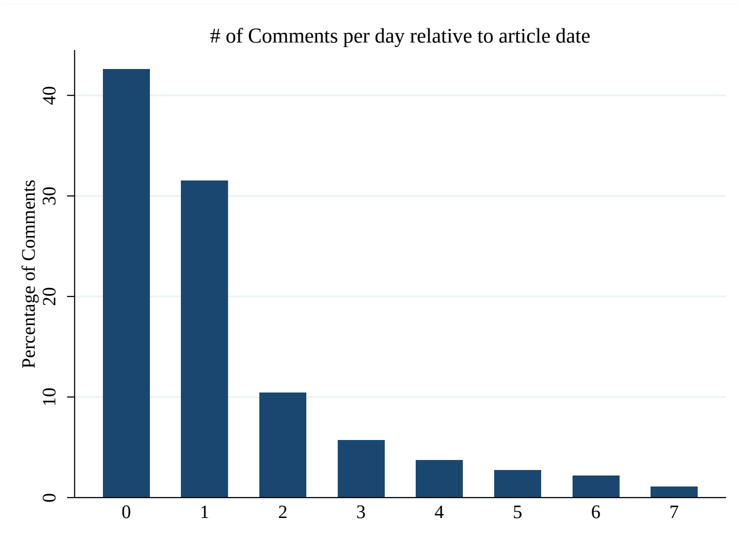
[34] "TripAdvisor uses a combination of technology and detective work to stop fake reviews reaching the site" Source: https://www.tripadvisor.com/TripAdvisorInsights/w3688

[35] "Yelp uses an automated filter to hide certain reviews in order to display only the most helpful and honest reviews. The purpose of the filter is to remove fake or illegitimate reviews. The filter is intended to protect businesses." Source: https://www.revlocal.com/blog/reviews/why-is-yelp-filtering-my-reviews-

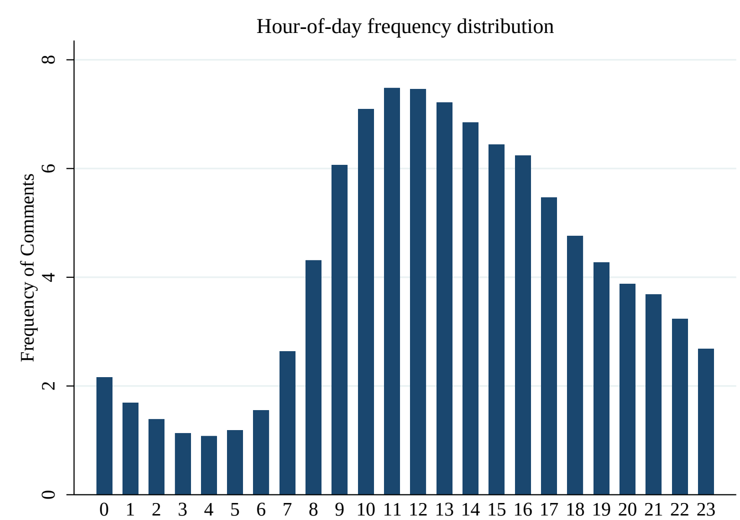
[36] As two examples of discussion regarding possibly changing the current legal protections of content moderation, see Jackson (2014), which argues "[P]ublic communications by users of social network websites deserve First Amendment protection because they simultaneously invoke three of the interests protected by the First Amendment: freedom of speech, freedom of the press, and freedom of association." Hudson (2019) writes "The First Amendment only limits governmental actors—federal, state, and local—but there are good reasons why this should be changed. Certain powerful private entities—particularly social networking sites such as Facebook, Twitter, and others—can limit, control, and censor speech as much or more than governmental entities."

