Journal of Financial Economics 145 (2022) 426–444

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

# Listening in on investors' thoughts and conversations☆

Hailiang Chen[a], Byoung-Hyoun Hwang[b,*]

[a] Faculty of Business and Economics, The University of Hong Kong, Hong Kong
[b] Nanyang Business School, NTU, Singapore

ARTICLE INFO

*Article history:*
Received 9 July 2020
Revised 21 June 2021
Accepted 15 July 2021
Available online 10 September 2021

*JEL classification:*
G11
G12
G14
G40

*Keywords:*
Social interactions
Social transmission bias
Asset prices

ABSTRACT

A large literature in neuroscience and social psychology shows that humans are wired to be meticulous about how they are perceived by others. In this paper, we propose that impression management considerations can also end up guiding the content that investors transmit via word of mouth and inadvertently lead to the propagation of noise. We analyze server log data from one of the largest investment-related websites in the United States. Consistent with our proposition, we find that investors more frequently share articles that are more suitable for impression management despite such articles less accurately predicting returns. Additional analyses suggest that high levels of sharing can lead to overpricing.



## 1. Introduction

People frequently draw on each other for financial advice and investment ideas (e.g., Shiller and Pound, 1989; Hong et al., 2004; Brown et al., 2008; Kaustia and Knüpfer, 2012; Lerner and Malmendier, 2013; Shue, 2013; Hvide and Östberg, 2015; Pool et al., 2015; Heimer, 2016; Fracassi, 2017; Hwang et al., 2019). Information transmitted through social interactions can lead to more informed financial decisions (Maturana and Nickerson, 2019; Haliassos et al., 2020). However, information transmitted via word of mouth can also mislead as propagators distort information or share only select, nonrepresentative types of information (Hirshleifer, 2015, 2020). The goal of this study is to propose and test for the relevance of one important feature that can cause social interactions to mislead. We also consider the implications of this feature for the quality of investors' investment decisions and for asset prices.

One of the most established theories in social psychology contends that, when considering what content to share in their social interactions, people primarily contemplate what impressions their sharing could create among receivers and whether those impressions are consistent with who they are or desire to be. People share to convey the impression that they are likeable; people also share to profess connoisseurship or project a certain image (e.g., Goffman, 1978; Baumeister, 1982; Leary and Kowalski, 1990; Berger and Schwartz, 2011; Berger and Milk-

☆ David Hirshleifer was the editor for this article. We are grateful to the editor and an anonymous reviewer for their extremely helpful comments and suggestions. We also would like to thank Diego Garcia, Bing Han, Soo Kim, Russell Jame, Wei Jiang, Justin Murfin, Manoj Thomas, and the seminar participants at Baruch College, the China International Conference in Finance (CICF) 2021, Cornell University, Korea University, Nanyang Technological University, Singapore, Singapore Management University, Stockholm University, and the University of Hong Kong for their many valuable suggestions and comments. Hailiang Chen acknowledges financial support from the Faculty of Business and Economics at the University of Hong Kong.

* Corresponding author.
*E-mail addresses:* chen19@hku.hk (H. Chen), bh.hwang@ntu.edu.sg (B.-H. Hwang).

https://doi.org/10.1016/j.jfineco.2021.09.004

man, 2012; Lovett et al., 2013; Packard and Wooten, 2013; Berger, 2014; Gilovich et al., 2019). In the literature, this concept is referred to as "self-presentation" or "impression management." Its roots can be traced to human evolution. As humans started to live in communities, to survive and thrive, they needed to socially affiliate with one another and create perceptions of power and competence (Lakin and Chartrand, 2003; Lakin et al., 2003; Baek et al., 2020).

In this paper, we propose that impression management considerations significantly determine investor conversations. That is, when deciding what content to share, investors do not always prioritize the value relevance of the content per se. Instead, they frequently pay greater attention to what impressions they make by sharing a particular content and whether such impressions agree with their actual or desired self.

To test this proposition, we examine investors' consumption and sharing of quantitative versus qualitative content. Opinions couched in numbers are often viewed as more thoughtful (Kadous et al., 2005; Koetsenruijter, 2011; McConway, 2016; Huang et al., 2018). They can also convey the idea that the source of the opinion is more knowledgeable (Roeh and Feldman, 1984; van Dijk 1988). In addition, a heavy reliance on numbers can create impressions of intelligence and competence (Weiss and Feldman, 2006; Cragg and Gilmore, 2014).

If impression management considerations are important to investors and if intelligence and thoughtfulness are desirable attributes to investors, then we should observe that investors predominantly share quantitative content over qualitative content even if investors, themselves, were to more frequently read content that is less reliant on numbers.

We test this hypothesis in two empirical settings. In our first setting, we use aggregated server log data from Seeking Alpha (SA; http://seekingalpha.com). SA is one of the largest investment-related websites in the United States. It hosts close to one million stock opinion articles and attracts more than 15 million unique visitors per month. For each article published on the SA website from August 2012 to March 2013, we have data on the number of times a reader scrolled to the bottom of the article ("number of read-to-ends") and the number of times the article was shared via e-mail ("number of shares").

The results from our first setting show that, while SA users substantially more frequently read articles that contain relatively few numbers, it is the articles that are heavily couched in numbers that generate the greater number of shares and, thus, become a predominant part of investors' conversations.

In our second setting, we conduct experiments on 840 actual investors, a significant number of whom have net investable assets worth more than $300,000. We select a relatively quantitative SA article and a relatively qualitative SA article, and we present both articles to all participants. For randomly selected participants, we upward manipulate their desire to manage their impressions and subsequently track whether these participants more frequently share the quantitative article compared with the qualitative article.

The results from our second setting show that, when impression management considerations are strengthened, investors substantially more frequently choose to share the quantitative article, but not the qualitative article.

When gauging which articles are more accurate, we find that it is the qualitative SA articles that more accurately predict returns. Conversely, quantitative SA articles less accurately predict returns.

In general, we find that the more read-to-ends a SA article receives, the more accurately the corresponding article predicts returns. This finding is consistent with prior work, suggesting that users of SA-type platforms are informed and can identify accurate from inaccurate articles (e.g., Chen et al., 2014; Avery et al., 2016; Jame et al., 2016). In stark contrast, the SA articles that receive the more frequent number of shares less accurately predict returns.

We conclude our paper with an examination of possible asset pricing implications. Our previous results suggest that impression management considerations are important to investors and can lead to the dissemination of less informative content. If receivers do not sufficiently account for this possibility and if investors more frequently buy a stock than short sell (Barber and Odean, 2008), then stocks that are mentioned more frequently in investors' conversations can become overpriced and, subsequently, earn abnormally low returns.

Consistent with this line of thinking, we find that viral stocks (i.e., stocks mentioned in SA articles that receive a disproportionately high number of shares) initially experience high returns. These high initial returns subsequently reverse. The initial price run-up and subsequent correction is substantially stronger for stocks that are likely to be short-sale constrained. We observe a similar pattern when considering a six-year sample containing all tweets regarding publicly traded firms in the United States. Our argument is further corroborated by survey evidence: Investors report taking investment ideas transmitted via word of mouth seriously and frequently act upon it. Many of the stocks that investors end up buying subsequently underperform.

Our study makes the following contribution to the literature. Given evidence that investors derive much of their information through social interactions (e.g., Shiller and Pound, 1989; Hong et al., 2004; Brown et al., 2008; Kaustia and Knüpfer, 2012; Hvide and Östberg, 2015; Pool et al., 2015; Heimer, 2016), there is good reason to determine whether information acquired from social interactions leads to better or worse investment decisions and what factors cause information acquired from social interactions to mislead.

Owing to the richness of our data, we are able to simultaneously observe both the types of content investors consume and the types of content investors subsequently choose to share with their peers. We use our data to show that the stories that investors most frequently consume can be strikingly different from those that they most frequently share. We find evidence that the content that investors consume is more accurate than the content that investors share. We point to one important motive that, at least partially, explains this finding. In their social interactions, investors are meticulous about the image that they project. Stories best suited for impression management are not always the most value relevant. Likewise, stories that are the most value relevant are not always well suited for impression management. This creates a wedge between content consumed and content shared with the latter generally being less informative. If listeners do not sufficiently account for this possibility, they end up putting too much weight on certain stories. This, in turn, can trigger excess buying activity and generate both temporary mispricing and poor investment performances.[1]

The rest of the paper is organized as follows. Section 2 describes the literature on impression management and presents our proposition that impression management considerations significantly determine investors' conversations. Sections 3 and 4 present the corresponding evidence utilizing server-log- and experimental data, respectively. Section 5 examines whether impression-management considerations lead to the transmission of less informative content. Section 6 examines asset-pricing implications. Section 7 concludes.

## 2. Literature review and hypothesis development

We begin the main body of our paper with an outline of the literature on impression management. We then build on this literature to develop our hypothesis.

### *2.1. Literature on impression management*

One of the key theories in social psychology contends that, when people interact with one another and contemplate about the content to share, they primarily consider how the message could be perceived by others. Gilovich et al. (2019) describe "a basic truth": "Our social self is often a dramatic performance in which we try to project a public self consistent with our hopes and aspirations. This public self is one that we actively create in our social interactions and that is shaped by the perceptions of other people and the perceptions we want others to have of us."

Similarly, Smith and Mackie (2000, p. 134) summarize that "*ingratiation*, trying to convey the impression that we are likeable, and *self-promotion*, trying to convey an impression of competence, are two of the most common goals of social interaction."

Impression management can be traced to human evolution (Baek et al., 2020). As nature nudged humans to live in large communities, humans became wired to socially affiliate with one another and create impressions of power and competence (Lakin and Chartrand, 2003; Lakin et al., 2003). Being liked and being seen as competent increase the odds that a human would be included when material goods are split up between members of the group. Being liked and being seen as competent also increase the odds that a human can pass on his or her genes to the next generation. In fact, evidence in anthropology and neuroscience suggests that the single most important reason that humans developed the largest brain size–to–body size ratio among all species is so that humans can form and manage large communities and survive and thrive within them (Dunbar, 1992, 1998; Hill and Dunbar, 1998; Lieberman, 2013).

### *2.2. Hypothesis development*

We propose that impression management considerations also significantly determine investors' conversations. This, in turn, can cause investors to inadvertently propagate noise with wide-ranging implications for the quality of investors' investment decisions and asset prices.

To test whether impression management considerations significantly determine investors' conversations, we consider investors' consumption and sharing of quantitative versus qualitative content. Figs. 1 and 2 present word clouds for the three topics that most frequently appear in quantitative SA articles and the three topics that most frequently appear in qualitative SA articles.[2] Quantitative (qualitative) articles are defined as articles for which the ratio of the total occurrences of numbers in an SA article to the total number of words is in the top (bottom) quartile of its distribution.

Figs. 1 and 2 show that the SA articles most heavily reliant on numbers generally discuss a company's earnings announcement and financial statement and use terms such as "quarter," "sales," "billion," "million," "average," and "analysts." SA articles least reliant on numbers generally emphasize consumer and investor sentiment and include terms such as "people," "like," "technology," "products," "shares," and "investors."[3]

We conjecture that, to investors, quantitative content is more useful to manage their impressions than qualitative content. Opinions couched in numbers are often viewed as more thoughtful (Kadous et al., 2005; Koetsenruijter, 2011; McConway, 2016). Quantification also increases concreteness and credibility (Huang et al., 2018). Research in journalism suggests that numbers are used primarily for their rhetorical effect, that is, to create an impression of credibility and authority, not for their information value (Roeh and Feldman, 1984; van Dijk, 1988).

