# Bias in the Crowds? An Examination of Whether Financial Positions Affect Investor Perceptions of Stock Opinions Transmitted Through Social Media

**John L. Campbell***
J.M. Tull School of Accounting, Terry College of Business
University of Georgia
Athens, GA 30602-6252
706.542.3595 (phone)
johnc@uga.edu

**Matthew D. DeAngelis**
School of Accountancy
J. Mack Robinson College of Business
Georgia State University

**James R. Moon, Jr.**
School of Accountancy
J. Mack Robinson College of Business
Georgia State University

March 2017

* Corresponding Author. We thank Vikas Agarwal, Kris Allee, Steve Baginski, Ted Christensen, James Chyz, Jenna Feagin, Jackie Hammersley, Frank Heflin, Todd Kravet, Santhosh Ramalingegowda, Bridget Stomberg, Quinn Swannquist, Steve Utke, Dana Wallace, Dave Weber, Ben Whipple, Mike Willenborg, Yuan Xie (discussant), Sunny Yang, Kexin Zheng, Ying Zhou, Flora Zhou, participants of the FARS 2017 midyear meeting, and workshop participants at the University of Connecticut, the University of Georgia, and Georgia State University for helpful comments and suggestions. We also thank Vic Lee and Bei Shi for excellent research assistance.

# Bias in the Crowds? An Examination of Whether Financial Positions Affect Investor Perceptions of Stock Opinions Transmitted Through Social Media

**ABSTRACT:** Motivated by concerns that information intermediaries' stock positions create a conflict of interest that impairs their objectivity, we examine investor perceptions of the stock positions of authors on the social media site SeekingAlpha (hereafter SA) and offer three primary findings. First, an author's long (short) positions are positively (negatively) associated with short-window returns surrounding the article's publication, suggesting that stock positions convey a signal about the author's overall opinion of the firm that is incremental to the analysis in the article. Second, we find that the price response attributed to article tone is significantly stronger for articles authored by individuals with stock positions (i.e., skin in the game). Our results suggest that investors perceive authors with short positions to be more credible *regardless* of their tone, but that authors with long positions are *only* more credible when expressing negative sentiment. Finally, investors appear to find positive tone written by authors with short positions (i.e., short authors that write with a tone that is opposite to their underlying stock position) to be the most credible of all. Overall, our results suggest that an author's stock positions do not impair that author's credibility with investors, and, in many cases, these positions actually enhance the author's credibility.

## 1. Introduction

Information intermediaries, such as the business press, professional financial analysts, and social media contributors, routinely provide analysis of firms' performance. Practitioners and regulators have expressed concern that the stock positions of these information intermediaries may impair the objectivity of their analysis. For example, with respect to financial analysts, the Securities and Exchange Commission (SEC) argues that stock positions "can create pressure on…independence and objectivity," although "the existence of these relationships does not necessarily mean…bias" (SEC 2016). To date, there is no comprehensive evidence on the extent to which stock positions impair information intermediaries' objectivity.

Using investor-authors on the social media site Seeking Alpha (www.seekingalpha.com, hereafter SA) as a proxy for information intermediaries, we examine whether information intermediaries' stock positions in the firms about which they write improves or impairs the information environment to which they contribute.[1] SA authors largely consist of independent financial "analysts" who are interested in establishing, building, and maintaining a reputation within the investment community, as well as individual and institutional investors who have information they wish to convey in order to accelerate price formation and enhance the value of their portfolios (e.g., Pasquariello and Wang 2016). Recognizing the potential for personal financial interests biasing authors' analysis, SA requires authors to disclose whether they have a long or short position in the firms named in articles that they write. Importantly, SA does not prohibit authors from trading in the stocks of firms named in their articles (i.e., there is no "blackout period," as with professional analysts). Thus, SA should be a powerful setting in which to test whether stock positions create a perceived conflict of interest. Consistent with regulators'

[1] We refer to SA as a "social media" platform because it allows users to share and discuss investment strategies and stock opinions. However, unlike Facebook or Twitter, SA has an editorial process, which we discuss later, that vets and curates the content we use in our analysis.

concerns, an author's stock positions could impair their analysis and reduce the quality of information conveyed. On the other hand, the author having "skin in the game" could enhance an author's analysis and align the interests of the author and other shareholders. Ultimately, whether investors perceive stock positions to impair or enhance the impact of the author's analysis is an empirical question.

We examine three specific research questions. First, does an author's financial position convey information incremental to the content of their analysis? If so, the direction of that position should reveal the author's private information to investors. On the other hand, if a financial position indicates bias on the part of the author then investors should not find the position to be incrementally informative. Then, we ask two final research questions. Do investors find an author's analysis more or less credible when the author has a financial position (i.e., the author has "skin in the game")? Finally, do investors respond most strongly to tone contrary to the author's position (i.e., to negative tone when the author has a long position, or to positive tone when the author has a short position)?

To investigate these questions, we extract the information conveyed in 114,120 SA articles covering 4,016 unique firms released on 80,589 unique firm-days from 2006 to 2015.[2] For each article, we identify the ticker symbol associated with the firm about which the article is written ("primary ticker"), extract the disclosure indicating whether the author has a position in the stock ("disclaimer"), and measure the tone of both the article and the comments to the article using Loughran and McDonald's (2011) list of positive and negative financial terms. We code the author's position as long (short) if the primary stock ticker appears in the author's listed long (short) holdings and no position otherwise. We then match each article to CRSP using the primary

[2] We begin with the full universe of articles on SA. Between 2006 and mid-2015, we identify nearly 500 thousand articles, excluding news flashes and other short-form updates, however the majority of them (approximately 80 percent) are from 2012 onwards. Table 1 describes the sample attrition from the total number of available articles to our sample size of 80,589 firm-day observations.

ticker and compute abnormal returns for the two days surrounding the article release.[3]

Because this data is not well explored in the academic literature, we begin by providing background information on SA authors, which we acquire from two sources. First, we obtain demographic information from SeekingAlpha for the full universe of SA authors. Second, we programmatically analyze the biographies of authors in our sample. Collectively, this data suggests that roughly seventy-five percent of SA authors reveal their name and place of employment, implying that the large majority of authors face reputational concerns not only online but also in their "day jobs." Furthermore, the average pay received by SA authors is quite small ($33.33 per month), and 30 percent of SA authors have their own independent investment blog.[4] Taken together, these results suggests that, on average, SA authors seek to establish a reputation within the financial community or accelerate stock price formation rather than to earn money directly from their SA activities. Finally, the average SA author in our sample has a following of about 4,500 users, each of whom receives an email whenever the author publishes an article, suggesting that SA provides a meaningful platform for promoting and disseminating an author's analysis.

To address our research questions, we begin by examining whether an author's disclosure of a financial position predicts abnormal stock returns in the two trading days surrounding the article release (i.e., the 0, +1 event window). If investors perceive that a disclosure of financial position provides a signal about the author's opinion of the firm that is incremental to the analysis in the article, we should find that returns are more positive when author has a long position and more negative when the author has a short position. We then examine how an author's stock

[3] Chen, De, Hu, and Hwang (2014) show that negative sentiment in SA articles predicts long-run (60 day) returns. We consider long-run returns in supplemental analysis.

[4] SA prohibits authors from publishing the full content of any article accepted by SA editors in other sources, such as a personal blog. Authors can, however, publish a summary of the analysis in other places, such as a personal blog, so long as they furnish a link to the SA article as well. While we do not expect this to influence our analyses, in a sensitivity analysis we exclude articles authored by those we believe to host personal websites from our sample and our results are quantitatively and qualitatively unchanged.

position affects investor perception of the credibility of an article by comparing the reaction to the tone of articles written by authors with positions to the reaction to tone authored by contributors with no positions. If holding a stock position enhances an author's credibility, investors should respond more strongly to the tone of an article when the author has a stock position compared to when they do not. Alternatively, if personal stock positions impair an author's credibility, investors should respond less strongly to the tone in an article when the author has a stock position compared to when they do not. Finally, we compare the reaction to an article's tone after partitioning separately on authors with long stock positions and authors with short stock positions. This analysis serves two purposes. First, it allows us to determine whether investors consider short and long authors equally credible. Second, by separately measuring positive and negative tone we can test whether investors react more strongly to tone contrary to an author's position.