Figure 1
Article and Comment Timing

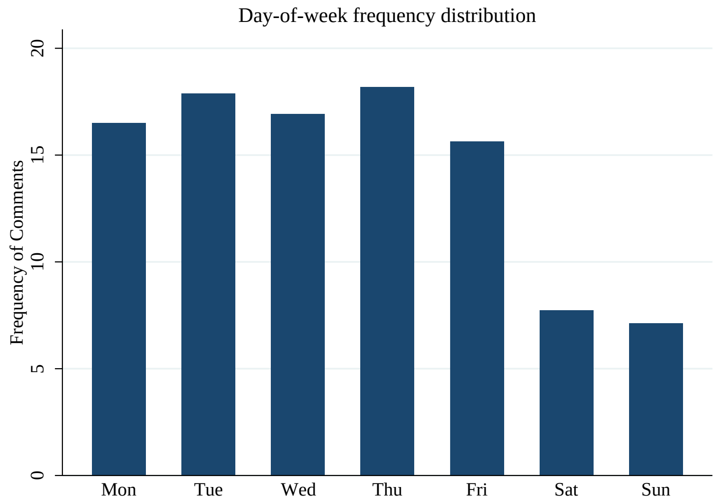
Panel A



Panel B



Panel C



Panel D

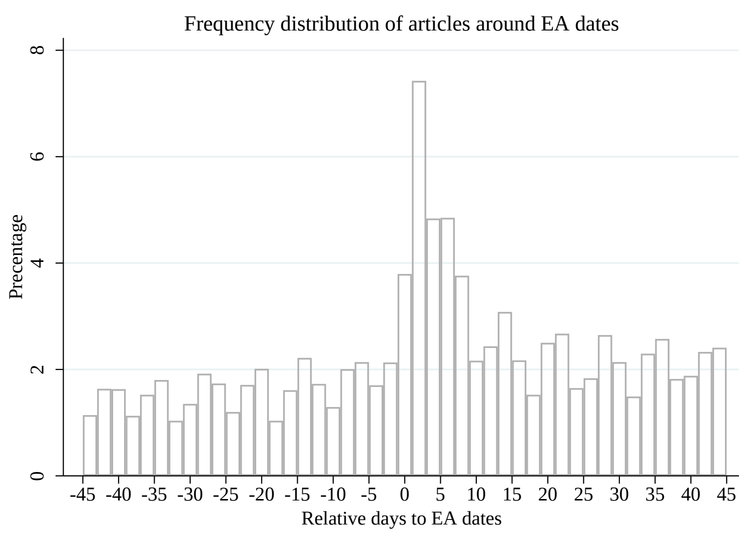


Figure 2
Predictions

	Measure		
	Absolute_Extremeness	*Relative_Extremeness*	*Strong_Language*
Increasing extremeness in opposing directions (***attitude polarization***)	+	+	+
Increasing extremeness in same direction (***group polarization***)	+	–	+
Decreasing extremeness/ Convergence (***depolarization)***	–	–	–

Table 1
Sample Description

Panel A: Sample selection

	# of SA articles	# of SA comments	# of firms
All types of articles from Seeking Alpha during April 2018 and April 2019	74,866		
Less:			
Articles with no tickers, covering multiple tickers, covering OTC stocks (including ETFs, index funds, and trusts), and articles that are transcripts of or slides from corporate earnings conferences, earnings release preview, webcasts	56,130		
Single-ticker opinion articles during our sample period	18,736	851,586	2,456
Less:			
Articles without comments or comments falling outside one week of article publications	1,534	103,179	246
Articles without corresponding Compustat and CRSP data	314	8,534	63
Articles for which there are only one comment (as we require article fixed effects)	816	816	121
Final Sample	16,072	739,057	2,026

Panel B: Sample distribution by industry

	# of SA articles	# of SA comments	# of firms		# of publicly traded common stock on the major exchanges during our sample period	
Consumer Non-Durables	878	28,769	93	4.59%	161	3.73%
Consumer Durables	999	188,441	54	2.67%	94	2.18%
Manufacturing	1,053	21,988	164	8.09%	329	7.62%
Energy	1,170	39,300	118	5.82%	159	3.68%
Chemicals and Allied Products	287	4,440	52	2.57%	92	2.13%
Business Equipment	3,825	196,889	349	17.23%	731	16.94%
Telephone and Television Transmission	591	46,590	42	2.07%	89	2.06%
Utilities	289	13,115	57	2.81%	97	2.25%
Wholesales, Retails, and Some Services	2,055	55,296	215	10.61%	324	7.51%
Healthcare, Medical Equipment, and Drugs	2,320	63,331	449	22.16%	856	19.84%
Finance	1,095	24,839	230	11.35%	827	19.17%
Other	1,510	56,059	203	10.02%	556	12.89%
Total	16,072	739,057	2,026	100.00%	4,315	100.00%

This table reports the sample selection and industry distribution. Panel A describes the criteria for articles to be included in our sample. Panel B reports the distribution of the sample by 12 Fama-French industries. The sample period spans from April 19, 2018 to April 9, 2019.