In addition to conveying reliability, thoughtfulness, and care, quantitative content could also be used to signal in-

[1] Our implications are related to those of prior literature, which suggests that, while investors are eager to share their investment successes, they are not forthcoming about their investment failures (Kaustia and Knüpfer, 2012; Heimer and Simon, 2017; Escobar and Pedraza, 2019; Lane et al., 2020). If receivers do not account for such communication decisions, they could erroneously infer that outperforming the market is easy and flock to active investment strategies (Han et al., 2021).

[2] In our topic modeling, we apply Latent Dirichlet Allocation to all the SA articles in our final sample. We find that 70 topics produce the best model fit. We then examine which three topics most frequently appear in quantitative articles (relative to qualitative articles) and which three topics most frequently appear in qualitative articles (relative to quantitative articles), and we plot the top 50 words from each of these topics.

[3] The articles deemed as "qualitative" in Fig. 2 are still rooted in numbers. For instance, an SA article that is at the 25th percentile in terms of its ratio of the total occurrences of numbers to the total number of words would be at the 59th percentile among the articles published in the *Wall Street Journal* in 2020, whereas an article that is at the 75th percentile among SA articles would be at the 98th percentile among the *Wall Street Journal* articles.

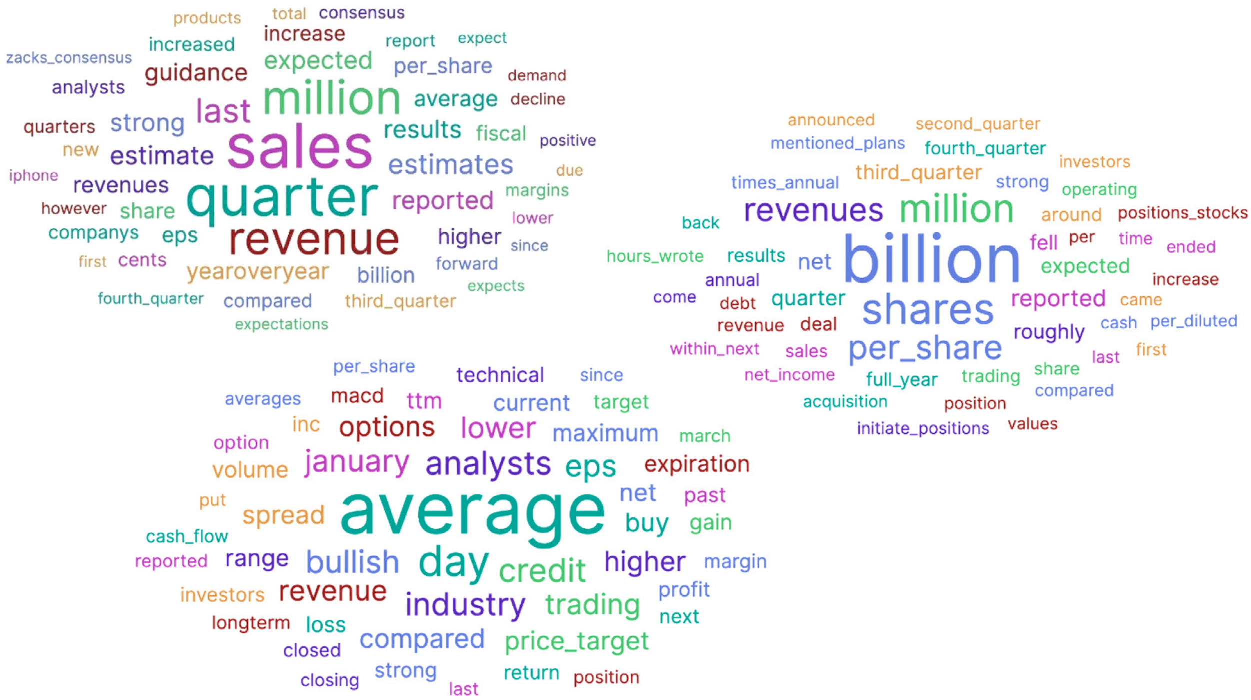


**Fig. 1.** Topics disproportionately appearing in quantitative Seeking Alpha articles.


**Fig. 2.** Topics disproportionately appearing in qualitative Seeking Alpha articles.

telligence and brilliance (Weiss and Feldman, 2006). Evidence in neuroscience shows that such signaling efforts are not unsubstantiated. The ability to process numbers is positively associated with the ability to reason, solve problems, and self-regulate (Cragg and Gilmore, 2014). The latter abilities are part of what one may call "intelligence." Related work shows that poor numeracy has a stronger negative impact on labor market outcomes and on mental and physical health than poor literacy (Parsons and Bynner, 2005).

If investors are not immune to impression management considerations and if, to investors, quantitative content is more suitable for impression management than qualitative content, then investors should more frequently share content that is couched in numbers even if, "in private," investors more frequently read content that is less reliant on numbers.

We do not believe that impression management manifests itself in the sharing of quantitative content only. We view the sharing of quantitative content merely as a testing ground to empirically assess the broader notion that impression management considerations are an important determinant of investors' conversations.

## 3. Impression management and investors' sharing behavior: evidence from server log data

We first test our hypothesis utilizing server log data from the SA website.

### *3.1. Server log data*

SA is a leading investments-related website in the United States. Anyone can submit a stock opinion article for possible publication on the SA website. These submissions are curated by a team of SA editors. If deemed of adequate quality and published on the SA website, the authors of the articles receive income based on the number of page views their articles generate.[4] As of January 2021, the website had published close to one million articles authored by 17,247 contributors.[5] From January through March 2019, the SA website attracted more than 15 million unique visitors a month, who, on average, spent seven minutes per visit.[6] Over the same period, SA reports that its audience had an average household income of $321,302 and that 65% of its audience traded at least once a month.[5]

Our server log data cover the period from August 1, 2012 to March 31, 2013. Our data are at the article level. We do not have data at the user level. For each article identification (ID), which is an identifier set by SA to uniquely identify SA articles, we have data on how often readers viewed an article, how often they scrolled to the bottom of the article, and how often the article was shared via e-mail. We obtain the server log data directly from SA. We augment our server log data with article-level data, which we scraped from the SA website. For each article, we scraped the article ID, title, full article text, date of publication, author's name, and stock ticker. Our final sample contains 16,446 single-stock opinion articles.

### *3.2. Variables and descriptive statistics*

#### *3.2.1. Dependent variables*

We use our server log data to construct two dependent variables. Our first dependent variable is the natural logarithm of one plus the number of times an article is read to the end, ln *(1 + # Read-to-Ends)*. We take the natural logarithm as *# Read-to-Ends* is highly right-skewed. We view our first dependent variable as a proxy for article consumption, albeit an upward-biased one, as we cannot rule out that some investors scroll to the bottom of the page without reading the article.

Our second dependent variable is the natural logarithm of one plus the number of times an article is shared via e-mail, ln *(1 + # Shares)*. We again take the natural logarithm as *# Shares* is highly right-skewed. To share an article, an investor is required to input the e-mail address of the recipient(s). The type of sharing we consider in this study is thus distinct from the mass sharing among loosely connected individuals, which we typically associate with social media outlets, such as Twitter or Facebook.[7]

#### *3.2.2. Independent variables*

Our main variable of interest is *Reliance on Numbers*, which is the ratio of the total occurrences of numbers in an SA article to the total number of words. In our regression analyses, we include the following controls (all of which are described in Table 1): *Length, Editors' Pick, Presentation, Long Score, Short Score,* ln *(1 + Analyst Coverage),* and ln *(1 + DJNS Coverage)*.[8]

#### *3.2.3. Descriptive statistics*

We present the descriptive statistics in Table 1. [9] Perhaps most interesting is that, on average, each read-to-end is accompanied with (only) 0.003 shares (= 5.46 / 2029.77). While this ratio could appear low, it is hardly unusual. For instance, Twitter reports that, for every tweet that is read by a user, there are only 0.0007 retweets.[10] Our ratio of 0.003 suggests that only few read-to-ends originate from individuals who receive e-mails about the articles. Instead, almost all read-to-ends originate from investors who encounter these articles through their own individual searches.

[4] The compensation scheme has evolved since 2013. Details on the current compensation scheme can be found at https://seekingalpha.com/pages/article_payments (accessed January 6, 2021).

[5] See https://seekingalpha.com/listing/contributors_stats and https://seekingalpha.com/listing/articles_stats (accessed January 6, 2021).

[6] See https://static.seekingalpha.com/uploads/2019/7/22/sa_media_kit_07_2019_generic.pdf (accessed January 6, 2021).

[7] We also have data on how many times an article was viewed by investors. In later analyses, we use the number of page views as a control. We do not consider the number of pages views to test our main hypothesis as, prior to viewing an article, investors observe only the article title and, in our sample, article titles rarely contain numbers.

[8] We thank Andrew Chen and Tom Zimmermann for generously sharing their anomaly data with us, which we use to construct *Long Score* and *Short Score*.

[9] We report a correlation matrix of our variables in Online Appendix Table OA1.

[10] See https://blog.twitter.com/en_us/a/2014/new-tweet-activity-dashboard-offers-richer-analytics.html (accessed January 6, 2021).

**Table 1**
Descriptive statistics.
In this table, we present summary statistics for our main variables at the article level. The sample contains 16,446 opinion articles written on a single stock published on Seeking Alpha (SA) from August 2012 through March 2013. *# Page Views* is the number of page views of an article. *# Read-to-Ends* is the number of times SA users scroll to the bottom of an article. *# Shares* is the number of times an article is shared via e-mail. *Reliance on Numbers* is the ratio of the total occurrences of numbers to the total number of words in an article. *Length* is the number of words used in an article. *Editors' Pick* equals one if an article is selected as an Editors' Pick and zero otherwise. *Presentation* is constructed as follows: Our server log data contain scores that SA editors assigned internally to each article based on how "actionable," "convincing," or "well presented" they perceived the article. Scores range from 1 through 5, with 1 being the lowest and 5 being the highest. *Presentation* is the sum of the above three scores. We provide more details about the *Presentation* variable in Online Appendix Fig. OA3. *Long Score* (*Short Score*) counts for how many out of 172 anomalies a stock resides in the long (short) leg. For the full list of 172 anomalies and a description of each anomaly variable, see Chen and Zimmermann (2020). *Analyst Coverage* is the number of analysts covering the stock in question. *DJNS Coverage* is the number of Dow Jones News Service (DJNS) articles about a given stock over the previous month.

| Variable | *N* | Mean | Standard deviation | 25th percentile | 50th percentile | 75th percentile |
|---|---|---|---|---|---|---|
| Article-level variables | | | | | | |
| *# Page Views* | 16,446 | 4512.94 | 4581.09 | 1765.00 | 3284.50 | 5732.00 |
| *# Read-to-Ends* | 16,446 | 2029.77 | 1982.82 | 758.00 | 1467.00 | 2638.00 |
| *# Shares* | 16,446 | 5.46 | 8.51 | 1.00 | 3.00 | 7.00 |
| *Reliance on Numbers* | 16,446 | 0.04 | 0.04 | 0.02 | 0.03 | 0.05 |
| *Length* | 16,446 | 1000.38 | 621.13 | 620.00 | 866.00 | 1187.00 |
| *Editors' Pick* | 16,446 | 0.08 | 0.28 | 0.00 | 0.00 | 0.00 |
| *Presentation* | 16,446 | 9.55 | 0.79 | 9.00 | 9.00 | 10.00 |
| Stock-level variables | | | | | | |
| *Long Score* | 16,446 | 10.43 | 5.52 | 6.00 | 10.00 | 14.00 |
| *Short Score* | 16,446 | 14.96 | 8.34 | 9.00 | 13.00 | 19.00 |
| *Analyst Coverage* | 16,446 | 5.01 | 9.44 | 0.00 | 0.00 | 6.00 |
| *DJNS Coverage* | 16,446 | 39.89 | 46.65 | 8.00 | 24.00 | 57.00 |

### 3.3. Main analyses

To assess how *Reliance on Numbers* relates to the number of read-to-ends, we estimate the following regression specification at the article level:

$$ln(1 + \# Read - to - Ends_i) = \alpha_k + \beta \ Relianceon \ Numbers_i + X_{i,j}\delta \ + \varepsilon_i, \tag{1}$$

where $j$ is the stock discussed in article $i$, $X_{i,j}$ includes our article- and firm-level controls, and $\alpha_k$ are author fixed effects. We include author fixed effects as authors who have a tendency to compose quantitative (or qualitative) articles can differ along other author characteristics. These other, unobserved author characteristics could in turn generate our findings. We estimate regression Eq. (1) initially including ln *(1 + # Page Views$_i$)* as an additional control and then without including ln *(1 + # Page Views$_i$)* as an additional control. The former shows the features that make it more likely that an article is read to the end, on the condition that the article is being viewed. The latter sheds light on the features that lead to more frequent read-to-ends, unconditionally. Unless further specified, *t*-statistics in this and the ensuing regression specifications are based on standard errors adjusted for heteroskedasticity and clustered by day of article publication.