We offer three main findings. First, author stock positions contribute directly to short-window returns surrounding the article's publication after controlling for an array of other measures of news, including the tone of both the article itself, comments written by others about the article, and the sentiment expressed in contemporaneously issued news by analysts, the firm, or the business press. These findings are not only statistically significant, but also economically significant: for instance, long-authored articles correspond to a positive 2-day return of 0.4 of 0.5 percent, while short-authored articles correspond to returns of 1.1 to 1.4 percent over the same period.[5] These findings suggest that investors perceive SA authors to be credible and, perhaps more importantly, that author positions convey information to investors. Second, we find that the price response attributed to article tone is significantly stronger for articles authored by individuals with stock positions (i.e., skin in the game). Interestingly, these effects hold for negative tone

[5] We also find that both long and short position disclosures predict absolute abnormal returns after controlling for author fixed effects, which isolate variation related to the authors' following, credibility, and other time-invariant characteristics.

across both long and short authors, though, for long authors, positive tone is not weighed more heavily. Together, these results suggest that an author's negative tone is perceived to be more credible when they have stock positions in the firms about which they write, *regardless* of whether they are long or short. However, an author's positive tone is more credible than authors with no financial positions only when they have a contrarian (i.e., short) stock position. Overall, our results suggest that investors perceive stock positions to enhance author credibility.

We conduct two sets of additional analyses. First, we perform several tests to further confirm our results are not attributable to other major events occurring concurrent to each article's publication. Namely, we restrict our sample to articles published in the early hours of the market. As we explain in Section 5.1, SA's editorial process makes it virtually impossible that early-morning articles are written about events occurring on the day of the article's publication. Results in this subsample are virtually identical to those previously discussed. Second, under the assumption that longer articles contain more analyses, we test whether the reaction to position disclosures strengthens with article length. Consistent with this expectation, we show that the positive (negative) association between long (short) position disclosures strengthens with the length of the article. Third, we examine whether the first-time disclosure of a position is more informative than subsequent disclosures, and find that the reaction to both short and long disclosures is stronger the first time an author discloses a position.[6] In our second set of analyses, we examine whether any of our results for short run returns extend to longer horizons. Unlike with short-window returns, we fail to find any relation between stock positions and long-run (60-day returns), suggesting that investor reactions surrounding the article release are complete and do not subsequently reverse. We do, however, observe some evidence that the relation between negative

[6] Repeat disclosures likely still provide information because they re-affirm an author's previous position.

tone in SA articles and future returns documented in prior research does not hold for articles written by long-position authors.

Our study provides several contributions to the finance and accounting literatures. First, we contribute to the literature on information intermediaries (including financial analysts). Sell-side financial analysts have been the subject of hundreds of academic studies because of analysts' role as sophisticated users and processers of financial data (Brown 2000; Bradshaw 2011). In his literature review, Bradshaw (2011) notes that one of the most prevalent beliefs in the capital markets is that analysts' behavior is dominated by conflicts of interest. While several papers have examined whether an analyst is more biased and less credible if their firm is also trying to generate investment banking revenues (e.g., Lin and McNichols 1999; Michaely and Womack 1999; Dechow, Hutton, and Sloan 2000; Bradshaw, Richardson and Sloan 2006), there is little evidence on the more basic question of whether an analyst's personal financial stake in the firms they follow makes their recommendations less credible. In an experimental setting, Taha and Petrocelli (2014) and Marley and Mellon (2015) find that participants consider analysts holding personal positions to be a conflict of interest, particularly when they are warned that analysts could mislead them. In contrast, we present comprehensive archival evidence finding that, on average, investors behave as if SA authors are more, not less, credible when they hold positions in the firms they follow, and that short positions appear to increase credibility more than long positions.[7]

Second, we contribute to the literature on the role of the business press in financial markets by investigating how a perceived author bias affects investors' use of information. Prior research largely examines how the financial press contributes to a firm's information environment through

[7] A contemporaneous working paper by Alfonso, Lin, and Yu (2016) uses hand-collected professional financial analyst disclosures of long positions and reaches similar conclusions as our paper. However, our study focuses on crowdsourced, social media authors, which many expect represents the future of financial analysis (e.g., Dediu 2011). Furthermore, our study is in the unique position of being able to assess the extent to which investors perceive long and short stock positions to create conflicts of interest in a setting where trading restrictions do not exist (and thus confound the inferences that can be drawn). Finally, given the absence of regulation or oversight, understanding how markets respond to crowdsourced news is interesting in its own right.

the dissemination of information (Bradshaw, Wang, and Zhou 2015; Engelberg and Parsons 2011; Fang and Peress 2009; Tetlock 2010; Tetlock 2007; Tetlock, Saar-Tsechansky, and Macskassy 2008). Collectively, this research suggests that the financial press plays an important role in the origination and dissemination of information, but does not consider whether conflicts of interest affect this process. Chen, De, Hu, and Hwang (2014) extend this research to a new information outlet—SA. Viewing SA as social media venue, they show that user-generated content and follow-up comments provide information about future stock returns and earnings surprises that is incremental to the information provided in the financial press. They conclude that, on average, SA authors provide valuable information, and do not mislead investors with their articles. We contribute to this research as the first to examine whether the disclosure of stock positions (and any accompanying perceived bias that is associated with that disclosure) affects investors' interpretation of information provided by intermediaries. SA is a unique setting in which to examine this question because, unlike the financial press, SA requires authors to disclose whether they, in fact, have personal positions in stocks about which they write.

Third, we contribute to the literature on the use of social media by market participants. In the absence of direct regulation, as with professional securities analysts, or established rules of conduct, as with financial journalists, users of social media must find alternative mechanisms for assessing the credibility of information. Chen et al. (2014) find that one of these mechanisms is an established track record of providing value relevant information. We add another: in general, investors perceive authors as more credible if they hold a position in the firm's stock. In addition, the effect of article tone on contemporaneous stock price reactions is affected by the author's stock positions.

Finally, our study should be of interest to regulators at the financial stock exchanges and the SEC.[8] Current disclosure rules require analysts to disclose whether they (or their employers) own one percent or more of the stock holdings in the firms that they follow, and prohibit analysts from trading in the stock of these firms from 30 days prior to five days after they issue a public report. Our evidence suggests that investors find disclosures regarding stock positions to be useful in interpreting the information provided by information intermediaries, that these financial positions do not impair the credibility of this information, and that, in most cases, these financial positions actually enhance the credibility of this information.

## 2. Background, Prior Literature, and Hypothesis Development

### *2.1 Social media and investor-sourced stock opinions from SeekingAlpha*

Social media allows investors to supplement information from traditional sources by communicating directly with one another. Although the method of information sharing makes a difference, with some venues (i.e., internet stock message boards) seeming to produce mostly noise and confusion (Antweiler and Frank 2004; Das and Chen 2007), recent research suggests that social media can produce and disseminate value-relevant information. Specifically, Blankespoor, Miller, and White (2014) find evidence that investors' use of Twitter reduces information asymmetry about publicly available news, while Chen et al. (2014) and Jame, Johnston, Markov, and Wolfe (2016) find evidence of social media communicating new information to the market. The latter study finds that "crowdsourced" earnings estimates are as accurate as professional analyst forecasts for some horizons, lending credence to speculation that the role of the paid, professional financial analyst might eventually become obsolete (Dediu 2011).