Table 2
Descriptive Statistics

Panel A: Descriptive statistics – general information

Variable	N	Mean	SD	1%	10%	25%	50%	75%	90%	99%
Article Level Information										
# of comments per article	16,072	45.98	75.92	2	3	7	18	49	117	395
Absolute_Extremeness_range	16,072	0.85	0.21	0.05	0.54	0.82	0.94	0.98	0.99	1.00
Strong_Language_range	16,072	0.82	0.22	0.06	0.50	0.77	0.91	0.96	0.98	0.99
Absolute_Extremeness_std	16,072	0.30	0.08	0.03	0.23	0.28	0.31	0.34	0.38	0.51
Strong_Language_std	16,072	0.28	0.07	0.04	0.21	0.26	0.29	0.31	0.35	0.47
Author Level Information										
# of articles per author	1,508	10.66	26.44	1	1	1	3	8	24	146
User Level Variables										
# of articles per user	45,895	7.33	22.6	1	1	1	2	5	15	97
# of comments within sample period	45,895	16.1	78.76	1	1	1	2	7	26	255
# of comments per article	45,895	2.2	3.8	1	1	1	1	2	4	17
Firm Level Variables										
# of articles per firm	2,026	7.93	24.34	1	1	1	3	6	16	78
# of comments per firm	2,026	364.79	4221.04	2	3	7	22.5	82	337	4977
Log_MVE	2,026	7.24	2.33	2.1	4.07	5.58	7.26	8.93	10.26	12.37
Log_Sales	2,026	6.66	2.59	-0.85	3.29	5.21	6.99	8.47	9.6	11.79
Log_Assets	2,026	7.28	2.35	2.31	4.12	5.62	7.33	8.88	10.41	12.58

	N	# of articles per firm	# of comments per firm	*Log_MVE*	*Log_Sales*	*Log_Assets*
NYSE	841	9.35	310.30	8.41	8.19	8.74
AMEX	61	2.38	47.11	4.79	4.32	4.89
Nasdaq	1,124	7.17	422.80	6.49	5.53	6.31

Panel B: Descriptive statistics – comment level variables

Variable	N	Mean	SD	1%	10%	25%	50%	75%	90%	99%
Absolute_Extremeness	739,057	0.370	0.309	0.002	0.022	0.086	0.291	0.625	0.859	0.984
Relative_Extremeness	722,996	0.369	0.294	0.005	0.052	0.130	0.289	0.564	0.801	1.165
Strong_Language	739,057	0.463	0.290	0.012	0.075	0.205	0.448	0.707	0.880	0.986
Stream_Relative_Extremeness	491,922	0.389	0.316	0.005	0.053	0.140	0.312	0.569	0.836	1.360
Log_Sequence	739,057	3.676	1.421	0	1.609	2.773	3.829	4.762	5.421	6.165
Log_Sequence_User	739,057	0.869	1.029	0	0	0	0.693	1.386	2.398	3.871
Log_Sequence_Stream	739,057	1.141	1.140	0	0	0	0.693	1.792	2.773	4.443
Parent_Comment	739,057	0.334	0.472	0	0	0	0	1	1	1
c_wd_count	739,057	55.471	67.009	2	8	16	34	69	125	333
Volatility_5_1	739,057	0.031	0.043	0.003	0.008	0.012	0.020	0.035	0.061	0.185
CAR_5_1	739,057	-0.008	0.147	-0.343	-0.089	-0.039	-0.002	0.029	0.078	0.271
CAR_30_6	739,057	-0.024	0.262	-0.668	-0.178	-0.090	-0.017	0.053	0.141	0.474

This table reports the distribution of analysis variables in the sample. Panel A details our sample distribution at article, author, user and firm level. The market value (*Log_MVE*), total assets (*Log_Assets*), and total sales (*Log_Sales*) are the most recently fiscal year-end numbers, which could be fiscal year of 2018 and 2019. All variables in Panel B are at the comment level. Definitions of the variables in this table are detailed in Appendix A.

Table 3
Extremeness of Opinion Dynamics

VARIABLES	(1) *Absolute_Extremeness*	(2) *Relative_Extremeness*	(3) *Strong_Language*
Log_Sequence	-0.008***	-0.022***	-0.010***
	(0.001)	(0.001)	(0.000)
Log_Word_Count	0.015***	-0.003***	0.025***
	(0.001)	(0.000)	(0.001)
Parent_Comment	0.024***	0.028***	0.020***
	(0.001)	(0.001)	(0.001)
Volatility_5_1	-0.010	0.004	-0.008
	(0.028)	(0.027)	(0.029)
CAR_5_1	0.007	0.007	0.007
	(0.010)	(0.010)	(0.010)
CAR_30_6	0.014	0.009	0.010
	(0.010)	(0.011)	(0.010)
Constant	0.340***	0.456***	0.404***
	(0.003)	(0.003)	(0.003)
Observations	721,490	704,782	721,490
Adjusted R-squared	0.073	0.068	0.101
Article FE	YES	YES	YES
User FE	YES	YES	YES
Cluster by Article	YES	YES	YES

This table presents the results of regressing opinion extremeness measures (i.e., *Absolute_Extremeness*, *Relative_Extremeness*, and *Strong_Language*) on the order of comment (i.e., *Log_Sequence*) within the article using full sample observations. The regressions are estimated with article and user fixed effects. Standard errors are clustered by article. Variable definitions are provided in Appendix A. The reported intercept term captures the average of included fixed effects. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively, based on two-tailed tests.