We report our results with regards to regression Eq. (1) in Columns 1 and 2 of Table 2. Table 2 shows that quantitative articles are substantially less frequently read to the end. The estimate in Column 1 shows that, on the condition that an article is being viewed, a one standard deviation increase in *Reliance on Numbers* leads to 5.1% less frequent read-to-ends (*t*-statistic = −5.92). The estimate in Column 2 shows that, unconditionally, a one standard deviation increase in *Reliance on Numbers* leads to 5.0% less frequent read-to-ends (*t*-statistic = −3.28).

The estimates for our control variables largely agree with expectations. Investors are less likely to finish reading an article if the text is long. Unconditionally, investors more frequently read Editors' Pick articles as well as articles on stocks that are receiving extensive media coverage. Investors less frequently read articles on stocks that already have existing investment recommendations in the form of analyst reports.

The above results shed light on how investors behave by themselves. Next, we consider how investors behave and interact with one another. We re-estimate regression Eq. (1) but replace the dependent variable with ln *(1 + # Shares$_i$)*:

$$ln \ (1 + \# Shares_i) = \alpha_k + \ \beta \ RelianceonNumbers_i + X_{i,j}\delta \ + \ \varepsilon_i. \tag{2}$$

As with regression Eq. (1), we estimate regression Eq. (2) when including and excluding ln *(1 + # Read-to-Ends$_i$)* as an additional control.

We report our results with regards to regression Eq. (2) in Columns 3 and 4 of Table 2. While our previous results show that quantitative articles receive fewer read-to-ends, investors appear keen to share these quantitative investment ideas with their peers. This tendency is so pronounced that quantitative investment ideas become a predominant part in investors' conversations despite the fact that quantitative investment ideas are less frequently consumed in the first place. The estimate for *Reliance on Numbers* reported in Column 3 shows that, on the condition that an article is being read to the end, a one standard deviation increase in *Reliance on Numbers* leads to 6.7% more

**Table 2**
Consumption and sharing of stock opinion articles.
In this table, we present coefficient estimates from regressions of measures of article consumption and article sharing levels on the characteristics of the corresponding article's content. The sample contains 16,446 opinion articles written on a single stock published on Seeking Alpha from August 2012 through March 2013. The dependent variable in Columns 1 and 2 is the natural logarithm of one plus the number of times SA users scroll to the bottom of an article. The dependent variable in Columns 3 and 4 is the natural logarithm of one plus the number of times an article is shared via e-mail. All variables are defined in Table 1. We include author fixed effects. *t*-statistics are reported in parentheses and are based on standard errors adjusted for heteroskedasticity and clustered by day of article publication. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

| | What makes an article more likely to be ... | | | |
|---|---|---|---|---|
| | Read-to-end? | | Shared? | |
| Variable | (1) | (2) | (3) | (4) |
| *Reliance on Numbers* | −1.267*** | −1.256*** | 1.685*** | 0.968*** |
| | (−5.92) | (−3.28) | (5.38) | (4.63) |
| *Length* | −0.208*** | 0.030 | 0.387*** | 0.404*** |
| | (−27.83) | (1.55) | (21.12) | (19.73) |
| *Editors' Pick* | −0.115*** | 0.215*** | 0.286*** | 0.409*** |
| | (−10.31) | (6.69) | (11.48) | (13.07) |
| *Presentation* | 0.003 | 0.046*** | 0.017* | 0.043*** |
| | (0.87) | (4.37) | (1.83) | (3.86) |
| *Long Score* | 0.001*** | 0.010*** | −0.010*** | −0.004*** |
| | (3.71) | (7.96) | (−8.38) | (−2.73) |
| *Short Score* | 0.001*** | 0.018*** | 0.001 | 0.011*** |
| | (3.90) | (21.37) | (1.58) | (12.43) |
| *ln (1 + Analyst Coverage)* | 0.005* | −0.028*** | −0.027*** | −0.043*** |
| | (1.81) | (−3.70) | (−5.19) | (−7.04) |
| *ln (1 + DJNS Coverage)* | −0.005*** | 0.041*** | 0.006 | 0.029*** |
| | (−3.54) | (8.74) | (1.34) | (6.19) |
| *ln (1+ # Page Views)* | 1.005*** | | | |
| | (272.84) | | | |
| *ln (1 + # Read-to-Ends)* | | | 0.571*** | |
| | | | (63.73) | |
| Number of observations | 16,446 | 16,446 | 16,446 | 16,446 |
| Adj. $R^2$ | 0.945 | 0.418 | 0.545 | 0.350 |

frequent article shares (*t*-statistic = 5.38). Meanwhile, the estimate for *Reliance on Numbers* reported in Column 4 shows that a one standard deviation increase in *Reliance on Numbers* leads to 3.9% more frequent article shares (*t*-statistic = 4.63), unconditionally.

Overall, the results presented in Table 2 are consistent with our hypothesis and suggest that impression management considerations are an important determinant of investors' conversations. Our results also show that the content that receives the more frequent number of read-to-ends and the content that receives the more frequent number of shares can be strikingly different.

### 3.4. Additional analyses

We conduct a host of sensitivity analyses. We begin by considering the binary relation between *Reliance on Numbers* and the level of article consumption and then the level of article sharing. As reported in Online Appendix Table OA2.1, the results are similar to those presented in Table 2.

As *Long Score* and *Short Score* include just about any firm characteristic, we do not separately control for firm characteristics in our regression equations. In Online Appendix Table OA2.2, we report estimates obtained from regressions, which exclude *Long Score* and *Short Score*, and, instead, include firm characteristics as separate control variables. The firm characteristics we consider are past one-month stock returns, previous month's stock return volatility, natural logarithm of previous month's turnover, and natural logarithm of market capitalization as of the previous month. We continue to make similar observations.

Online Appendix Table OA2.3 reports results without author fixed effects and shows that the results are similar to those reported in Table 2.

Finally, we consider the possibility that articles with a greater reliance on numbers differ along other textual attributes and that these other attributes generate our results. We consider the article tone as well as the extent to which such tone differs from the average tone of articles on the corresponding stock over the previous month. The results presented in Online Appendix Table OA2.4 show that the inclusion of these textual attributes has almost no impact on our main results.

In Online Appendix Table OA2.5, we consider how more frequent sharing of quantitative articles moderates with an article's tone. One alternative interpretation of our results is that investors with a long (short) position in a stock have a financial incentive to share positive (negative) articles to encourage other investors to also buy (short) this stock. Quantitative articles could be regarded as more persuasive. This in turn could explain why investors more frequently share articles with a greater reliance on numbers.

We do not have data on whether sharers of an article have a financial stake in the stock discussed in the article. However, SA users should be significantly more likely to hold a long position than a short position. Consequently, if financial incentives were the primary determinant of our sharing result, then our pattern should be concentrated among articles that have a positive tone. As shown in Online Appendix Table OA2.5, the propensity to share quantitative content over qualitative content is of similar magnitude, regardless of whether the tone of the underlying article is positive or negative.[11]

## 4. Impression management and investors' sharing behavior: experimental evidence

While our regression-based results are consistent with impression management considerations entering investors' conversations, to provide causal evidence, we also conduct four experiments.

### *4.1. Data*

Our first set of experiments, Experiment 1 and Experiment 2, is based on 540 actual investors. Among these, we recruit 230 through Prolific (https://www.prolific.co) and the remaining 310 investors through CoreData Research (https://coredataresearch.com). Our second set of experiments, Experiment A1 and Experiment A2, is based on three hundred actual investors, all of whom we recruit through Prolific.

Prolific is a platform that allows researchers to recruit prescreened participants for online experiments. We require that participants recruited through Prolific be US residents, list English as their first language, and answer "Yes" to the following two questions: (1) "Have you ever made investments (either personally or through your employment) in the common stock or shares of a company?" and (2) "Have you invested in any of the following types of investment in the past?—Stock market."

Given that, as of the time of this study, Prolific's minimum wage was $6.50 per hour, we suspect that most of the investors we recruit through Prolific have low to medium levels of net investable assets.[12] In an attempt to include investors with high levels of net investable assets and to make the wealth distribution of our participant pool resemble that of the US retail investor population, we augment our Prolific participant pool with participants recruited through CoreData Research.[13]

CoreData Research is a market research firm that conducts investor surveys for large financial institutions. All the participants recruited through CoreData Research reside in the US, are fluent in English, and are above the age of 22. The company provides each participant's age, gender, and net investable assets. The average age of the participants in our CoreData Research sample is 57. Sixty-six percent of the participants are male. As tabulated in Online Appendix Table OA3.1, among the 310 investors we recruit through CoreData Research, 46% possess net investable assets worth more than $100,000, 32% possess net investable assets worth more than $300,000, and 12% possess net investable assets worth more than $500,000.[14]

Going forward, we refer to investors with net investable assets worth equal to or less than $300,000 as "mass-market investors" and investors with net investable assets worth more than $300,000 as "affluent investors." In this study, we assume that all investors recruited through Prolific are mass-market investors.

### *4.2. Main analyses*

Experiment 1 has 280 investors, whom we randomly assign to two groups of 140 investors each. Each group contains 110 mass-market investors and 30 affluent investors.[15]

We ask all investors to read a quantitative and a qualitative article. Both articles recommend that investors buy shares of Tesla. The two articles are published within one week of another. The key difference is that the quantitative article's *Reliance on Numbers* is above the median and the qualitative article's *Reliance on Numbers* is below the median. We report both articles in Online Appendix Fig. OA1.

At the end of each article, investors in the first group (treatment group) are asked: "Is this an article you would share with a coworker?" Those in the second group (control group) are asked: "Is this an article you would share with a very close friend?" Prior work suggests that people have stronger desires to manage their impressions while interacting with their coworkers than with their very close friends (Dubois et al., 2016). If impression management considerations are important to investors and lead to the sharing of quantitative content, then investors in the treatment group should substantially more frequently share the quantitative article than investors in the control group. In contrast, little difference should exist in the level of sharing for the qualitative article.