One of the largest social media platforms, Seeking Alpha (http://seekingalpha.com), has become a popular venue for individuals to share the results of their own analysis of financial

[8] The current analyst disclosure rules were initially developed by the New York Stock Exchange (NYSE) and the Financial Industry Regulatory Authority (FINRA), and then adopted by the SEC in 2002 (SEC 2016).

securities (i.e., "analysts" that are not necessarily paid professionals). SA is rapidly becoming one of the most referenced sources for financial news and analysis. Investopedia.com ranks it third, behind only Google Finance and Yahoo Finance, and users of "The top tens" rank SA first, ahead of both the Wall Street Journal and Financial Times. Citing Chen et al. (2014), the Wall Street Journal even speculates that SA contributors could replace professional financial analysts (Chernova 2014).[9] SA reports an average of 7 million unique visitors per month and states its mission is to provide "opinion and analysis rather than news…written by investors…rather than journalists" (Seeking Alpha 2016). They publish an average of 200 to 250 articles per day, which, with their subscription base, corresponds to 200 million email or mobile alerts going out each month. SA does not generally solicit opinions or content, but does pay contributors based on the number of users accessing their content. Chen et al. (2014) suggest that the long form of SA articles, combined with the curation of content by SA's editorial board, results in the identification of authors with something valuable to say and an opportunity for them to say it. Consistent with this suggestion, they find that the fraction of negative words in a Seeking Alpha article is negatively associated with both stock returns over the following three months and subsequent earnings surprises.

In conclusion, prior research establishes that, on average, social media in general represents an important and emerging venue for value relevant news, and SeekingAlpha articles, in particular, provide information that accurately predicts a firm's future earnings and future stock prices (Chen et al. 2014). However, while this evidence suggests that SA articles represent credible sources of information, no prior study has examined circumstances under which the contemporaneous reaction to these articles might differ cross-sectionally.

---

[9] See http://www.investopedia.com/articles/investing/112514/top-sites-latest-stock-market-news.asp and http://www.thetoptens.com/financial-news-websites/. Both sites accessed in Summer 2016.

## 2.2 *Information Intermediaries and the Credibility of their Disclosures*

Information relevant to a firm's performance comes from several types of intermediaries, such as professional financial analysts and the business press. In addition to these traditional sources, social media participants, like SA authors, represent a relatively new type of information intermediary. Each of these groups face varying incentives, which potentially impacts the credibility of any disclosed news. We next discuss these unique incentives and how they may bias disclosures.

Professional financial analysts face a number of conflicts of interest that could bias the credibility of their reports and recommendations. They are paid professionals with multiple incentives including *firm-related conflicts*, such as the generation of investment banking fees, trading commissions, trading gains/losses, etc. (e.g., Lin and McNichols 1999; Michaely and Womack 1999; Dechow, Hutton, and Sloan 2000; Bradshaw, Richardson and Sloan 2006), and well as *personal conflicts*, such as their compensation structure, long-term reputation, ingratiating themselves to managers and powerful investor groups and personal trading gains/losses (Bradshaw 2011). In other words, professional analysts face a number of conflicts, which may impair the credibility of their work.

Business press professionals face fewer financial incentives than professional analysts, but are heavily incentivized to develop their reputations by, for instance, being the first to break a big story or by doing detailed investigative work. To the business press reporter, credibility is important, and legal and professional code of ethics restrictions exist. Nevertheless, conflicts of interest can arise as media firms rely on advertising revenues from major clients, and this leads to media outlets having certain clienteles (e.g., some outlets cater to politically left-leaning viewers while other outlets cater to politically right-leaning viewers).

Investor-analysts (like those who write articles on SeekingAlpha) have their own incentives to provide private information. In a single period anonymous setting, investors might provide biased information to maximize their profits at the expense of other investors. However, in a multiperiod setting without anonymity (largely like the setting for SeekingAlpha authors), providing credible private information to others can increase an investor-analyst's investment income by accelerating price discovery for their stock positions (Pasquariello and Wang 2016). Their reputation within the investing community may also be valuable, providing opportunities to sell their analysis to others.

*2.3 Personal stock positions and credibility of published analysis*

Having a stock position in a firm about which they provide information creates a potential bias for any of the above parties. On one hand, as previously mentioned, regulators worry that stock positions of analysts could create a bias that impairs the credibility of reports and recommendations (SEC 2016). Similar scenarios likely exist with respect to the press and SA authors. On the other hand, personal stock holdings could enhance an intermediary's credibility if it implies that the individual is willing to "put their money where their mouth is."

Although prior research finds evidence that analysts are systematically optimistically biased (Bradshaw 2011), the literature is mixed as to whether stock positions are a source of that bias. Two experimental papers investigate conflicts of interest arising from analysts' personal financial positions (Taha and Petrocelli 2014; Marley and Mellon 2015). These studies suggest that analysts with stock positions are *less* credible, but the fact that both experiments represent a single-period game in a laboratory setting removes any possible effects of analyst reputation. Furthermore, participants cannot discern from personal experience whether analysts with positions are indeed more or less credible. In contrast, in actual market settings investors can usually observe an analyst's recommendation history before acting on any analysis. In this setting, Alfonso et al. (2016) find that investors find analysts *more* credible when forced to disclose long positions in the stocks they cover. However, their event-study examines the effect of a basket of regulations taking effect around 2002, which not only require analysts to disclose their positions but also impose strict rules on the timing and nature of those positions.[10] In sum, the literature on whether stock positions impact investors' perceptions of an analyst's credibility is inconclusive.

With respect to the business press, to our knowledge no prior research examines whether a journalist's stock positions affect investor reaction to the content of their articles. This lack of evidence could be attributable to the profession's legal and ethical regulations prohibiting such positions and/or the disclosure of such positions. That is, there is a lack of archival data upon which to examine such questions.

As it relates to investors, the prior archival research is limited to professional/institutional investors and finds that these investors try to restrict disclosures rather than voluntarily provide private information. For instance, Agarwal, Jiang, Tang, and Yang (2013) find that when hedge funds ask the SEC to keep their ownership levels confidential, these positions are associated with information-sensitive events and higher information asymmetry. They conclude that stock positions of hedge funds convey information about their private information. Similarly, Aragon, Hertzel, and Shi (2013) conclude that hedge fund managers seek confidentiality to protect proprietary information. In other words, the literature seems to suggest that investors try to conceal disclosures rather than voluntarily provide them and, thus, there is no evidence on the extent to which stock positions affect the credibility of an investor's *voluntary* disclosures.

In contrast, investors posting to SA are voluntarily disclosing their own analysis ("private information"). In addition, these non-insider investors generally face no trading restrictions (i.e.,

[10] Except for shares subject to trading restrictions, analysts were not required to disclose their stock positions prior to 2002. Using a sample of 43 analysts owning IPO-shares subject to such restrictions, Johnston (2013) finds no evidence of positive bias among stockholding analysts and actually finds some evidence that they are less favorable than non-owning analysts. He finds similar evidence using 41 analysts employed at a single brokerage firm who voluntarily disclosed employee positions. However, Johnston provides no evidence as to whether investors believe analysts with personal stock positions are more or less credible.

blackout periods) as explicitly exist with analysts or implicitly exist with journalists (i.e., ethical codes of conduct), and do not face conflicts of interest related to their *firms'* business interests unlike analysts (e.g., investment banking relationships or generation of commission revenues) and the business press (e.g., business relationships with other firms). These features make SA a powerful setting to uniquely assess whether other investors find the credibility of voluntary disclosure to be impaired or enhanced when the author holds a personal stock position in the firm about which the disclosure is being made.