Table 4
Extremeness of Opinion Dynamics, Around Earnings Announcements

	(1)	(2)	(3)
VARIABLES	*Absolute_Extremeness*	*Relative_Extremeness*	*Strong_Language*
Log_Sequence	-0.008***	-0.026***	-0.010***
	(0.002)	(0.002)	(0.001)
Log_Word_Count	0.015***	-0.002	0.025***
	(0.002)	(0.001)	(0.002)
Parent_Comment	0.021***	0.028***	0.018***
	(0.003)	(0.003)	(0.003)
Volatility_5_1	-0.179*	-0.133	-0.139
	(0.104)	(0.106)	(0.102)
CAR_5_1	-0.026	-0.042	-0.018
	(0.033)	(0.033)	(0.030)
CAR_30_6	0.021	0.032	0.022
	(0.030)	(0.032)	(0.028)
Constant	0.351***	0.473***	0.421***
	(0.010)	(0.009)	(0.009)
Observations	85,777	83,299	85,777
Adjusted R-squared	0.078	0.075	0.109
Article FE	YES	YES	YES
User FE	YES	YES	YES
Cluster by Article	YES	YES	YES

This table presents the results of regressing opinion extremeness measures (i.e., *Absolute_Extremeness*, *Relative_Extremeness*, and *Strong_Language*) on the order of comment (i.e., *Log_Sequence*) within the article using comments that linked to the articles published within the [+1,+5] window surrounding earnings announcements (day 0). The regressions are estimated with article and user fixed effects. Standard errors are clustered by article. Variable definitions are provided in Appendix A. The reported intercept term captures the average of included fixed effects. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively, based on two-tailed tests.

Table 5
Extremeness of Opinion Dynamics, by User, Within Article

Panel A: User and article fixed effects

VARIABLES	(1) *Absolute_Extremeness*	(2) *Relative_Extremeness*	(3) *Strong_Language*
Log_Sequence_User	-0.007***	-0.011***	-0.006***
	(0.001)	(0.001)	(0.001)
Log_Word_Count	0.012***	-0.008***	0.021***
	(0.001)	(0.001)	(0.001)
Parent_Comment	-0.003	0.011***	-0.010***
	(0.002)	(0.002)	(0.002)
Volatility_5_1	-0.056	-0.038	-0.057
	(0.040)	(0.037)	(0.039)
CAR_5_1	0.012	0.012	0.009
	(0.014)	(0.013)	(0.013)
CAR_30_6	0.029**	0.035**	0.026*
	(0.015)	(0.014)	(0.014)
Constant	0.329***	0.392***	0.385***
	(0.004)	(0.003)	(0.004)
Observations	296,918	295,901	296,918
Adjusted R-squared	0.065	0.051	0.093
Article FE	YES	YES	YES
User FE	YES	YES	YES
Cluster by User	YES	YES	YES

Panel B: User-Article fixed effects

VARIABLES	(1) *Absolute_Extremeness*	(2) *Relative_Extremeness*	(3) *Strong_Language*
Log_Sequence_User	-0.008***	-0.013***	-0.008***
	(0.001)	(0.001)	(0.001)
Log_Word_Count	0.011***	-0.008***	0.021***
	(0.001)	(0.001)	(0.001)
Parent_Comment	-0.003	0.011***	-0.010***
	(0.002)	(0.002)	(0.002)
Volatility_5_1	-0.064	-0.054	-0.066
	(0.044)	(0.039)	(0.042)
CAR_5_1	0.022	0.019	0.023*
	(0.014)	(0.014)	(0.013)
CAR_30_6	0.043***	0.049***	0.043***
	(0.016)	(0.015)	(0.015)
Constant	0.333***	0.397***	0.389***
	(0.004)	(0.003)	(0.004)
Observations	296,918	295,901	296,918
Adjusted R-squared	0.075	0.059	0.107
User-Article FE	YES	YES	YES
Cluster by User	YES	YES	YES

This table presents the results of regressing opinion extremeness measures (i.e., *Absolute_Extremeness*, *Relative_Extremeness*, and *Strong_Language*) on the order of comments (i.e., *Log_Sequence_User*) posted by a particular user within an article. In Panel A, the regressions are estimated with article and user fixed effects. In Panel B, the regressions are estimated using user-article fixed effect. Standard errors are clustered by user in both panels. Variable definitions are provided in Appendix A. The reported intercept term captures the average of included fixed effects. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively, based on two-tailed tests.