Experiment 2 has a different set of 260 investors, whom we randomly assign to two groups of 130 investors each. Each group, in turn, contains 110 mass-market investors and 20 affluent investors.[16]

[11] We also consider how more frequent sharing of quantitative articles moderates with the corresponding stock's past performance and an article's tone. The results presented in Online Appendix Table OA2.6 show little interaction effects. We do find that positive articles predicting a reversal from recent negative returns are substantially more frequently shared than positive articles predicting a continuation of positive returns. Conversely, negative articles predicting a reversal from recent positive returns are less frequently shared than negative articles predicting a continuation of negative returns.

[12] In our experiments, we paid participants the equivalent of $16 per hour.

[13] In its 2016 Survey of Consumer Finances, the Federal Reserve Board reports that 31% of US households hold investable assets of greater than $100,000, 18% hold investable assets of greater than $250,000, and 11% hold investable assets of greater than $500,000.

[14] Given cost considerations, we use CoreData Research only to help recruit participants for our first set of experiments, which produce our baseline results. To recruit participants for our second set of experiments, we rely on Prolific. Our agreement with CoreData Research disallows us from disclosing information regarding the amount paid to CoreData Research and the amount paid to the participants recruited through CoreData Research.

[15] Each group has 55 investors recruited through Prolific and 85 investors recruited through CoreData Research.

[16] Each group has 60 investors recruited through Prolific and 70 investors recruited through CoreData Research.

The investors in the first group (treatment group) are given the following task: "Please think about a situation where you felt you did not look as knowledgeable in the eyes of your coworkers as you would have liked. Briefly describe the situation in the box below" [minimum of 2 sentences]. Those in the second group (control group) are given the following task: "Please think about what a 'perfect' office would look like to you. Briefly describe this 'perfect' office in the box below" [minimum of 2 sentences]. Thereafter, as part of a seemingly unrelated second task, all investors read the same quantitative article and the same qualitative article as in Experiment 1. At the end of each article, investors are asked whether they would share this article with a coworker.

A large body of literature finds that people engage in more impression management when they perceive deficiencies in their self (Wicklund and Gollwitzer, 1981, 1982; Rucker and Galinsky, 2008; Gao et al., 2009; Mead et al., 2011; Packard and Wooten, 2013). If impression management considerations are important to investors and lead to the sharing of quantitative content, then, compared with investors in the control group, investors in the treatment group should substantially more frequently choose to share the quantitative article. No such difference should emerge in the level of sharing for the qualitative article.

We report our findings for Experiment 1 in Panel A of Table 3. Panel A shows that, among the investors who were asked whether they would share the quantitative article with a coworker, 60.7% report that they would do so. Among the investors who were asked whether they would share the quantitative article with a very close friend, only 45.7% report that they would do so. The difference, $\Delta_{quant}$, is 15.0% ($t$-statistic = 2.53). We observe virtually no difference in the level of sharing for the qualitative article ($\Delta_{qual}$ = 0%, $t$-statistic = 0.00). The difference between $\Delta_{quant}$ and $\Delta_{qual}$, $\Delta\Delta$, which measures the disproportionate increase in the level of sharing of quantitative content versus nonquantitative content when impression management considerations are strengthened, is 15% ($t$-statistic = 2.12).

Panel D of Table 3 reports our findings for Experiment 2. Column 1 shows that, among the investors who were tasked to write about a perceived deficiency in the self and, subsequently, were asked whether they would share the quantitative article with a coworker, 50% report that they woulddo so. The corresponding fraction among the investors who were tasked to write about their perfect office is a mere 39.2%. The difference, $\Delta_{quant}$, is 10.8% ($t$-statistic = 1.75). The difference in the level of sharing is substantially smaller for the qualitative article ($\Delta_{qual}$ = 1.5%, $t$-statistic = 0.25). The difference in differences, $\Delta\Delta$, is 9.3% ($t$-statistic = 1.46).[17]

[17] Online Appendix Table OA3.2 shows that, for both Experiments 1 and 2, the difference in the level of sharing of the quantitative article between investors in the treatment condition and investors in the control condition is roughly twice as strong for affluent investors than for mass-market investors. Our finding is consistent with prior literature suggesting that impression management considerations strengthen with wealth (Rucker, Hu, and Galinsky, 2014).

Overall, the results from our experiments are consistent with our hypothesis and provide causal evidence that impression management considerations are important to investors.

### 4.3. Additional analyses

In additional analyses, we gauge the robustness and generalizability of our results. More important, we examine to what degree financial incentives could explain our findings. We perform two experiments. Experiment A1 mimics Experiment 1 and has 150 investors. Experiment A2 mimics Experiment 2 and has a different set of 150 investors. To gauge the robustness and generalizability of our previous results, in Experiments A1 and A2, we replace the quantitative and the qualitative articles on Tesla with a quantitative and a qualitative article on Delta Air Lines. Both articles recommend that investors buy shares of Delta Air Lines. The articles are published within one week of each other. The key difference is that, while the quantitative article's *Reliance on Numbers* is above the median, the qualitative article's *Reliance on Numbers* is below the median. We report both articles in Online Appendix Fig. OA2. Unlike in Experiments 1 and 2, in Experiments A1 and A2, we now also ask participants whether they own shares of Delta Air Lines. We also include questions regarding participants' gender and investment experience.

We present our findings for Experiment A1 in Table 3. We report the results separately for the eight investors who declare that they hold shares of Delta Air Lines (Panel B) and the 142 investors who declare that they do not have a financial stake in Delta Air Lines (Panel C). We find that when impression management considerations are strengthened, investors substantially more frequently share the quantitative article even when they have no financial stake in the underlying company.

Our findings for Experiment A2 are also presented in Table 3. We report the results separately for the 12 investors who declare that they hold shares of Delta Air Lines (Panel E) and the 138 investors who declare that they do not hold shares of Delta Air Lines (Panel F). Similar to our previous experiment, we find that when impression management considerations are strengthened, investors substantially more frequently share the quantitative article even when they have no financial stake in the underlying company.

Overall, the results confirm that, when impression management considerations are strengthened, investors substantially more frequently share quantitative content over qualitative content with their peers. In fact, the results in Experiments A1 and A2 are noticeably stronger, both statistically and economically, than those in Experiments 1 and 2.

## 5. Do impression management considerations lead to the transmission of less informative content?

Having found evidence that investors are not immune to impression management considerations, we now turn to the question of whether impression management consider-

**Table 3**
Effect of impression management considerations on the sharing of stock opinion articles.
In this table, we present responses from four online experiments. (1) In Experiment 1, tabulated in Panel A, we randomly assign 280 investors to two groups of 140 investors each. Each group, in turn, has 110 investors with reported or estimated net investable assets worth equal to or less than $300,000 and 30 investors with reported net investable assets worth more than $300,000. All investors are asked to read the same quantitative article and the same qualitative article about Tesla (presented in Online Appendix Fig. OA1). At the end of each article, investors in the first group (treatment group) are asked whether they would share this article with a coworker. Investors in the second group (control group) are asked whether they would share this article with a very close friend. The first (second) column reports the fraction of treatment group investors who would share the quantitative (qualitative) article, the fraction of control group investors who would share the quantitative (qualitative) article, and the difference in the two fractions. The third column reports the difference between the differenced fraction of investors choosing to share the quantitative article and the differenced fraction of investors choosing to share the qualitative article. (2) In Experiment A1, tabulated in Panels B and C, we perform Experiment 1 on a different set of 150 investors. We now consider a quantitative and a qualitative article about Delta Air Lines (presented in Online Appendix Fig. OA2). We report our results separately for the eight investors who report to hold shares of Delta Air Lines (Panel B) and for the 142 investors who report not to hold shares of Delta Air Lines (Panel C). (3) In Experiment 2, tabulated in Panel D, we randomly assign a different set of 260 investors to two groups with 130 investors each. Each group, in turn, has 110 investors with reported or estimated net investable assets worth equal to or less than $300,000 and 20 investors with reported net investable assets worth more than $300,000. The investors in the first group (treatment group) are given the following task: "Please think about a situation where you felt you did not look as knowledgeable in the eyes of your coworkers as you would have liked. Briefly describe the situation in the box below" [minimum of 2 sentences]. The investors in the second group (control group) are given the following task: "Please think about what a 'perfect' office would look like to you. Briefly describe this 'perfect' office in the box below" [minimum of 2 sentences]. All investors are subsequently asked to read the same two articles about Tesla as in Experiment 1. At the end of each article, investors are asked whether they would share this article with a coworker. (4) In Experiment A2, tabulated in Panels E and F, we perform Experiment 2 on a different set of 150 investors. We consider the same two articles about Delta Air Lines as in Experiment A1. We report our results separately for the 12 investors who report to hold shares of Delta Air Lines (Panel E) and for the 138 investors who report not to hold any shares of Delta Air Lines (Panel F). *t*-statistics are reported in parentheses and are based on standard errors adjusted for heteroskedasticity. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

| | Fraction of investors responding that they would share the … | | |
|---|---|---|---|
| | quantitative article … (1) | qualitative article … (2) | ΔΔ |
| *Panel A: Experiment 1 (N = 280)* | | | |
| with a coworker (Treatment). | 60.7% | 55.7% | |
| with a very close friend (Control). | 45.7% | 55.7% | |
| Treatment – Control | 15.0%** | 0.0% | 15.0%** |
| | (2.53) | (0.00) | (2.12) |
| *Panel B: Experiment A1, investors with financial incentives (N = 8)* | | | |
| with a coworker (Treatment). | 100.00% | 50.0% | |
| with a very close friend (Control). | 50.0% | 50.0% | |
| Treatment – Control | 50.0%* | 0.0% | 50.00% |
| | (1.73) | (0.00) | (1.00) |
| *Panel C: Experiment A1, investors without financial incentives (N = 142)* | | | |
| with a coworker (Treatment). | 62.7% | 44.8% | |
| with a very close friend (Control). | 21.3% | 32.0% | |
| Treatment – Control | 41.4%*** | 12.8% | 28.6%*** |
| | (5.42) | (1.56) | (3.62) |
| *Panel D: Experiment 2* | | | |
| with a coworker after writing about a perceived deficiency in the self (Treatment). | 50.0% | 49.2% | |
| with a coworker after writing about a neutral topic (Control). | 39.2% | 47.7% | |
| Treatment – Control | 10.8%* | 1.5% | 9.3% |
| | (1.75) | (0.25) | (1.46) |
| *Panel E: Experiment A2, investors with financial incentives (N = 12)* | | | |
| with a coworker after writing about a perceived deficiency in the self (Treatment). | 100.00% | 40.0% | |
| with a coworker after writing about a neutral topic (Control). | 28.6% | 28.6% | |
| Treatment – Control | 71.4%*** | 11.4% | 60.0%* |
| | (3.82) | (0.38) | (1.84) |
| *Panel F: Experiment A2, investors without financial incentives (N = 138)* | | | |
| with a coworker after writing about a perceived deficiency in the self (Treatment). | 61.7% | 38.3% | |
| with a coworker after writing about a neutral topic (Control). | 38.5% | 37.2% | |
| Treatment – Control | 23.2%*** | 1.2% | 22.1%** |
| | (2.76) | (0.14) | (2.35) |

ations lead to the transmission of more or less informative content.

We construct three measures of article accuracy, *Article Accuracy*. Our first measure equals one if either an article's tone is above the median and the corresponding stock earns positive abnormal returns over the ensuing one month, or, an article's tone is below the median and the corresponding stock earns negative abnormal returns; the measure equals zero otherwise. We construct two analog measures of article accuracy based on abnormal returns over the ensuing three and six months, respectively.