*2.4 Reaction to financial stock position disclosures in SA articles – H1*

As previously discussed, Chen et al. (2014) find that negative tone in a SA article is associated with lower returns over the following sixty trading days. We believe their evidence implies that investors likely react to the information conveyed by SA articles *at the time* the article is published online, consistent with SA's own claim that "Seeking Alpha articles frequently move stocks" (Seeking Alpha 2016). Therefore, our first hypothesis tests whether there is a contemporaneous price reaction to the disclosure of financial position (i.e., author is long, is short, or holds no position), controlling for the tone of the article. We identify at least two reasons why the contemporaneous price reaction to an SA article would vary with disclosure of an author's stock position. First, an author's stock position is a signal about the author's belief of the company (i.e., like an analyst's buy/sell recommendation or a manager's forecast) that is incremental to the information conveyed by the article. Second, the financial position could reveal the author's private information that is not otherwise disclosed in the article. For these two reasons, we expect author stock positions to convey information that is incremental to the tone of the article such that the stock price increases for the disclosure of long positions and decreases for the disclosure of short positions. This leads to our first set of hypotheses:

**H1a**: Holding constant the tone in an article, investors respond positively to the disclosure of long positions by SA authors.

**H1b**: Holding constant the tone in an article, investors respond negatively to the disclosure of short positions by SA authors.

If stock positions introduce bias into SA articles, however, we do not expect to find support for either of the above hypotheses.

*2.5 Stock position as a moderator of investor reaction to tone – H2*

In H1, we examine whether investors react to position disclosures in SA articles. The next natural question is the one posed by regulators: is the credibility of voluntary disclosure impaired or enhanced when the discloser has a financial position? As previously discussed, if investors perceive that a stock position creates a conflict of interest, the reaction to a SA author's tone might be *weaker* (i.e., investors would assume bias and discount the tone of the information being conveyed) as compared to when the author takes no position. On the other hand, if the author has a position that is directionally consistent with the article's tone, this indicates that the author has "skin in the game" and is "putting his/her money where his/her mouth is" by acting as if s/he agrees with the information in their article. If this is the case, one might expect a *stronger* reaction to positive (negative) tone when the author has a long (short) position (i.e., there is enhanced credibility) as compared to when the author takes no position. Ultimately, whether investors perceive the tone from an author with a stock position that is "directionally-consistent" with the tone of their article as more or less credible is an empirical question. Thus, our second set of hypotheses examines whether (1) any position results in a stronger reaction to the tone of the article, and (2) whether these effects vary by long or short position. Our second hypotheses follow:

**H2a:** Investors respond more strongly to tone in SA articles authored by those with positions than by those with no positions.

**H2b**: Investors respond more strongly to negative tone in short-authored SA articles than in no-position articles

**H2c**: Investors respond more strongly to positive tone in long-authored SA articles than in no-position articles.

*2.6 Agreement between tone and stock positions – H3*

Finally, we examine whether an author's credibility is higher when expressing sentiment contrary to his or her personal interest (stock position) in the firm about which they are writing. In this case, there is no conflict of interest; in fact, the author is going out of his or her way to convey information that goes against their position.[11] If investors assign more credibility in this situation, investor reaction should be stronger when an author with a long position in a firm writes an article with a negative sentiment (compared to when the author has a short position). Similarly, investor reaction should be stronger when an author with a short position in a firm writes an article with a positive sentiment (compared to when the author has a long position). Our final hypotheses follow:

**H3a**: Investors respond more strongly to negative tone when the author holds a long position than when the author holds a short position.

**H3b**: Investors respond more strongly to positive tone when the author holds a short position than when the author holds a long position.

We test each hypothesis using short-window returns surrounding the release of each SA article, as we discuss in the next two sections. However, in additional analyses we also examine long-window returns (60 trading days) as in Chen et al. (2014) to assess whether any short-window effects persist (or reverse) in explaining future returns. We discuss these results in Section 5.2.

## 3. Data and research design

*3.1 Seeking Alpha data*

We obtain news content from Seeking Alpha by systematically downloading all content published before July 7, 2015 (the date we performed the query). To ensure that we capture new analysis provided to the markets, we download "articles" (available at seekingalpha.com/article)

[11] We do not mean to imply that short (long) authors frequently write highly positive (negative) articles. However, we measure positive and negative tone separately, consistent with investors weighing positive and negative news differently depending on expected conflicts of interest.

rather than "news" (available at seekingalpha.com/news). The former represents long-form analysis whereas the latter shorter, news-flash-like content. Table 1 describes this beginning sample and sources of data-loss. In total, we obtained 487,197 SA articles.[12] We then parsed each article to identify the article title, timestamp, referenced stocks (tickers), article content, author, and position disclosure, each clearly delineated with specific HTML tags. The header information of the SA articles identifies tickers for referenced companies in two categories, "Primary" and "About." Primary tickers are only identified when a single company is the focus of the article and analysis, and the About tickers capture other mentioned companies. In order to match articles to returns and financial statement data, we require identification of a "Primary" ticker, resulting in a loss of 280,219 articles.[13] Because our primary interest is in the price effects of author positions, we delete another 58,378 articles with no position disclosure.[14]

Disclosures generally, though not universally, follow the same basic format. At the beginning or end of each article, the author includes a statement such as "I am/we are long XXX," I/we have no positions in any stocks mentioned, and no plans to initiate any positions within the next 72 hours," or "I am/we are short XXX." However, in other instances positions are less clear, as the author may disclose complex option holdings or multiple positions in different stocks (i.e., long XXX and short YYY).[15] Therefore, we use a two-stage procedure to code these disclosures.

First, we identify long positions by searching for the terms "long," "hold," or "own

[12] SA offers a "pro" subscription that gives subscribers early access to content selected by editors. Per our discussion with the SA editor, this access lasts 24 hours after which the article is made public for 30 days. After that 30 day period, the article is archived behind a paywall and available only to pro-subscribers. Due to our sampling procedures, we did download a limited number of "pro" articles that were published in the 30 days preceding our query of SA, but these articles were not accessible when we extracted comments at a later date and are therefore excluded from our final sample. Thus, our sample is fully comprised of articles which were available to the public on the timestamp appearing in the article.

[13] Most of these deletions relate to articles that provide general commentary on industry trends and other more macro-economic events. Based on our review, articles focusing on one particular firm almost universally identify a primary ticker.

[14] Disclosure of positions were relatively rare until 2012 (no more than 20 percent each year). In more recent years, most (over 90 percent) of articles with a primary ticker designation include disclosures.

[15] For brevity, we refer to articles authored by investors with long (short) positions as "long articles" ("short articles"), and to those holding no position as "no position articles".

stock/shares." We then capture the text following those words, stopping when a period or the word "may" or "short" is encountered. The latter two words indicate the beginning of a new position disclosure (i.e., "I am long … and may…"). We repeat this procedure for possible short positions, looking for the word "short" and then capturing tickers until the word long or may, or a period. Note that for both long and short positions, we do not allow negating or qualifying words (no, not, none, neither, never, nobody, may, or plan) to occur within the five words preceding the position indicator. Finally, we search for cues that the author holds no position in any stocks. These include the terms "No Position," "None," or "May."

Inspection of results suggests these search procedures are relatively accurate, but we do encounter complex disclosures that yield multiple classifications (i.e., long, short, and/or no position) or instances where we fail to identify any of the three positions. Further, disclosure of long or positions could be in reference to stocks other than the stock about which the article is primarily written. Therefore, we further refine our disclosure coding as follows. First, in order to confirm a long or short position, we require the primary ticker of the article to match one of the tickers identified in the position disclosure. If the tickers do not match, we code the disclosure as "no position." Second, inspection of disclosures where we fail to identify long, short, or no position cues suggests these are almost universally no-position disclosures, so we code them as such. Finally, we manually inspect 370 disclosure that our procedure tagged as both long and short. Based on this inspection, we code 80 of these articles as long, 44 as short. The remaining 246 disclosures correspond to unclear positions, usually involving both equity and option positions (i.e., own stock in X and short calls in X). We drop those from our sample, leaving 148,354 coded SA articles. We next attempt to match the primary tickers to the CRSP header and history files. We lose an additional 33,585 articles for firms where we cannot identify a CRSP identifier (i.e., permno) upon which to merge. As discussed in the next section, we use a separate script to obtain reader-commentary on SA articles. We lose 90 articles for which we are unable to download the comments section of the SA articles. Finally, we lose another 559 articles that are missing any one of our basic control variables, also described in the next section. In total, we are left with 114,120 articles. In our analyses, we collapse this dataset down to 94,565 unique firm-trading day combinations of which we lose 13,976 observations missing returns for any of our windows (described next). Our final sample totals 80,589 observations.