Table 6
Extremeness of Opinion Dynamics, Within a Parent-Reply Comment Stream

Panel A: Distribution of the number of comments in a parent-reply stream

	N	**Mean**	**SD**	**1%**	**10%**	**25%**	**50%**	**75%**	**90%**	**99%**
Stream_Comments (*at the comment level*)	739,057	13.763	27.465	1	1	2	5	12	31	149
Stream_Comments (*at the unique parent level*)	247,298	2.989	5.675	1	1	1	2	3	6	22

Panel B: Regression analysis

	(1)	(2)	(3)
VARIABLES	*Absolute_Extremeness*	*Stream_Relative_Extremeness*	*Strong_Language*
Log_Sequence_Stream	-0.004***	-0.043***	-0.002***
	(0.001)	(0.001)	(0.001)
Log_Word_Count	0.022***	-0.011***	0.031***
	(0.001)	(0.001)	(0.001)
Volatility_5_1	-0.108**	-0.182***	-0.104**
	(0.050)	(0.057)	(0.048)
CAR_5_1	0.046***	0.059***	0.055***
	(0.015)	(0.018)	(0.014)
CAR_30_6	0.044**	0.040*	0.050***
	(0.018)	(0.021)	(0.016)
Constant	0.293***	0.506***	0.353***
	(0.003)	(0.004)	(0.003)
Observations	599,034	425,825	599,034
Adjusted R-squared	0.113	0.098	0.146
Parent FE	YES	YES	YES
User FE	YES	YES	YES
Cluster by Parent	YES	YES	YES

This table presents the results of regressing opinion extremeness measures (i.e., *Absolute_Extremeness*, *Stream_Relative_Extremeness*, and *Strong_Language*) on the order of comments (i.e., *Log_Sequence_Stream*) within a parent-reply stream. Panel A displays the distribution of the number of comments in a parent-reply stream. Panel B presents the results of regression model (3). Regressions are estimated with parent and user fixed effects. Standard errors are clustered by parent. Variable definitions are provided in Appendix A. The reported intercept term captures the average of included fixed effects. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively, based on two-tailed tests.

Table 7
Cross-Sectional Variation in Extremeness of Opinion Dynamics

Panel A: Article Language Extremeness

VARIABLES	(1) *Absolute_Extremeness*	(2) *Relative_Extremeness*	(3) *Strong_Language*
Log_Sequence × *Art_Strong_Language_High*	0.005***	0.005**	0.003*
	(0.002)	(0.002)	(0.001)
Log_Sequence	-0.009***	-0.023***	-0.010***
	(0.001)	(0.001)	(0.001)
Log_Word_Count	0.015***	-0.003***	0.025***
	(0.001)	(0.000)	(0.001)
Parent_Comment	0.024***	0.028***	0.020***
	(0.001)	(0.001)	(0.001)
Volatility_5_1	-0.011	0.003	-0.009
	(0.028)	(0.027)	(0.029)
CAR_5_1	0.007	0.007	0.007
	(0.010)	(0.010)	(0.010)
CAR_30_6	0.014	0.009	0.010
	(0.010)	(0.011)	(0.010)
Constant	0.340***	0.456***	0.404***
	(0.003)	(0.003)	(0.003)
Observations	721,490	704,782	721,490
Adjusted R-squared	0.073	0.068	0.101
Article FE	YES	YES	YES
User FE	YES	YES	YES
Cluster by Article	YES	YES	YES

Panel B: User and author descriptive statistics

Word count of commenters' biographical description	**N**	**Mean**	**SD**	**1%**	**10%**	**25%**	**50%**	**75%**	**90%**	**99%**
All commenters	46,223	13	59	0	0	0	0	2	36	207
All commenters with a non-empty biography	12,129	51	106	1	3	8	23	58	121	414
User-type commenters	43,713	9	47	0	0	0	0	0	20	155
Author-type commenters	2,510	92	139	2	12	29	56	109	194	596