We measure an article's tone as the number of positive words in the article minus the number of negative words, scaled by the total number of words. We account for negation. As has become the norm for textual analysis in the finance field, we use the lists of positive and negative words as in Loughran and McDonald (2011). Following Chen et al. (2014), when computing cumulative abnormal returns, we skip the first two days of article publication and compute abnormal returns as the difference between raw returns and returns on a value-weighted portfolio of firms with similar size, book-to-market ratio, and past return (Daniel et al., 1997), hereafter referred to as DGTW-adjusted returns.

We estimate the following regression specification at the article level:

$$Article\,Accuracy_i = \alpha + \beta\,Reliance\,on\,Numbers_i + X_{i,j}\delta + \varepsilon_i, \tag{3}$$

where $j$ is the stock discussed in article $i$ and $X_{i,j}$ includes our article- and firm-level controls.

If SA authors' comparative advantage lies in the evaluation of numbers, such as a company's earnings and financial statements, then we should observe a positive coefficient estimate for *Reliance on Numbers*. Meanwhile, if SA authors' comparative advantage lies in the evaluation of softer aspects of a firm (e.g., consumer sentiment), then we should observe a negative coefficient estimate for *Reliance on Numbers*.

We report our results in Table 4. As shown in Columns 1, 4, and 7, quantitative articles are comparatively less accurate. Depending on the return horizon over which accuracy is measured, the estimates for *Reliance on Numbers* are −0.215 ($t$-statistic = −1.92), −0.312 ($t$-statistic = −3.17), and −0.314 ($t$-statistic = −3.31). That is, an article characteristic that leads to the more frequent sharing of an article (yet to less frequent read-to-ends) strongly negatively associates with the corresponding article's investment value.

The remaining columns in Table 4 report the results when considering the levels of reading and sharing more broadly and replacing *Reliance on Numbers* with either ln *(1 + # Read-to-Ends)* or ln *(1 + # Shares)*. We find that the articles that SA users more frequently read to the end are substantially more accurate than those that SA users less frequently finish reading. Depending on the return horizon over which an article's accuracy is computed, the estimates for ln *(1 + # Read-to-Ends)* are 0.014 ($t$-statistic = 2.96), 0.014 ($t$-statistic = 2.71), and 0.010 ($t$-statistic = 2.13). Our results are consistent with prior studies, suggesting that actual and possible users of SA-type platforms are informed and capable of separating the accurate SA articles from the inaccurate ones (Kaniel et al., 2008, 2012; Kelley and Tetlock, 2013, 2017; Chen et al., 2014; Avery et al., 2016; Barrot et al., 2016; Jame et al., 2016; Boehmer et al., 2021).

In sharp contrast, we find that the articles that investors more frequently share are less accurate on average. Put differently, while the estimates for ln *(1 + # Read-to-Ends)* are strongly positive, those for ln *(1 + # Shares)* are all strongly negative. Depending on the return horizon over which an article's accuracy is computed, the estimates are −0.012 ($t$-statistic = −2.52), −0.007 ($t$-statistic = −1.46), and −0.010 ($t$-statistic = −2.24).

In Online Appendix Table OA4.1, we present results based on a different empirical design. We build on Chen et al. (2014), who provide evidence that the average SA article has investment value in that the overall tone revealed in an SA article positively predicts the corresponding firm's subsequent stock market performance. We estimate a regression of cumulative abnormal returns on an article's tone, *Reliance on Numbers*, ln *(1 + # Read-to-Ends)*, ln *(1 + # Shares)*, and the article's tone interacted with these variables. The estimates for the interaction terms indicate whether the articles that have greater investment value are those that are more heavily rooted in numbers, those with more frequent read-to-ends, or those with more frequent shares.

Consistent with the findings presented in Table 4, the estimates for the interactions between article tone and *Reliance on Numbers* are negative regardless of the return horizon, suggesting that the articles more heavily couched in numbers have lower investment value. More broadly, the estimates for the interactions between article tone and ln *(1 + # Read-to-Ends)* are all positive and statistically significant, suggesting that the articles that investors more frequently read to the end have greater investment value. In stark contrast, the estimates for the interaction term with ln *(1 + # Shares)* are all negative. That is, while the tone in the articles with the more frequent read-to-ends more accurately predicts abnormal returns, the tone in the articles with the more frequent shares does not.

What could explain the fact that, relatively speaking, the articles with the more frequent number of read-to-ends more accurately predict returns, whereas the articles with the more frequent number of shares less accurately predict returns?

We believe that two factors play a role. First, an article can exhibit characteristics that, on one hand, make it highly appealing for impression management and, thus, lead to more frequent sharing but that, on the other hand, happens to lower the article's investment value. The evidence in our previous analyses points to one such characteristic. Quantitative content is more useful for impression management and, thus, more frequently shared. At the same time, our evidence suggests that SA authors' comparative advantage does not lie in the analysis of a company's earnings announcement or recent financial statement. Quantitative SA articles are thus less accurate on average.

We provide another example in Online Appendix Table OA4.2. We find that SA articles discussing large stocks are

**Table 4**
Which stock opinion articles are more accurate?
In this table, we present coefficient estimates from regressions of measures of article accuracy on measures of article consumption and article sharing as well as article characteristics. The sample contains 16,446 opinion articles written on a single stock published on Seeking Alpha from August 2012 through March 2013. The dependent variable in Columns 1–3 is *Article Accuracy*, which equals one if either an article's tone is above the median and the cumulative Daniel et al. (1997; DGTW)-adjusted stock return over one month following article publication (while skipping the first two trading days) is positive, or, an article's tone is below the median and the cumulative DGTW-adjusted stock return is negative; *Article Accuracy* equals zero otherwise. Tone is the number of positive words minus the number of negative words divided by the total number of words in an article. The dependent variable in Columns 4–6 measures an article's accuracy over three months following article publication. The dependent variable in Columns 7–9 measures an article's accuracy over six months following article publication. All other variables are defined in Table 1. We do not report the intercept. *t*-statistics are reported in parentheses and are based on standard errors adjusted for heteroskedasticity and clustered by day of article publication. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively. DJNS = Dow Jones News Service.

| | One month | | | Three months | | | Six months | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| *Reliance on Numbers* | −0.215* | | | −0.312*** | | | −0.314*** | | |
| | (−1.92) | | | (−3.17) | | | (−3.31) | | |
| *ln (1 + # Read-to-Ends)* | | 0.014*** | | | 0.014*** | | | 0.010** | |
| | | (2.96) | | | (2.71) | | | (2.13) | |
| *ln (1 + # Shares)* | | | −0.012** | | | −0.007 | | | −0.010** |
| | | | (−2.52) | | | (−1.46) | | | (−2.24) |
| *Length* | 0.011 | 0.013 | 0.017* | 0.008 | 0.011 | 0.014 | 0.017** | 0.020** | 0.023*** |
| | (1.29) | (1.50) | (1.93) | (1.04) | (1.36) | (1.64) | (1.96) | (2.27) | (2.62) |
| *Editors' Pick* | 0.064*** | 0.057*** | 0.063*** | 0.029 | 0.021 | 0.025 | 0.019 | 0.011 | 0.016 |
| | (3.21) | (2.92) | (3.17) | (1.52) | (1.09) | (1.31) | (1.07) | (0.61) | (0.88) |
| *Presentation* | −0.008 | −0.010 | −0.007 | 0.000 | −0.002 | 0.001 | 0.000 | −0.002 | 0.000 |
| | (−1.09) | (−1.37) | (−0.97) | (0.07) | (−0.24) | (0.10) | (−0.03) | (−0.29) | (0.06) |
| *Long Score* | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.000 | 0.000 | 0.000 |
| | (0.90) | (0.67) | (0.89) | (1.35) | (1.15) | (1.37) | (0.34) | (0.21) | (0.35) |
| *Short Score* | 0.001 | 0.000 | 0.001* | 0.000 | 0.000 | 0.000 | 0.001** | 0.001* | 0.001*** |
| | (1.31) | (0.75) | (1.79) | (−0.22) | (−0.76) | (0.33) | (2.12) | (1.82) | (2.64) |
| *ln (1 + Analyst Coverage)* | −0.005 | −0.005 | −0.007* | −0.006* | −0.006* | −0.008** | −0.008** | −0.008** | −0.009*** |
| | (−1.57) | (−1.35) | (−1.94) | (−1.80) | (−1.68) | (−2.14) | (−2.42) | (−2.36) | (−2.83) |
| *ln (1 + DJNS Coverage)* | 0.003 | 0.002 | 0.003 | 0.008*** | 0.007** | 0.008*** | 0.003 | 0.002 | 0.003 |
| | (0.97) | (0.64) | (1.13) | (2.90) | (2.57) | (3.00) | (1.06) | (0.85) | (1.20) |
| Number of observations | 16,446 | 16,446 | 16,446 | 16,446 | 16,446 | 16,446 | 16,446 | 16,446 | 16,446 |
| Adj. $R^2$ | 0.001 | 0.002 | 0.001 | 0.001 | 0.001 | 0.001 | 0.002 | 0.001 | 0.001 |

shared more frequently. One possible explanation is that people prefer communicating about topics that are familiar to others (Fast et al., 2009; Berger, 2014). The results presented in Online Appendix Table OA4.2 also reveal that articles discussing large-cap stocks are less accurate than those discussing smaller stocks. One possible explanation is that it is difficult for SA authors to uncover unique and important insights for large-cap stocks as they are heavily scrutinized by finance professionals already.

The second factor that we believe helps explain the lower accuracy of the articles that are more frequently shared is that the people most active in content sharing are less informed on average. For the most part, investors should be wary of transmitting an investment opinion that subsequently turns out to be incorrect.[18] Some investors nevertheless choose to share investment ideas with family, friends, and coworkers. Possible reasons for such behavior include overconfidence and limited exposure to the challenges of outperforming the stock market. Prior work, in turn, suggests that overconfident and inexperienced investors are less informed on average (Barber and Odean, 2001; Seru et al., 2010).

To test whether overconfident and inexperienced investors are more active in content sharing, we estimate a regression equation within the sample of three hundred investors we recruit for Experiments A1 and A2. In addition to asking participants whether they would share the quantitative and qualitative article, at the conclusion of Experiments A1 and A2, we pose the following two questions: "What is your gender?" and "Which of the following categories best describes your investment experience?" Possible answer choices to the first question are "Female," "Male," and "Prefer not to say," and possible answer choices to the second question are "Novice investor," "Investor with intermediate experience," "Professional investor," and "Prefer not to say."

The regression equation is estimated at the investor level. The dependent variable is the number of shared articles. Given that each investor is asked to read a quantitative article and a qualitative article, the dependent variable ranges from zero to two. The independent variables are *Male* and *Novice Investor. Male* equals one if the investor reports to be male. Lundeberg, Fox, and Puncochar (1994) and Barber and Odean (2001), among others, provide evidence that men are more overconfident than women. *Novice Investor* equals one if the investor describes himself or herself as a "Novice investor."

We present the results in Online Appendix Table OA4.3. The estimates for *Male* and *Novice Investor* are 0.415 (*t*-

[18] Combined with the inherent difficulty of predicting the stock market, this could explain why the number of shares relative to the number of read-to-ends, both reported in Table 1, is low.

statistic = 4.44) and 0.418 ($t$-statistic = 3.87), respectively. Our results thus suggest that overconfident and inexperienced investors are indeed more active in content sharing.

## 6. Asset pricing implications

If receivers are aware that certain article characteristics lead to more frequent sharing, but also come with lower investment value, or if receivers are aware that, as Plato said, "empty vessels make the loudest sound," then the fact that content shared is less informative need not be a problem. In this section, we examine the degree to which receivers account for such a possibility and its implications for investor performance and asset prices.