*3.2 Descriptive information on SA authors*

Before moving to our research design and results, we first present some basic descriptive information about SA authors. We obtain this data from two sources. First, upon request, SA's Executive Editor provided us with basic demographic information they collect about the universe of SA authors. To supplement this data, we use a series of Python scripts to analyze the biographies posted on SA for authors in our sample. We present the SA provided information (the information we generated) in Panel A (Panel B) of Table 2.

Based on SA's descriptive data, 75 percent of SA authors reveal their name and place of employment, suggesting that a majority of authors face reputational concerns not only online but also in their "day jobs." As mentioned, SA pays authors earn a nominal wage based on the number of page-views. SA reports that this averages roughly $33.33 per month, per author. Thus, for most authors, monetary rewards do not appear to be the primary driver of producing high quality content. SA also reports that 30 percent of SA authors have their own independent investment blog, suggesting a substantial portion of authors invest significant time in the investment community beyond SA.[16] Taken together, these results suggests that, on average, SA authors are seeking to establish a reputation within the financial community or accelerate stock price formation

[16] SA prohibits authors from publishing the full content of their analysis in other locations. Therefore, it is unlikely articles written by authors with their own blog could be published in advance of clearing the editorial process at SA. Nonetheless, in a sensitivity analysis we exclude these observations from our sample. All of our results are quantitatively and qualitatively unchanged.

rather than to earn money directly from their SA activities.

We also independently collected biographical data for authors in our sample as of the fourth quarter of 2016. Panel B provides this information for all authors as well as statistics by position (Short, No Position, Long). We first manually code each author as an individual, a company (a private investment firm, advisor, etc.), or anonymous. Similar in Panel A, approximately 27 percent of authors use an alias while the remaining 73 percent identify themselves. These statistics are similar across positions, except that authors we identify as companies more frequently disclose no positions. We also search for certain keywords in contributor biographies and find approximately 14 percent mention “Analyst”, 9 percent mention “MBA”, and 7 percent mention “CFA”. These references are also similar across all positions, except that short-positioned authors appear twice as likely to have an MBA. Finally, SA reports “followers” for each contributor, much like FaceBook, or Twitter, and these followers are notified when contributors publish new content. In our sample, the average SA has a following of about 4,500 accounts, suggesting fairly wide dissemination of new content.

*3.3 Empirical models*

To test our hypotheses, we regress short-window abnormal returns on author position, article sentiment, and a series of controls, as presented below in [1] (*i* and *t* denote firm and year subscripts, respectively, and *[ ]* signify a multi-day range):

$$AbRet_{i,[t,t+1]} = a_0 + a_1 Long_{i,t} + a_2 Short_{i,t} + a_3 NegPct_{i,t} + a_4 PosPct_{i,t} + a_5 lWordCount_{i,t} + a_6 ComNegPct_{i,[t,t+1]} + a_7 ComPosPct_{i,[t,t+1]} + a_8 Upgrades_{i,t} + a_9 Downgrades_{i,t} + a_{10} PosES_{i,t} + a_{11} NegES_{i,t} + a_{12} Volatility_{i,t} + a_{13} AbRet_{i,[t-60,t-3]} + a_{14} AbRet_{i,t-2} + a_{15} AbRet_{i,t-1} + a_{16} DJPosPct_{i,t} + a_{17} DJNegPct_{i,t} + a_{18} IDJ_{i,t} + e_{i,t} \quad [1]$$

The dependent variable in [1] is the firm’s return measured over the two days starting on the day the article was published, adjusted by a matching size and market-to-book portfolio return over the same period. If the article was published after-hours, on a weekend, or a holiday, we begin our return window on the first trading day following the article's release.

In some cases, a stock has multiple articles written about it on the same day. If so, we collapse the SA derived data into firm-day observations to avoiding including certain observations multiple times in our regressions. For instance, we compute *Long* and *Short* as the average number of articles on a given day that disclose long and short positions, respectively.[17] To measure article tone, we sum the number of positive and negative words (from Loughran and Mcdonald 2011) in all articles corresponding to a given trading day and divide each count by the total word count across articles, yielding *PosPct* and *NegPct*, respectively.[18]

Based on H1a (H1b), we expect a positive (negative) coefficient on *Long* (*Short*). To test H2 and H3, we estimate [1] separately for four different cross-sections, which we denote with *Position*. Position equals 0 if both *Long* and *Short* equal 0 on a given trading day. Position equals -1 (1) if *Short* exceeds *Long* (*Long* exceeds *Short*) on day *t*. If *Short* and *Long* are equal and non-zero, we exclude these days from our estimations in tests of H2 and H3 (this occurs 156 times). H2a, H2b, and H2c suggest that investors respond more strongly to tone expressed by authors with "skin in the game." Specifically, H2a (H2b, H2c) predicts that the coefficient on both *NegPct* and *PosPct* exhibits stronger significance (in the expected direction) in the $|Position| = 1$ ($Position = -1$, $Position = 1$) partition relative to the $Position = 0$ partition. Finally, H3a (H3b) predicts that tone contrary (consistent) with the author's position is valued less (more). Therefore, H3a suggests a stronger coefficient on *NegPct* in the $Position = 1$ partition compared to the $Position = -1$ group, and H3b suggests the opposite for *PosPct* (stronger in $Position = -1$).

Control variables attempt to isolate other news that may affect current period returns. First, we control for the natural log of the total number of words in the article (*lWordCount*) in case there

[17] We provide detailed variable definitions in Appendix A.
[18] Following Loughran and McDonald (2011), we do not code words as positive or negative if they are preceded by a negating word (no, not, none, neither, never, or nobody).

is asymmetry in how investors respond to length. Second, we control for other "news" appearing alongside the SA article. Chen et al. (2014) find that comments following SA articles provide value relevant information. Therefore, we separately download comments for each article in our sample and code positive and negative linguistic tone for comments (*ComPosPct* and *ComNegPct*) using the same procedure as *PosPct* and *NegPct*. We restrict our comment sample to those disclosed between the date of the article and the second trading day in our return window.

Remaining variables, largely based on Chen et al. (2014), control for contemporaneous news events or for circumstances that may prompt an SA article and explain stock performance.[19] We control for analyst upgrades and downgrades (*Upgrades* and *Downgrades*) and positive and negative earnings surprises (*PosES* and *NegES*) occurring on the date of the article (or between article date and the first trading day, if different). We also measure pre-disclosure volatility (*Volatility*), which captures uncertainty in the calendar month preceding the article's release, and pre-article stock performance over three separate windows (day *t-60* to *t-3*, day *t-2*, and day *t-1*). Finally, we control for the tone of the business press using news disseminated by the Dow Jones newswire (*DJPosPct* and *DJNegPct*). We also include an indicator, *IDJ*, equaling 0 on days where there is no Dow Jones content (*DJPosPct* and *DJNegPct* set to 0 on these days).