Panel C: Self-Identified User-Type Commenters (Non-blank biography

VARIABLES	(1) *Absolute_Extremeness*	(2) *Relative_Extremeness*	(3) *Strong_Language*
Log_Sequence × %User_Detail_High	-0.005*	-0.014***	-0.006**
	(0.003)	(0.004)	(0.003)
Log_Sequence	-0.008***	-0.022***	-0.010***
	(0.001)	(0.001)	(0.000)
Log_Word_Count	0.015***	-0.003***	0.025***
	(0.001)	(0.000)	(0.001)
Parent_Comment	0.024***	0.028***	0.020***
	(0.001)	(0.001)	(0.001)
Volatility_5_1	-0.010	0.004	-0.009
	(0.028)	(0.027)	(0.029)
CAR_5_1	0.007	0.007	0.007
	(0.010)	(0.010)	(0.010)
CAR_30_6	0.013	0.009	0.009
	(0.010)	(0.011)	(0.010)
Constant	0.340***	0.456***	0.404***
	(0.003)	(0.003)	(0.003)
Observations	721,492	704,782	721,492
Adjusted R-squared	0.073	0.068	0.101
Article FE	YES	YES	YES
User FE	YES	YES	YES
Cluster by Article	YES	YES	YES

Panel D: Self-Identified Author-Type Commenters (Non-blank biography)

VARIABLES	(1) *Absolute_Extremeness*	(2) *Relative_Extremeness*	(3) *Strong_Language*
Log_Sequence × *%Author_Detail_High*	-0.014***	-0.013***	-0.011***
	(0.003)	(0.004)	(0.003)
Log_Sequence	-0.008***	-0.022***	-0.009***
	(0.001)	(0.001)	(0.000)
Log_Word_Count	0.015***	-0.003***	0.025***
	(0.001)	(0.000)	(0.001)
Parent_Comment	0.024***	0.028***	0.020***
	(0.001)	(0.001)	(0.001)
Volatility_5_1	-0.009	0.005	-0.007
	(0.027)	(0.027)	(0.029)
CAR_5_1	0.007	0.007	0.007
	(0.010)	(0.010)	(0.010)
CAR_30_6	0.014	0.009	0.009
	(0.010)	(0.011)	(0.010)
Constant	0.339***	0.456***	0.404***
	(0.003)	(0.003)	(0.003)
Observations	721,492	704,782	721,492
Adjusted R-squared	0.073	0.068	0.101
Article FE	YES	YES	YES
User FE	YES	YES	YES
Cluster by Article	YES	YES	YES

Panel E: Horserace between User-Type and Author-Type Commenters

VARIABLES	(1) *Absolute_Extremeness*	(2) *Relative_Extremeness*	(3) *Strong_Language*
Log_Sequence × *%User_Detail_High*	-0.006*	-0.014***	-0.006**
	(0.003)	(0.004)	(0.003)
Log_Sequence × *%Author_Detail_High*	-0.014***	-0.014***	-0.011***
	(0.003)	(0.004)	(0.003)
Log_Sequence	-0.008***	-0.022***	-0.009***
	(0.001)	(0.001)	(0.001)
Log_Word_Count	0.015***	-0.003***	0.025***
	(0.001)	(0.000)	(0.001)
Parent_Comment	0.024***	0.028***	0.020***
	(0.001)	(0.001)	(0.001)
Volatility_5_1	-0.009	0.005	-0.007
	(0.027)	(0.027)	(0.029)
CAR_5_1	0.007	0.007	0.007
	(0.010)	(0.010)	(0.010)
CAR_30_6	0.014	0.009	0.009
	(0.010)	(0.011)	(0.010)
Constant	0.339***	0.456***	0.404***
	(0.003)	(0.003)	(0.003)
Observations	721,492	704,782	721,492
Adjusted R-squared	0.073	0.068	0.101
Article FE	YES	YES	YES
User FE	YES	YES	YES
Cluster by Article	YES	YES	YES

This table presents the results of regressing opinion extremeness measures (i.e., *Absolute_Extremeness*, *Relative_Extremeness*, and *Strong_Language*) on the order of comments (i.e., *Log_Sequence*) and the interaction between article/commenters' characteristics and *Log_Sequence*. *Art_Strong_Language_High* indicates that the article is in the top decile of extreme language use. *%User_Detail_High* (*%Auther_Detail_High*) indicates that the percentage of comments from user-type commenters (author-type commenters) with non-blank biographies on Seeking Alpha is in the top decile of all articles. The regressions are estimated with article and user fixed effects. (Article fixed effects subsume *Art_Strong_Language_High* in Panel A, and *%User_Detail_High* and *%Auther_Detail_High* in Panels C, D, and E.) Standard errors are clustered by article. Additional variable definitions are provided in Appendix A. The reported intercept term captures the average of included fixed effects. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively, based on two-tailed tests.