If, upon receiving stock recommendations, receivers do not sufficiently account for impression management considerations (Enke, 2020), and if receivers more frequently buy stocks, not short sell (Barber and Odean, 2008), then the stocks that are more frequently mentioned in investors' conversations could become overpriced and subsequently earn abnormally low returns.[19] The initial price run-up and subsequent correction should be more pronounced among stocks that are likely to be short-sale constrained.

To test this hypothesis, we first compute a level of virality. $Virality_{SA}$ is the total number of times SA articles regarding stock $i$ published on day $t$ are shared through e-mail, scaled by the total number of times these articles are viewed or read to the end. We scale by the total number of page views or read-to-ends to exclude any effects arising from increased investor attention that is unrelated to sharing.

At the end of each day, we sort all stocks by $Virality_{SA}$. Following the methodology by Jegadeesh and Titman (1993), we then create two portfolios: a low-virality portfolio, which is the equal-weighted portfolio of 10% of the stocks with the lowest level of $Virality_{SA}$, and a high-virality portfolio, which is the equal-weighted portfolio of 10% of the stocks with the highest level of $Virality_{SA}$. We report the average daily raw returns for each of the two portfolios and the corresponding long-short portfolio. We also report the average daily DGTW-adjusted returns for each of the two portfolios and the corresponding long-short portfolio.

In our sample, around 88% of page views and read-to-ends occur in the first week of article publication.[20] As the time period over which virality may build, we therefore choose the first week of article publication. As the time period over which overpricing may correct, we choose the ensuing three weeks. In asset pricing terms, our high- and low-virality portfolios are held for either one week or three weeks (starting from week 2). $t$-statistics are based on Newey-West standard errors with five lags, which is the equivalent to one calendar week.

Panel A in Table 5 reveals that, if $Virality_{SA}$ is calculated as the total number of shares scaled by the total number of page views, in the first week of article publication, stocks with high virality earn daily DGTW-adjusted returns of +0.13% on average ($t$-statistic = 3.23). This outperformance reverses over weeks 2–4 as high-virality stocks earn average DGTW-adjusted returns of −0.04% a day ($t$-statistic = −1.92). Compared with low-virality stocks, high-virality stocks have +0.19% ($t$-statistic = 2.54) higher daily returns in week 1 and −0.08% ($t$-statistic = −2.21) lower daily returns over weeks 2–4. As the portfolio-holding period is three times longer in the reversal stage, these numbers imply that whatever cumulative outperformance high-virality stocks accrue in the first week is more than eliminated over the ensuing three weeks.

We make the same observations when calculating $Virality_{SA}$ as the total number of shares scaled by the total number of read-to-ends (Panel B in Table 5). Here, in week 1, high-virality stocks earn daily DGTW-adjusted returns that are 0.24% higher than those experienced by low-virality stocks ($t$-statistic = 3.21). Over weeks 2–4, high-virality stocks under-perform low-virality stocks by 0.08% ($t$-statistic = −2.04).

As shown in Table 5, the under-performance is of almost identical magnitude when considering raw returns. In untabulated analyses, we find that, depending on the measure of $Virality_{SA}$, the fraction of stocks that experience negative raw returns is 5% or 6% higher among high-virality stocks than low-virality stocks.

Fig. 3 plots the cumulative differential performance of high-virality stocks over low-virality stocks in event days from trading day 1 to trading day 60 since article publication. The plots show that virality and overpricing continuously build in the first week of article publication and peak on trading day 5 or 6, depending on whether we measure performance through raw returns or DGTW-adjusted returns and whether we scale the total number of shares by the total number of page views or the total number of read-to-ends. Overpricing corrects over the ensuing one to three weeks.

Online Appendix Tables OA5.1 and OA5.2 report the results from various sensitivity analyses. In Online Appendix Table OA5.1, instead of reporting the average DGTW-adjusted returns, we present alphas based on both the Fama-French three-factor model and the Fama-French five-factor model (Fama and French, 1993, 2015). The results are similar, both economically and statistically, compared with those reported in Table 5. In Online Appendix Table OA5.2, we report results from Fama-MacBeth regressions of returns on a high-virality indicator and various firm characteristics. Again, the results are similar to those reported in Table 5.

### 6.1. Role of short-sale constraints

In Table 6, we separately consider stocks that are likely to be short-sale constrained. We measure short-sale constraints based on the average daily lending fees (from Markit), and we consider stocks as likely to be short-sale constrained if their average daily lending fees in the month of article publication are in the top 30% of their distribution at a particular time.

[19] If propagators of information are more likely to send buy recommendation to overpricing, we no longer require the second assumption that (upon receiving a stock recommendation) investors more frequently buy (not short).

[20] This information was provided to us by SA.

**Table 5**
Virality and stock returns.
This table reports the daily average raw returns or daily average Daniel et al. (1997; DGTW)-adjusted returns for portfolios formed by the level of virality a stock receives. The sample is based on 16,446 opinion articles written on a single stock published on Seeking Alpha (SA) from August 2012 through March 2013. For each stock, at the end of each day, we compute the level of virality as the total number of times SA articles about the corresponding stock were shared through e-mail, scaled by the total number of times SA articles were viewed (Panel A) or read to the end (Panel B). Each day, we rank stocks based on their level of virality. "Low virality" is the equal-weighted portfolio of 10% of the stocks with the lowest level of virality. "High virality" is the equal-weighted portfolio of 10% of the stocks with the highest level of virality. Portfolios are held either for one week or for three weeks (starting from week 2). *t*-statistics are based on Newey-West standard errors (five lags) and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

| | Week 1 | | Weeks 2–4 | |
|---|---|---|---|---|
| Virality | Raw returns | DGTW-adjusted returns | Raw returns | DGTW-adjusted returns |
| *Panel A: Virality = # Shares / # Page Views* | | | | |
| Low virality | 0.04% | −0.06% | 0.15%** | 0.04% |
| | (0.45) | (−0.82) | (2.18) | (1.31) |
| High virality | 0.24%*** | 0.13%*** | 0.05% | −0.04%* |
| | (3.00) | (3.23) | (0.83) | (−1.92) |
| High minus low | 0.20%*** | 0.19%** | −0.10%** | −0.08%** |
| | (2.75) | (2.54) | (−2.43) | (−2.21) |
| Number of observations | 169 | 169 | 181 | 181 |
| *Panel B: Virality = # Shares / # Read-to-Ends* | | | | |
| Low virality | 0.01% | −0.09% | 0.15%** | 0.04% |
| | (0.09) | (−1.17) | (2.22) | (1.45) |
| High virality | 0.26%*** | 0.16%*** | 0.07% | −0.03% |
| | (3.38) | (3.74) | (1.09) | (−1.41) |
| High minus low | 0.26%*** | 0.24%*** | −0.09%** | −0.08%** |
| | (3.40) | (3.21) | (−2.13) | (−2.04) |
| Number of observations | 169 | 169 | 181 | 181 |

**Table 6**
Virality and stock returns among stocks likely to be short-sale constrained.
This table replicates Table 5 for stocks for which average daily lending fees in the month of article publication are in the top 30% of their distribution. DGTW = Daniel et al. (1997).

| | Week 1 | | Weeks 2–4 | |
|---|---|---|---|---|
| Virality | Raw returns | DGTW-adjusted returns | Raw returns | DGTW-adjusted returns |
| *Panel A: Virality = # Shares / # Page Views* | | | | |
| Low virality | −0.00% | −0.11% | 0.20% | 0.08% |
| | (−0.00) | (−0.45) | (1.72) | (0.97) |
| High virality | 0.38%** | 0.27%** | −0.06% | −0.16%** |
| | (2.38) | (2.11) | (−0.55) | (−2.19) |
| High minus low | 0.38% | 0.38% | −0.26%** | −0.24%** |
| | (1.45) | (1.45) | (−2.29) | (−2.22) |
| Number of observations | 166 | 166 | 179 | 179 |
| *Panel B: Virality = # Shares / # Read-to-Ends* | | | | |
| Low virality | −0.01% | −0.12% | 0.20% | 0.09% |
| | (−0.03) | (−0.51) | (1.89) | (1.16) |
| High virality | 0.31%** | 0.22% | −0.02% | −0.13%* |
| | (1.99) | (1.63) | (−0.20) | (−1.81) |
| High minus low | 0.32% | 0.34% | −0.22%** | −0.22%** |
| | (1.20) | (1.29) | (−2.01) | (−2.09) |
| Number of observations | 166 | 166 | 179 | 179 |

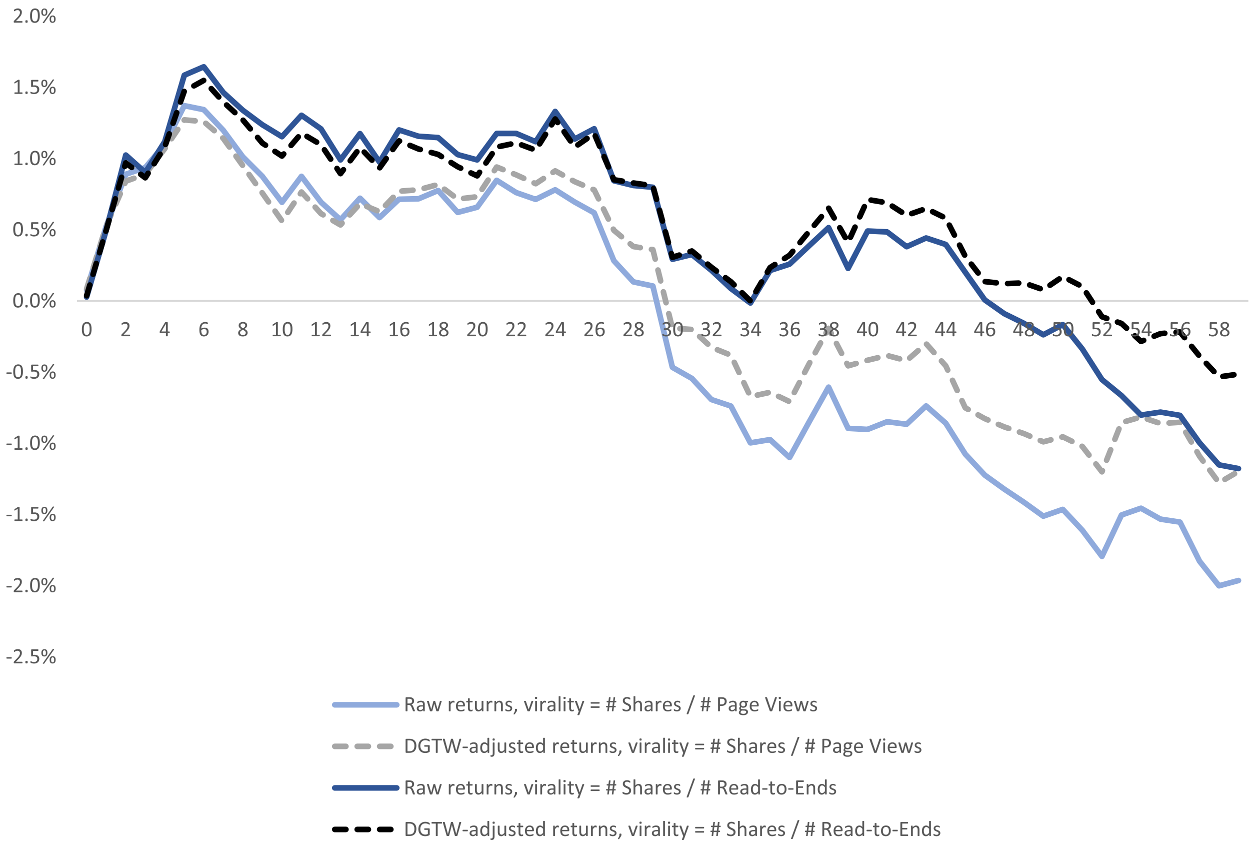


**Fig. 3.** Virality and cumulative stock returns in event time. This figure mimics Table 5, but we now plot the average cumulative daily returns, in event time, when going long the high-virality portfolio and going short the low-virality portfolio. The solid lines are based on raw returns, and the dashed lines are based on Daniel et al. (1997; DGTW)-adjusted returns.