*3.4 Descriptive Statistics*

Table 3 presents descriptive statistics for our 80,589 firm-day observations. Variables marked with "*" are scaled by 100 to facilitate presentation of descriptive statistics. Means and medians for each return metric (*AbRet*) all hover about 0, suggesting fairly symmetric return distributions. Statistics for *Short* (*Long*) suggest that approximately 2.7 (27.2) percent of articles

[19] The long-form nature of SA articles makes it unlikely that an event on day *t* leads to an article written on day *t*. Further, discussions with an editor at SA suggest that the editorial process can be lengthy—as long as 12 hours in some cases. This delay makes it unlikely that reactions to SA content reflect contemporaneously issued news. Nonetheless, we control for several aspects of contemporaneous news in our models and we conduct several additional analyses in Section 5 to rule out the alternative explanation that contemporaneous events explain our findings.

are authored by authors with a short (long) position. Our sentiment measures (*PosPct* and *NegPct*) suggest that authors use only about 1.3 (1.5) percent of negative (positive) words in articles. Chen et al. (2014) report similar statistics for negative words (they do not report statistics for positive words). We observe similar statistics for the language used in the comments section to the articles. We also find that 1.6 (3.7) percent of articles co-occur with an analyst upgrade (downgrade), and 2.7 (1.2) percent correspond to the day of a positive (negative) earnings surprise. Finally, roughly half of SA-article days have at least one other news-item as tracked by the Dow Jones newswire (*IDJ*), and the use of tone-words in these articles is relatively sparse (less than one percent for both positive and negative words).

Table 4 presents correlations among our variables. Italicized correlations are insignificantly different from zero. Consistent with H1a (H1b), we observe a significantly positive (negative) correlation of 0.06 (-0.06) between *Long* (*Short*) and $AbRet_{i,[t,t+1]}$. We also observe positive (negative) correlations between $Abret_{i,[t,t+1]}$ and both *PosPct* and *ComPosPct* (*NegPct* and *ComNegPct*), implying market movement in the direction consistent with the SA articles. The tone of Dow Jones content similarly corresponds to short-window returns in expected directions (-0.03 and 0.01 for negative and positive tone, respectively). Interestingly, few non-SA related variables relate significantly to *Short* and *Long.* We observe positive correlations between *RetVol* and both *Long* (0.03) and *Short* (0.09), implying that authors with positions more likely publish content when uncertainty is relatively high. We also observe a correlation of 0.02 between the abnormal return over the prior quarter ($AbRet_{i,[t-60,t-3]}$) and *Short*, suggesting that past news plays, at most, a minor role in these authors' decisions to publish. With respect to SA-article derived variables, we find a greater (smaller) intensity of negative (positive) words for short authors, suggesting that these authors write content consistent with their positions. Interestingly, we find that *Long* authors use fewer negative words, though not necessarily more positive words. The tone of comments tends to follow the tone of articles, and *Short* (*Long*) authors tend to incite more negative and less positive (more negative and more positive) comment sentiment. In sum, these correlations suggest that readers of SA content respond differently depending on the authors' positions.

## 4. Empirical Results

### *4.1 Test of H1*

H1a predicts that there is an investor reaction to position disclosures in SA articles at the time they are published. Specifically, H1a predicts that the disclosure of a long position (*Long*) generates a positive abnormal return in the short-window surrounding the articles release, while H1b predicts that the disclosure of a short position (*Short*) does the opposite. We report results for these hypotheses in Table 5 using equation [1].[20]

Columns 1 through 4 present results using various sets of control variables. We begin with a baseline model in column 1 that includes variables measured from SA articles themselves: *PosPct*, *NegPct*, *Long*, *Short*, *lWordCount*, *ComPosPct*, and *ComNegPct*. As presented, we find strong support for H1a and H1b as the coefficient on *Long* is significantly positive (0.44, *t*-statistic = 10.3) and the coefficient on *Short* is significantly negative (-1.39, *t* = -11.8). We also observe significant coefficients on both *PosPct* and *NegPct* (14.53, *t* = 8.0 and -22.3, *t* = -10.5, respectively). These coefficients imply a 2-day abnormal return of 0.4 (-1.4) percent attributable to the disclosure of a long (short) position, and a one-standard deviation increase in *PosPct* (*NegPct*) corresponds to a return of 0.1 (-0.2) percent. We next introduce controls for non-Dow Jones related news content (column 2), and DJ content (column 3) and continue to find strong support for H1a and H1b. Coefficients on *Long* all hover around 0.4 (*t > 10.0*) while coefficients on *Short* center around -1.1 (*t* > 8.7). Coefficients on *NegPct* and *PosPct* also exhibit similar

[20] In all tables, we multiply the dependent variable by 100 to facilitate presentation of coefficient estimates. All specifications also include year-month fixed effects.

magnitudes to column 1. In columns 4 and 5, we re-estimate our full model (corresponding to column 3) after removing observations with Dow Jones content (*IDJ* = 1) and after removing observations with DJ content and earnings surprises (*IDJ = 1* or *PosES = 1* or *NegES = 1*). All results continue to hold and the coefficients on *Long* and *Short* exhibit slightly larger magnitudes compared to the column 4 estimates. Overall, the results presented in Table 5 provide strong support for H1a and H1b, suggesting that stock positions convey information about the author's overall opinion of the firm and that investors perceive SA authors to be credible.

One immediate concern related to these results is that author positions correlate with some unobserved author characteristic, which drives the results found in support for H1a and H1b. To address this possibility, we repeat our tests including author fixed effects to control for unobserved author characteristics that may correlate with stock positions. Specifically, we treat each article as an observation (rather than firm-day) and regress $|AbRet_{i,[t,t+1]}|$ on indicators identifying long and short authors, all other variables in [1], and author fixed effects. We observe significantly positive coefficients (untabulated) on both of these indicator variables, further supporting H1a and H1b.

*4.2 Test of H2*

H2a predicts that investors respond more strongly to the tone (*PosPct* and *NegPct*) of articles authored by those with stock positions compared to those with no position. H2b (H2c) similarly predicts a stronger reaction for long-authored (short-authored) articles. We use *Position* as a partitioning variable and estimate equation [1] using four separate subsamples: (1) *Position = 0*, (2) *|Position| = 1*, (3) *Position = -1*, and (4) *Position = 1*. Table 6 presents these results. We use columns 1 and 2 to test H2a, columns 1 and 3 to test H2b, and columns 1 and 4 to test H2c. As shown, the coefficient on *NegPct* in column 2 (-28.2) exceeds that in column 1 (-20.8), and this difference is highly significant ($p < 0.01$). We observe a similar pattern for *PosPct*. The column 2 coefficient (16.8) significantly exceeds the column 1 value (12.7, difference significant at $p = 0.01$

level). Thus, we find strong support for H2a, suggesting that investors respond more strongly to the tone in an article when the author has a stock position in the firm about which they are writing compared to when they have no such position. That is, investors perceive SA authors to be more credible when they have "skin in the game."

Moving to H2b, we compare reactions to short-authored articles (*Position = -1*) and articles written by authors with no position (*Position = 0*). As with H2a, we observe coefficients in column 3 that are significantly larger in magnitude for both *NegPct* (-39.2) and *PosPct* (67.3) than their column 1 counterparts. The difference in *PosPct* coefficients is highly significant ($p = 0.01$) and the difference in *NegPct* coefficients is significant ($p = 0.04$). For H2c, we compare column 4 to column 1. Column 4 magnitudes on *NegPct* and *PosPct* again exceed their column 1 counterparts. However, only the difference in *NegPct* (-28.5 vs. -20.8) is significant ($p = 0.04$). This is again consistent with investors interpreting position-consistent tone as failing to enhance credibility, at least when authors have long positions (though, even then, it does not appear to impair credibility).

Overall, we find evidence that article tone by authors with skin-in-the-game is valued more significantly than those without (H2a). Similarly, short-authors appear to be more credible on both positive and negative tone (H2b), though long authors' credibility is only enhanced for position-inconsistent, negative tone. We next compare tone by long vs. short authors to test H3.