Table 8 Sample Construction for Financial Market Effects Tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Main Sample, Seeking Alpha Articles with Comments		Seeking Alpha Articles with Comments, Excluding Both Analyst Forecast and High-News-Count Days		Analyst Forecast Sample		High-News-Count Sample	
	# of firm-article-days	# of firms	# of firm-article-days	# of firms	# of firm-analyst forecast-days	# of firms	# of firm-news-days	# of firms
Starting observations [37]	12,657	1,866	5,877	1,644	17,283	1,753	79,071	2,049
Remove trading days with overlapping event in the [-7, -1] and [+1, +7] window[38]	8,436	53	2356	35	11,902	73	72,977	166
	4,221	1,813	3,521	1,609	5,381	1,680	6,094	1,883
Remove if missing *DIVOP_Market* measures in the [-2, +5] trading-day window around the given event	4	1	2	0	1	1	1,085	93
	4,217	1,812	3,519	1,609	5,380	1,679	5,009	1,790
× 8 daily observations per event	× 8		× 8		× 8		× 8	
Sample size, firm-day observations	33,736		28,152		43,040		40,072	

This table describes the sample attrition for samples used in testing the implications of online discussion on market *DIVOP*. Column (1) and (2) display the sample construction for article days sample. Column (3) and (4) show the sample construction for article days. We exclude observations if there is any issuance of analyst forecast or if the news-count is high on the same day when the article published. Column (5) and (6) are the sample construction for analyst forecast events. Column (7) and (8) show the sample construction for high news-count days. For Column (5) to (8), we exclude observations if there is any Seeking Alpha article released on the same day.   A high-news-count day is defined as a day in the top tercile of daily news count for the firm.

[37] The starting samples exclude observations within the [-7, +7] trading day window around an earnings announcement, to minimize the confounding effect of earnings announcements.

[38] The "overlapping event" removed for each column is as follows: for columns (1) through (4) it is the publication of another Seeking Alpha article, for columns (5) and (6) it is the release of another analyst earnings forecast on the day when there is no Seeking Alpha article published, and for columns (7) and (8) it is another high-news day when there is no Seeking Alpha article published.

Table 9 Financial Market Effects

Panel A: Dependent variable = *SUV*

VARIABLES	(1) Seeking Alpha Articles with Comments	(2) SA Articles with Comments, excluding article-days with analyst-forecasts or High-news-count	(3) Analyst forecast days	(4) High news-count days
Day_m1	0.033	-0.001	0.010	0.008
	(0.026)	(0.025)	(0.019)	(0.017)
Day_0	0.029	-0.056**	0.185***	0.088***
	(0.025)	(0.022)	(0.019)	(0.018)
Day_1	-0.065***	-0.047**	-0.001	0.008
	(0.025)	(0.024)	(0.018)	(0.017)
Day_2	-0.109***	-0.080***	-0.023	-0.002
	(0.024)	(0.024)	(0.018)	(0.017)
Day_3	-0.098***	-0.081***	-0.023	0.005
	(0.025)	(0.025)	(0.019)	(0.017)
Day_4	-0.097***	-0.065***	-0.027	0.003
	(0.026)	(0.025)	(0.018)	(0.017)
Day_5	-0.087***	-0.057**	-0.051***	0.008
	(0.025)	(0.024)	(0.019)	(0.017)
Volatility_5_1	4.179***	5.002***	5.524***	4.747***
	(1.051)	(0.846)	(0.828)	(0.479)
CAR_5_1	-0.069	0.013	-0.555***	-0.099
	(0.235)	(0.203)	(0.177)	(0.125)
CAR_30_6	0.150**	0.121*	-0.017	0.079**
	(0.065)	(0.065)	(0.071)	(0.039)
Lag_SUV	0.339***	0.320***	0.318***	0.272***
	(0.012)	(0.012)	(0.010)	(0.010)
Observations	33,720	28,150	37,901	40,043
Adjusted R-squared	0.407	0.356	0.348	0.285
Date FE	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Cluster by firm	YES	YES	YES	YES