The result we find particularly striking is that, among short-sale constrained stocks, the seeming overpricing that develops during week 1 is so high that, over weeks 2–4, stocks with high virality experience not only negative DGTW-adjusted returns but also negative raw returns.

We would like to stress that we do not believe that the anomalous returns shown in Tables 5 and 6 are generated by the sharing of SA articles per se. In this study, we merely assume that the sharing behavior of SA readers provides a representative glimpse of the sharing behavior of non-negligible parts of the investor population. That is, we assume that, when SA readers tend to share certain ideas, so do non-negligible parts of the investor population.

We must also stress that our attempt at deriving asset pricing implications using an eight-month sample period is ambitious. Our results should thus be interpreted with caution.

### 6.2. Additional evidence based on twitter

In response to the above caveat, we repeat our previous analysis on a six-year sample containing all tweets regarding publicly traded firms in the United States. We write a web scraping program that uses Twitter's Advanced Search function (https://twitter.com/search-advanced). We then search for all tweets written in English and containing the dollar sign ($) followed by a ticker of a stock traded in the United States (e.g., $AAPL).[21] In total, our sample contains 90,994,236 tweets on 10,068 stocks from January 2013 through December 2018. For each of the 91 million tweets, we obtain the following information: tweet ID, date and time the tweet was posted, user ID (i.e., screen_name on Twitter), content, and number of retweets. We make our Twitter data aggregated to the stock/day level freely available on our personal websites.[22]

At the end of each day, we compute the total number of tweets about each stock $i$, as well as how many times those tweets are retweeted. The level of virality

[21] Since July 31, 2012 (https://twitter.com/twitter/status/230098997010911233), users can make a ticker symbol clickable by adding the $ symbol (e.g., $AAPL). This feature makes it easier to identify tweets that discuss stocks. However, some noise still exists. In our analysis, we exclude 13 “tickers” from our sample that have more than one million tweets per year: ABC, EBAY, DON, ETSY, FOX, GAGA, GOP, JACK, KIM, LFC, PTI, VOX, and WWE. To put the one million tweets per year in perspective, AAPL, which is one of the most popular tickers tweeted in our sample, has (only) 276,256 tweets in 2018. We believe the reason these 13 “tickers” have so many tweets is that users use the $ symbol not only to denote a stock ticker. All of the 13 “tickers” are names or part of the names (or usernames) of products and celebrities.

[22] http://www.hlchen.com and http://www.bhwang.com

**Table 7**
Virality and stock returns: evidence based on tweets and retweets.
This table reports the daily average raw returns or daily average Daniel et al. (1997; DGTW)-adjusted returns for portfolios formed by the level of virality a stock receives. The sample is based on 91,262,601 tweets covering 10,079 stocks from January 2013 through December 2018. For each stock, at the end of each day, we compute the level of virality as the total number of retweets of tweets about the corresponding stock, scaled by the total number of tweets. Each day, we rank stocks based on their level of virality. "Low virality" is the equal-weighted portfolio of stocks for which virality equals zero. "High virality" is the equal-weighted portfolio of stocks for which virality is greater than zero. Portfolios are held either for one week or for three weeks (starting from week 2). *t*-statistics are based on Newey-West standard errors (five lags) and are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

| | Week 1 | | Weeks 2–4 | |
|---|---|---|---|---|
| Virality | Raw returns | DGTW-adjusted returns | Raw returns | DGTW-adjusted returns |
| Low virality | 0.04%* | −0.00% | 0.04%* | 0.00% |
| | (1.81) | (−1.00) | (1.88) | (0.61) |
| High virality | 0.09%*** | 0.04%*** | 0.03% | −0.01%** |
| | (3.44) | (7.99) | (1.27) | (−2.23) |
| High minus low | 0.04%*** | 0.04%*** | −0.01%** | −0.01%** |
| | (7.58) | (8.16) | (−2.18) | (−2.50) |
| Number of observations | 1511 | 1511 | 1507 | 1507 |

$Virality_{Twitter}$ is the total number of retweets, scaled by the total number of tweets. At the end of each day, we sort all stocks by $Virality_{Twitter}$. Unlike those of $Virality_{SA}$, most of the realizations of $Virality_{Twitter}$ are zero.[23] Therefore, instead of considering the returns of stocks in the top and bottom decile, we contrast the performance of stocks for which $Virality_{Twitter}$ is equal to zero to that of stocks for which $Virality_{Twitter}$ is greater than zero. We report the average daily raw returns for each of the two portfolios, the corresponding long-short portfolio, and the average daily DGTW-adjusted returns. Portfolios are again held for either one week or three weeks (starting from week 2).

The advantage of considering retweets over the sharing of SA articles is that the sample period is much longer. A potential disadvantage is that, compared with SA articles, tweets could be less informative and less representative of actual investors' viewpoints. In line with this notion, in Online Appendix Table OA6, we show that, compared with the tone of tweets, the tone of articles published on SA much more strongly correlates with contemporaneous abnormal returns.

Table 7 reports our results. We find that, in week 1, high-virality stocks earn DGTW-adjusted returns of +0.04% a day (*t*-statistic = 7.99). The initial outperformance reverses in weeks 2–4, as high-virality stocks earn daily DGTW-adjusted returns of −0.01% (*t*-statistic = −2.23). Compared with low-virality stocks, high-virality stocks have +0.04% (*t*-statistic = 8.16) higher daily DGTW-adjusted returns in week 1 and −0.01% (*t*-statistic = −2.50) lower daily DGTW-adjusted returns over weeks 2–4. The results are similar for raw returns.

Overall, our Twitter analysis produces results similar to those produced when considering SA data. Consistent with tweets less accurately reflecting actual investors' viewpoints, the economic magnitude of the Twitter-based results is substantially smaller than that of the SA-based results.

### 6.3. Additional evidence based on an investor survey

Our final test provides additional survey-based evidence. At the conclusion of Experiments 1 and 2, we ask all 540 investors Q1: "Over the past 12 months, did a coworker, friend, or family member mention a stock to you that they thought you might be interested in buying?" The answer choices are "Yes" and "No."

All investors who respond "Yes" are asked Q2: "How thoroughly do you think that person researched that stock before mentioning it to you?" and Q3: "Did you end up buying the stock?" The scales for Q2 are 1 ("rather casually") to 5 ("rather thoroughly"). The answer choices for Q3 are "Yes" and "No."

All investors who respond "Yes" to Q3 are asked Q4: "What was or has been your overall return since you bought the stock? If you are not sure, please answer 'don't know.'" The answer choices are "(1) less than −20%, (2) between −20% and −10.01%, (3) between −10% and −0.01%, (4) between 0% and +10%, (5) between +10.01% and +20%, (6) greater than +20%, (7) don't know."

In Column 1 of Table 8, we report the responses across all investors. In Columns 2 and 3, we report the responses separately for mass-market and affluent investors, respectively.

Consistent with social interactions playing an important role in the transmission of investment ideas, 43.7% of investors respond "Yes" to Q1. Consistent with investors taking signals received through word-of-mouth seriously, responses to Q2 and Q3 show that more investors believe that the sender "rather thoroughly" researched the corresponding stock than "rather casually" and 33.9% of investors report to have bought the underlying stock.

The responses from affluent investors are similar to those from mass-market investors. While a smaller portion of affluent investors reports to have received an investment

[23] The distribution of $Virality_{Twitter}$ is such that $Virality_{Twitter}$ starts becoming greater than zero only after the 74th percentile.

**Table 8**
Virality and stock returns: evidence based on an investor survey.
In this table, we present responses from an investor survey. At the conclusion of the two experiments (Experiment 1 and Experiment 2) described in Table 3, we ask all 540 investors Q1: "Over the past 12 months, did a coworker, friend, or family member mention a stock to you that they thought you might be interested in buying?" All investors who respond "Yes" to Q1 are asked Q2: "How thoroughly do you think that person researched that stock before mentioning it to you?" and Q3: "Did you end up buying the stock?" All investors who respond "Yes" to Q3 are asked Q4: "What was or has been your overall return since you bought the stock? If you are not sure, please answer, 'don't know.'" In Column 1, we report the responses across all investors. In Columns 2 and 3, we report the responses separately for investors with net investable assets worth equal to or less than $300,000 and worth more than $300,000, respectively.

| | All investors (1) | Investors with net investable assets | |
|---|---|---|---|
| Investors' responses | | ≤ $300,000 (2) | > $300,000 (3) |
| Q1: "Over the past 12 months, did a coworker, friend or family member mention a stock to you that they thought you might be interested in buying?" | | | |
| "Yes"- | 43.7% | 45.9% | 34.0% |
| Q2: "How thoroughly do you think that person researched that stock before mentioning it to you?" | | | |
| "1 (rather casually)" | 8.1% | 8.4% | 5.9% |
| "2″ | 25.0% | 26.7% | 14.7% |
| "3″ | 23.7% | 22.8% | 29.4% |
| "4″ | 27.5% | 26.2% | 35.3% |
| "5 (rather thoroughly)" | 15.7% | 15.8% | 14.7% |
| Q3: "Did you end up buying the stock?" | | | |
| "Yes" | 33.9% | 32.7% | 41.2% |
| Q4: "What was or has been your overall return since you bought the stock? If you are not sure, please answer 'don't know.'" | | | |
| "Less than −20%" | 2.5% | 1.5% | 7.1% |
| "Between −20% and −10.01%" | 11.3% | 9.1% | 21.4% |
| "Between −10% and −0.01%" | 10.0% | 10.6% | 7.1% |
| "Between 0% and +10%" | 41.3% | 45.5% | 21.4% |
| "Between +10.01% and +20%" | 15.0% | 12.1% | 28.6% |
| "Greater than +20%" | 8.8% | 10.6% | 0.0% |
| "Don't know" | 11.3% | 10.6% | 14.3% |

idea through word-of-mouth (34.0% for affluent investors compared with 45.9% for mass-market investors), affluent investors more frequently indicate that they acted on the idea and ended up buying the underlying stock (41.2% for affluent investors compared with 32.7% for mass-market investors).

We conduct a simulation on what returns ten thousand investors would have realized as of the survey completion date if each investor, over the previous 12 months, had randomly purchased a stock from the Center for Research in Security Prices universe. We plot the distribution of realized returns in Online Appendix Table OA7. The simulation results show that 12% of the investors would have realized a raw return of less than −10% if they had randomly purchased a stock and 53% of them would have realized a raw return of greater than +10%.

In our survey, among the investors who respond to Q4 and do not choose "don't know" as their answer, 16% experienced a raw return of less than −10% (compared with 12% in our simulation). 27% experienced a raw return of greater than +10% (compared with 53% in our simulation). That is, compared with the investors in our simulation who randomly purchased a stock, those in our survey more frequently experienced strongly negative returns and substantially less frequently experienced strongly positive returns. As presented in Online Appendix Table OA7, the simulation results are similar when varying the universe from which we randomly select stocks and when imposing different assumptions regarding investors' holding period.

Overall, a comparison of the survey responses with the simulation results shows that financial advice and investment ideas transmitted through social interactions can lead to poor investment performances.