*4.3 Test of H3*

Our final set of hypotheses, H3a and H3b, tests whether investors perceive negative (positive) bias in the tone of short (long) SA authors. H3a suggests that negative tone (*NegPct*) is more highly valued for long-authored articles than short-authored articles, consistent with authors being most credible when they write articles with tone that is opposite to their stock positions. Similarly, H3b suggests that positive tone (*PosPct*) generates a larger response when authors are short compared to when authors are long. Results for these hypotheses are also presented in Table

6. We focus on tests of differences between columns 3 (short-authors) and 4 (long-authors).

We do not find support for H3a, as there is no difference between the coefficient for long-authored negative tone and the coefficient for short-authored negative tone ($p = 0.33$). One possible explanation for this result is that readers of SA content consider short-authors to be more sophisticated and knowledgeable, and this knowledge offsets any perceived bias compared to long authors. We do, however, find support for H3b. The association between *PosPct* and $AbRet_{i,[t,t+1]}$ in column 3 (67.3, $t = 3.0$) is significantly larger than in column 4 (14.0, $t = 3.2$; difference significant at $p = 0.01$ level).[21] Thus, positive tone by authors with a long-position does not appear to enhance credibility (but, in any case, it does not appear to impair credibility consistent with a perceived conflict of interest).[22]

To summarize, we find evidence that investors impound information generated by other investors into price. Perhaps more importantly, we also find consistent evidence that knowledge of an author's position has two effects on how information is impounded in stock price. First, the disclosure itself serves as a price-relevant signal to investors. Second, knowledge of the author's position moderates sentiment pricing. Compared to authors with no-position, the pricing of article sentiment is stronger when an author has a skin in the game. Further, positive-tone in long-authored articles appears to be discounted compared to short-authored articles. Finally, investors appear to find authors with short positions that write articles with a positive tone (i.e., short authors that write with a tone that is opposite to their underlying stock position) to be the most credible of all. In the

[21] An alternative means of testing H3 is to compare the difference in difference in coefficient magnitudes between columns 3 and 4. In column 3, the magnitude of the coefficient on *PosPct* (67.3, labeled *a*) exceeds the coefficient on *NegPct* (39.2, labeled *b*). Conversely, in column 4 the pattern reverses, as the magnitude of the coefficient on *PosPct* (14.0, labeled *c*) is less than *NegPct* (28.5, labeled *d*). The difference in differences (i.e., *a-b* compared to *c-d*) is significant ($p = 0.05$). This further supports the notion tone contrary to one's position is viewed as more credible.

[22] Recall that we collapse article level data to the firm-day level. This results in aggregating tone across multiple articles (and thus multiple authors). While this correctly represents the body of work published by SA about a given firm on a given day, it adds noise to our partitioning procedure. To ensure this does not impact of results, we repeat tests of H1, H2, and H3 after limiting the sample to days on which only a single article is published. Results are generally unchanged, except that the difference in *NegPct* coefficients in columns 1 and 3 of Table 6 declines in significance ($p=0.13$).

following section, we conduct several additional analyses to expand upon this evidence.

## 5. Additional analysis

*5.1 Alternative explanations: Contemporaneous events leading to observed stock price reactions*

One possible concern with our results is that our variables of interest capture some other article attribute or contemporaneous event. We believe this to be unlikely because the long-form of SA articles, which frequently include extensive tables, charts, and links to detailed analysis, makes it unlikely that an author observes an event and immediately produces an article. In addition, SA articles undergo an editorial process which, per our discussions with an SA executive, can take up to 12 hours. Finally, with respect to H2 and H3, contemporaneous events would need to correlate with not only the content of SA articles, but also the position of the author writing the article, and an inspection of SA articles reveals no systematic differences in articles authored by positioned authors (*Short* or *Long* equaling 1) compared to no-position authors. Nonetheless, in this section we perform several additional tests designed to mitigate the likelihood that our results are due to contemporaneous firm economic events.

*5.1.1 Controlling for firm news around the SA article release date and time*

As just mentioned, an alternative explanation for our results is that the stock returns we observe are not a reaction to the SA article but are instead a reaction to contemporaneous firm news event. In our main tests, we attempt to rule out this explanation by including a number of control variables, including the tone of business press articles issued around the SA article. To further mitigate this issue, we take advantage of the fact that the SA editorial process take about 4.5 hours to complete, and assume that any article published on SA before 1 pm (i.e., the first 3.5 trading hours of the day) must have been submitted to SA before the market opens for the day. We then redefine the day 0 return as (closing price – opening price) / opening price (all measured on day 0). Therefore, any overnight news impounded into opening price is excluded from our returns.

Using this revised measure of returns and limiting the sample to articles posted in the first few hours of trading, we repeat our analyses from Table 5 and present results in Table 7. As shown, we continue to find all results, and the economic significance of these effects is relatively unchanged. These results add further assurance that our stock price results are not explained by contemporaneous news events and are, instead, a reaction to the SA articles themselves.

*5.1.2 Interactions between length and position*

While we contend that disclosure of position provides a value relevant signal to market participants, we recognize that simply saying "I am long…" without any other support would likely garner little reaction. Therefore, we expect that the reaction to author positions increases with the amount of information presented alongside their position-disclosure. To test this conjecture, we use article length (*lWordCount*) as a proxy for total information and interact this variable with both *Long* and *Short* in equation [1]. We expect that each interaction to load in the same direction as the position (i.e., positive for *Long*lWordCount* and negative for *Short*lWordCount*).[23]

Results from these analyses are presented in Table 8. We include the same subsamples as in earlier analyses (all observations, excluding DJ, excluding DJ and earnings surprises). In all three columns, coefficients on *Long* and *Short* remain significant (*t*-statistics between 6.3 and 10.1), and magnitudes are similar to those shown in Table 5. Most importantly, we observe highly significant coefficients on interactions in the expected direction. The interaction between *Long* and *lWordCount* is significantly positive (*t*-statistics between 2.6 and 3), and the interaction between *Short* and *lWordCount* is significantly negative (*t*-statistics between -3.4 and -4.2). Thus, pricing of author positions appears to increase with article length, as expected, which further corroborates our expectation that author positions themselves convey meaningful information to readers.

[23] We center *lWordCount* about 0 for this analysis to maintain interpretability of coefficients on *Long* and *Short*. In other words, main effects on *Long* and *Short* represent the response to position for the average-length article.

*5.1.3 First-time vs. Repeated Disclosures*

To further support that author position, and not some correlated omitted variable, explains the reaction to SA content, we next examine whether the reaction to position is stronger in the author's first article publishing his or her stock position. Specifically, we sort our sample of SA content by firm (i.e., primary ticker), author, and date, and identify the first time an author discloses a position about the subject firm. We denote this article using an indicator variable, *FirstDisc.* We then estimate equation [1], including interactions between *FirstDisc* and both *Long* and *Short.* We expect that the relations between *Short* and *Long* are stronger when *FirstDisc* equals 1.[24] In addition, we include interactions between *FirstDisc* and other article attributes (*lWordCount*, *NegPct*, *PosPct*, *ComNegPct*, and *ComPosPct*) because relations between those variables and returns may vary depending on how often the author writes about a given firm. However, we make no predictions related to these interactions.

Table 9 presents results from this analysis. For brevity, we only include coefficients on SA-related variables (those interacted with *FirstDisc*) and suppress tabulation of other coefficients. As in Table 9, we include the same subsamples used throughout the paper. Consistent with expectations, we observe a highly significant, negative coefficient on the interaction between *FirstDisc* and *Short* (*t*-statistics between -4.2 and -6.1) and a significantly positive coefficient on the interaction between *FirstDisc* and *Long* (*t*-statistics between 3.2 and 3.5). This evidence provides further confirmation that disclosures themselves provide value relevant information to investors. As for other interactions, we observe some evidence that investors price tone (*NegPct* and *PosPct*) *less* the first time an author writes a positioned-article, consistent with credibility

[24] One may argue that the first-time disclosure of a position should be the only time this knowledge matters. However, multiple articles disclosing the same position affirm the author's beliefs over time, thus providing additional positive signals. We also considered instances where authors changed positions, but found these events occur relatively infrequently. For example, if an author sells a stock in which he or she is long, it appears that author often ceases to write about that firm.

being gained over time. However, these interactions are only significant in column 1.