Panel B: Dependent variable = *AbLogVol*

VARIABLES	(1) Seeking Alpha Articles with Comments	(2) SA Articles with Comments, excluding article-days with analyst-forecasts or High-news-count	(3) Analyst forecast days	(4) High news-count days
Day_m1	0.008	-0.014	0.022**	-0.008
	(0.012)	(0.012)	(0.009)	(0.012)
Day_0	0.010	-0.027**	0.150***	0.078***
	(0.011)	(0.011)	(0.010)	(0.013)
Day_1	-0.040***	-0.033***	-0.027***	-0.013
	(0.010)	(0.011)	(0.008)	(0.011)
Day_2	-0.055***	-0.046***	-0.004	-0.012
	(0.010)	(0.011)	(0.008)	(0.011)
Day_3	-0.039***	-0.041***	-0.008	-0.006
	(0.011)	(0.012)	(0.009)	(0.011)
Day_4	-0.050***	-0.044***	-0.004	-0.013
	(0.011)	(0.011)	(0.008)	(0.011)
Day_5	-0.049***	-0.043***	-0.008	-0.013
	(0.011)	(0.012)	(0.008)	(0.011)
Volatility_5_1	0.023	0.427*	0.065	1.961***
	(0.468)	(0.248)	(0.311)	(0.349)
CAR_5_1	0.026	0.053	-0.044	0.063
	(0.098)	(0.074)	(0.079)	(0.088)
CAR_30_6	0.052	0.044	-0.059	0.133***
	(0.041)	(0.057)	(0.043)	(0.040)
Lag_AbLogVol	0.483***	0.464***	0.509***	0.445***
	(0.017)	(0.016)	(0.016)	(0.016)
Observations	33,317	27,764	37,287	39,023
Adjusted R-squared	0.590	0.541	0.529	0.413
Date FE	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Cluster by firm	YES	YES	YES	YES

Panel C: Dependent variable = *Ab_AdjTurnover*

VARIABLES	(1) Seeking Alpha Articles with Comments	(2) SA Articles with Comments, excluding article-days with analyst-forecasts or High-news-count	(3) Analyst forecast days	(4) High news-count days
Day_m1	-0.001	-0.034**	0.026***	0.005
	(0.018)	(0.015)	(0.010)	(0.007)
Day_0	-0.017	-0.076***	0.136***	0.060***
	(0.017)	(0.014)	(0.010)	(0.008)
Day_1	-0.084***	-0.062***	-0.023***	-0.015**
	(0.016)	(0.014)	(0.009)	(0.007)
Day_2	-0.096***	-0.071***	0.002	-0.003
	(0.015)	(0.014)	(0.008)	(0.006)
Day_3	-0.054***	-0.045***	-0.005	0.005
	(0.016)	(0.014)	(0.008)	(0.006)
Day_4	-0.090***	-0.068***	0.002	-0.005
	(0.017)	(0.015)	(0.008)	(0.007)
Day_5	-0.079***	-0.065***	-0.003	0.007
	(0.016)	(0.015)	(0.008)	(0.006)
Volatility_5_1	0.148	0.037	0.495**	0.820**
	(0.466)	(0.359)	(0.249)	(0.354)
CAR_5_1	0.199*	0.077	0.042	0.119*
	(0.117)	(0.087)	(0.077)	(0.067)
CAR_30_6	0.007	-0.018	-0.088*	-0.086***
	(0.066)	(0.084)	(0.050)	(0.032)
Lag_Ab_AdjTurnover	0.443***	0.478***	0.476***	0.541***
	(0.017)	(0.017)	(0.015)	(0.023)
Observations	33,327	27,774	37,287	39,092
Adjusted R-squared	0.441	0.440	0.451	0.631
Date FE	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Cluster by firm	YES	YES	YES	YES

This table presents the results of OLS estimation of equation (4) separately for several samples. Earnings announcement windows are excluded in all columns. Column (1) shows results for days around Seeking Alpha articles with comments. Column (2) shows results for the subset of Seeking Alpha articles with comments, excluding any articles for which there is an analyst forecast or high-news-count day during the event window. Column (3) shows results around analysts forecast days, and Column (4) shows results around high-news-count days. The sample construction processes are detailed in Table 8. *Day_m1* is an indicator variable that equals to one if trading day *d* is one day preceding the event date, and zero otherwise. *Day_N* (*N* = 0 to 5) is an indicator variable that equals to one if trading day *d* is the *N*th trading day subsequent to the event date, and zero otherwise. Regressions are estimated with date and firm fixed effects. Standard errors are clustered by firm. All variables are defined in Appendix A. The reported intercept term captures the average of included fixed effects. *, **, *** denote significance at the 10%, 5%, and 1% levels, respectively, based on two-tailed tests.