## 7. Conclusion

Given how much people rely on information transmitted through social interactions, it is crucial we understand (1) what types of content propagate in financial markets through social interactions, (2) whether any selections in content represent social transmission biases or simply mechanisms to aggregate information efficiently, and (3) the overall implications of these factors for investor trading and asset prices.

Our study begins to shed light on these issues by applying one of the building blocks in social psychology, the impression management concept, to the financial setting. We provide evidence that investors' communication decisions are affected by impression management considerations. Given that stories well suited for impression management are not always value relevant, self-presentation purposes can cause investors to inadvertently propagate noise, leading to both price dislocations and poor investment decisions.

There are likely economic considerations as well as other social-psychological motives that govern investors' sharing decisions. Berger (2014) and Gilovich et al. (2019) provide an overview of some of the social-psychological motives. Exploring which of these motives are the most relevant in the financial setting could prove to be an interesting avenue for future research in social finance.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jfineco.2021.09.004.

## References


Avery, C.N., Chevalier, J.A., Zeckhauser, R.J., 2016. The "CAPS" prediction system and stock market returns. Rev. Financ. 20, 1363–1381.
Baek, E.C., Scholz, C., Falk, E.B., 2020. The neuroscience of persuasion and information propagation. In: Floyd, K., Weber, R. (Eds.), The Handbook of Communication Science and Biology. Routledge, New York, pp. 122–133.
Barber, B.M., Odean, T., 2001. Boys will be boys: gender, overconfidence, and common stock investment. Q. J. Econ. 116, 261–292.
Barber, B.M., Odean, T., 2008. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev. Financ. Stud. 21, 785–818.
Barrot, J.N., Kaniel, R., Sraer, D., 2016. Are retail traders compensated for providing liquidity? J. Financ. Econ. 120, 146–168.
Baumeister, R.F., 1982. A self-presentational view of social phenomena. Psychol. Bull. 91, 3–26.
Berger, J., 2014. Word of mouth and interpersonal communication: a review and directions for future research. J. Consum. Psychol. 24, 586–607.
Berger, J., Milkman, K.L., 2012. What makes online content viral? J. Mark. Res. 49, 192–205.
Berger, J., Schwartz, E.M., 2011. What drives immediate and ongoing word of mouth? J. Mark. Res. 48, 869–880.
Boehmer, E., Jones, C.M., Zhang, X., Zhang, X., 2021. Tracking retail investor activity. J. Finance 76, 2249–2305. doi:10.1111/jofi.13033, forthcoming**[Has this been published?]**.
Brown, J.R., Ivković, Z., Smith, P.A., Weisbenner, S., 2008. Neighbors matter: causal community effects and stock market participation. J. Finance 63, 1509–1531.
Chen, H., De, P., Hu, Y.J., Hwang, B.H., 2014. Wisdom of crowds: the value of stock opinions transmitted through social media. Rev. Financ. Stud. 27, 1367–1403.
Chen, A.Y., Zimmermann, T., 2020. Publication bias and the cross section of stock returns. Rev. Asset Pricing Stud. 10, 249–289.
Cragg, L., Gilmore, C., 2014. Skills underlying mathematics: the role of executive function in the development of mathematics proficiency. Trends Neurosci. Educ. 3, 63–68.
Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. J. Finance 52, 1035–1058.
Dubois, D., Bonezzi, A., De Angelis, M., 2016. Sharing with friends versus strangers: how interpersonal closeness influences word-of-mouth valence. J. Mark. Res. 53, 712–727.
Dunbar, R.I.M., 1992. Neocortex size as a constraint on group size in primates. J. Hum. Evol. 22, 469–493.
Dunbar, R.I.M., 1998. The social brain hypothesis. Evolut. Anthropol. Issues News Rev. 6, 178–190.
Enke, B., 2020. What you see is all there is. Q. J. Econ. 135, 1363–1398.
Escobar, L., Pedraza, A., 2019. Active trading and (poor) performance. World Bank, Washington, DC Unpublished working Paper.
Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33, 3–56.
Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. J. Financ. Econ. 116, 1–22.
Fast, N.J., Heath, C., Wu, G., 2009. Common ground and cultural prominence: how conversation strengthens culture. Psychol. Sci. 20, 904–911.
Fracassi, C., 2017. Corporate finance policies and social networks. Manag. Sci. 63, 2420–2438.
Gao, L., Wheeler, S.C., Shiv, B., 2009. The "shaken self": product choices as a means of restoring self-view confidence. J. Consum. Res. 36, 29–38.
Gilovich, T., Keltner, D., Chen, S., Nisbett, R.E., 2019. Social psychology. W. W. Norton & Company, New York.
Goffman, E., 1978. The presentation of self in everyday life. Harmondsworth, London, UK.
Haliassos, M., Jansson, T., Karabulut, Y., 2020. Financial literacy externalities. Rev. Financ. Stud. 33, 950–989.
Han, B., Hirshleifer, D., Walden, J., 2021. Social transmission bias and investor behavior. J. Financ. Quant. Anal. 1–23. doi:10.1017/S0022109021000077, forthcoming.
Heimer, R.Z., 2016. Peer pressure: social interaction and the disposition effect. Rev. Financ. Stud. 29, 3177–3209.
Heimer, R.Z., Simon, D., 2017. Facebook finance: how social interaction propagates active investing. Boston College, Chestnut Hill, MA Unpublished working Paper.
Hill, R.A., Dunbar, R.I.M., 1998. An evaluation of the roles of predation rate and predation risk as selective pressures on primate grouping behaviour. Behaviour 135, 411–430.
Hirshleifer, D., 2015. Behavioral finance. Ann. Rev. Financ. Econ. 7, 133–159.
Hirshleifer, D., 2020. Presidential address: social transmission bias in economics and finance. J. Finance 75, 1779–1831.
Hong, H., Kubik, J.D., Stein, J.C., 2004. Social interaction and stock market participation. J. Finance 59, 137–163.
Huang, X., Nekrasov, A., Teoh, S.H., 2018. Headline salience, managerial opportunism, and over- and underreactions to earnings. Account. Rev. 93, 231–255.
Hvide, H.K., Östberg, P., 2015. Social interaction at work. J. Financ. Econ. 117, 628–652.
Hwang, B.H., Liberti, J.M., Sturgess, J., 2019. Information sharing and spillovers: evidence from financial analysts. Manag. Sci. 65, 3624–3636.
Jame, R., Johnston, R., Markov, S., Wolfe, M.C., 2016. The value of crowdsourced earnings forecasts. J. Account. Res. 54, 1077–1110.
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. J. Finance 48, 65–91.
Kadous, K., Koonce, L., Towry, K.L., 2005. Quantification and persuasion in managerial judgement. Contemp. Account. Res. 22, 643–686.
Kaniel, R., Liu, S., Saar, G., Titman, S., 2012. Individual investor trading and return patterns around earnings announcements. J. Finance 67, 639–680.
Kaniel, R., Saar, G., Titman, S., 2008. Individual investor trading and stock returns. J. Finance 63, 273–310.
Kaustia, M., Knüpfer, S., 2012. Peer performance and stock market entry. J. Financ. Econ. 104, 321–338.
Kelley, E.K., Tetlock, P.C., 2013. How wise are crowds? insights from retail orders and stock returns. J. Finance 68, 1229–1265.
Kelley, E.K., Tetlock, P.C., 2017. Retail short selling and stock prices. Rev. Financ. Stud. 30, 801–834.
Koetsenruijter, A.W.M., 2011. Using numbers in news increases story credibility. Newsp. Res. J. 32, 74–82.
Lakin, J.L., Chartrand, T.L., 2003. Using nonconscious behavioral mimicry to create affiliation and rapport. Psychol. Sci. 14, 334–339.
Lakin, J.L., Jefferis, V.E., Cheng, C.M., Chartrand, T.L., 2003. The chameleon effect as social glue: evidence for the evolutionary significance of nonconscious mimicry. J. Nonverbal Behav. 27, 145–162.
Lane, J., Lim, S.S., Uzzi, B., 2020. Biased information transmission in investor social networks: evidence from professional traders. Harvard University, Cambridge, MA Unpublished working paper..
Leary, M.R., Kowalski, R.M., 1990. Impression management: a literature review and two-component model. Psychol. Bull. 107, 34–47.
Lerner, J., Malmendier, U., 2013. With a little help from my (random) friends: success and failure in post–business school entrepreneurship. Rev. Financ. Stud. 26, 2411–2452.
Lieberman, M.D., 2013. Social: why our brains are wired to connect. Crown Publishers, New York.
Loughran, T., McDonald, B., 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-Ks. J. Finance 66, 35–65.
Lovett, M.J., Peres, R., Shachar, R., 2013. On brands and word of mouth. J. Mark. Res. 50, 427–444.
Lundeberg, M.A., Fox, P.W., Puncochar, J.M., 1994. Highly confident but wrong: gender differences and similarities in confidence judgments. J. Educ. Psychol. 86, 114–121.
Maturana, G., Nickerson, J., 2019. Teachers teaching teachers: the role of workplace peer effects in financial decisions. Rev. Financ. Stud. 32, 3920–3957.
McConway, K., 2016. Statistics and the media: a statistician's view. Journalism 17, 49–65.

Mead, N.L., Baumeister, R.F., Stillman, T.F., Rawn, C.D., Vohs, K.D., 2011. Social exclusion causes people to spend and consume strategically in the service of affiliation. J. Consum. Res. 37, 902–919.
Packard, G., Wooten, D.B., 2013. Compensatory knowledge signaling in consumer word-of-mouth. J. Consum. Psychol. 23, 434–450.
Parsons, S., Bynner, J., 2005. Does numeracy matter more?. National Research and Development Centre for Adult Literacy and Numeracy, London, UK.
Pool, V.K., Stoffman, N., Yonker, S.E., 2015. The people in your neighborhood: social interactions and mutual fund portfolios. J. Finance 70, 2679–2732.
Roeh, I., Feldman, S., 1984. The rhetoric of numbers in front-page journalism: how numbers contribute to the melodramatic in the popular press. Text 4, 347–368.
Rucker, D.D., Galinsky, A.D., 2008. Desire to acquire: powerlessness and compensatory consumption. J. Consum. Res. 35, 257–267.
Rucker, D.D., Hu, M., Galinsky, A.D., 2014. The experience versus the expectations of power: a recipe for altering the effects of power on behavior. J. Consum. Res. 41, 381–396.
Seru, A., Shumway, T., Stoffman, N., 2010. Learning by trading. Rev. Financ. Stud. 23, 705–739.
Shiller, R.J., Pound, J., 1989. Survey evidence on diffusion of interest and information among investors. J. Econ. Behav. Organ. 12, 47–66.
Shue, K., 2013. Executive networks and firm policies: evidence from the random assignment of MBA peers. Rev. Financ. Stud. 26, 1401–1442.
Smith, E.R., Mackie, D.M., 2000. Social psychology. Taylor & Francis, Philadelphia, PA.
van Dijk, T.A., 1988. News as discourse. Lawrence Erlbaum Associates, Hillsdale, NJ.
Weiss, B., Feldman, R.S., 2006. Looking good and lying to do it: deception as an impression management strategy in job interviews. J. Appl. Soc. Psychol. 36, 1070–1086.
Wicklund, R.A., Gollwitzer, P.M., 1981. Symbolic self-completion, attempted influence, and self-deprecation. Basic Appl. Soc. Psychol. 2, 89–114.
Wicklund, R.A., Gollwitzer, P.M., 1982. Symbolic self-completion. Lawrence Erlbaum Associates, Hillsdale, NJ.