*5.2 SA content and long-run returns*

As discussed, Chen et al. (2014) document a significantly negative return between the percentage of negative words in SA articles and returns over the subsequent quarter (approximately 60 trading days). In our main tests, we focus on contemporaneous pricing of SA content and how position affects this pricing under the assumption that markets are generally efficient. We now repeat these analyses (Tables 5 and 6) replacing our short window return ($AbRet_{i,[t,t+1]}$) with the post-event return over the following 60 days ($AbRet_{i,[t+3,t+60]}$), as in Chen et al. (2014).

Table 10 replicates Table 5 using the 60-day return. We use all controls from [1] and add $AbRet_{i,[t,t+1]}$ and tone from comments over the duration of the full return window. We include the same five columns as in Table 5. Like Chen et al. (2014), we observe strong, negative associations between *NegPct* and post-event returns, consistent with SA content providing value relevant information that is not immediately impounded in stock prices. *PosPct* exhibits similar positive associations, but these exhibit much lower statistical significance (*t*-statistics between 1.4 and 1.8). Results in column 1 indicate that returns in the quarter following short-authored (*Short*) articles tend to be abnormally negative ($t$ = -2.5). Interestingly, the coefficient on *Long* is negative, though statistically insignificant ($t$ = -1.3). Columns 2 through 4 introduce our control variables, which subsume the position-disclosure effects observed in column 1. Thus, unlike in Table 5, position itself does not provide an indication of long-run performance. Columns 5 and 6 drop observations with potentially confounding contemporaneous events, and we again fail to observe any relation between *Short* or *Long* and $Abret_{i,[t+3,t+60]}$.

Table 11 replicates Table 6, again replacing the short-window return corresponding to the SA article's release with the abnormal return over the following quarter. Like in Table 6, we focus on differences in coefficients on *NegPct* and *PosPct* across the position-partitions. For *NegPct*, we observe an interesting pattern of results. Having a position does not appear to influence the predictive power of negatively toned SA content. We observe negative coefficients in columns 1 and 2 ($t$ = -2.3 and $t$ = -1.8, respectively). The magnitude of the effect is larger in column 2, but these differences are insignificant ($p$ = 0.5). Column 3 reports results for short-authored articles. Here, the magnitude of the coefficient on *NegPct* (-116.3) is three to four times that of the column 1 (-29.6) and 2 (-35.6) effects. Further, the difference between columns (1) and (3) is marginally significant ($p$ = 0.09), consistent with short-authors providing information that is more value-relevant than other authors. In contrast, column 4 reveals that negative sentiment by long authors exhibits no significant association with future returns ($t$ = -1.0). This coefficient is significantly different than in column 3 ($p$ = 0.09), though not column 1 ($p$ = 0.7). In sum, our results suggest that most of the predictive power for long-run returns in SA articles is focused on no-position and short-authored articles.

## 6. Conclusion

Motivated by concerns that an information intermediary's stock positions in firms they follow present a conflict of interest that impairs their objectivity, we examine investor perceptions of the stock positions of investor-authored articles on the social media outlet SeekingAlpha (www.seekingalpha.com). We use textual analysis to extract the information conveyed in 114,120 SA articles that cover 4,016 unique firms released on 80,589 unique firm-days from 2006 to 2015, and offer three main findings. First, author positions contribute directly to short-window returns surrounding the article's publication, holding constant the tone of the article, article comments, and contemporaneously issued news. Economically, long-authored articles correspond to a positive 2-day return of 0.4 of 0.5 percent, while short-authored articles correspond to returns of 1.1 to 1.4 percent over the same period. These findings suggest that author positions convey information to investors, and that investors perceive SA authors to be credible. Second, we find that the price response attributed to article tone is significantly stronger for articles authored by individuals with stock positions (i.e., skin in the game). Interestingly, these effects hold for negative tone across both long and short authors. For long authors, positive tone is not weighted more heavily. Together, these results suggest that investors perceive authors with short positions to be credible *regardless* of their tone, but that authors with long positions are *only* credible when they write articles with negative tone. Finally, we find that the response to positive tone is stronger for short authors compared to long authors. These results suggest that investors find authors who write articles with a tone that is opposite to their underlying stock position to be most credible of all. Overall, our results suggest that investors perceive an author's stock positions to enhance their credibility, as these authors "put their money where their mouth is" and have skin in the game.

Three additional analyses bolster our findings. First, we show that limiting our sample to a subset of articles which are unlikely to be affected by contemporaneous news has no effect on our results. Second, we suspect that position disclosures are more informative when the author provides more information relevant to the firm, increasing his or her credibility. Consistent with this conjecture, we show that the positive (negative) association between long (short) position disclosures strengthen with article length. Thus, position disclosures matter more when accompanied by more analysis. Finally, we suspect that first-time disclosures of positions garner greater responses because the information is "new." Consistent with this expectation, we show that reaction to both short and long disclosures is significantly stronger the first time an author discloses a position.

In our final analysis, we examine whether any of our results exhibit drift or reversal over the 60 days following publication. We fail to find this is the case, though we provide some evidence that evidence in Chen et al. (2014) is mostly focused in articles authored by individuals disclosing either a short position or no position.

## References

Agarwal, V., W. Jiang, Y. Tang, and B. Yang. 2013. "Uncovering Hedge Fund Skill from the Portfolio Holdings They Hide." *The Journal of Finance* 68 (2): 739–83. doi:10.1111/jofi.12012.

Alfonso, E., S. Lin, and Y. Yu. 2016. "Analyst Stock Ownership and Stock Recommendations." Working paper, Florida International University and the University of Texas at Austin.

Antweiler, W., and M. Z. Frank. 2004. "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards." *The Journal of Finance* 59 (3): 1259–94. doi:10.2307/3694736.

Aragon, G. O., M. Hertzel, and Z. Shi. 2013. "Why Do Hedge Funds Avoid Disclosure? Evidence from Confidential 13F Filings." *Journal of Financial and Quantitative Analysis* 48 (05): 1499–1518. doi:10.1017/S0022109013000483.

Blankespoor, E., G. S. Miller, and H. D. White. 2014. "The Role of Dissemination in Market Liquidity: Evidence from Firms' Use of Twitter™." *The Accounting Review* 89 (1): 79–112. doi:10.2308/accr-50576.

Brown, L., ed. 2000. *I/B/E/S Research Bibliography: Sixth Edition*. I/B/E/S International Inc., New York, NY.

Bradshaw, M. 2011. Analysts' Forecasts: What Do We Know after Decades of Work? Working paper, Boston College.

Bradshaw, M., S. Richardson, and R. Sloan. 2006. The relation between corporate financing activities, analysts' forecasts and stock returns. *Journal of Accounting & Economics* 42 (1-2): 53-85.

Bradshaw, M. T., Xue Wang, and D. Zhou. 2015. "Analysts' Assimilation of Soft Information in the Financial Press." Working Paper, Boston College.

Chen, H., P. De, Y. Hu, and B.-H. Hwang. 2014. "Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media." *Review of Financial Studies* 27 (5): 1367–1403. doi:10.1093/rfs/hhu001.

Chernova, Y. 2014. "Study: Crowdsourced Stock Opinions Beat Analysts, News -- WSJ Blog." *Dow Jones Institutional News*, March 19.

Das, S. R., and M. Y. Chen. 2007. "Yahoo! For Amazon: Sentiment Extraction from Small Talk on the Web." *Management Science* 53 (9): 1375–88. doi:10.1287/mnsc.1070.0704.

Dechow, P., A. Hutton, and R. Sloan. 2000. The relation between analysts' forecasts of long-term earnings growth and stock price performance following equity offerings. *Contemporary Accounting Research* 17 (1): 1-32.

Dediu, H. 2011. "A New Era in Financial Analysis Is Dawning." *Asymco*. January 19. http://www.asymco.com/2011/01/19/an-new-era-in-financial-analysis-is-dawning/.

Engelberg, J. E., and C. Parsons. 2011. "The Causal Impact of Media in Financial Markets." *The Journal of Finance* 66 (1): 67–97.

Fang, L., and J. Peress. 2009. "Media Coverage and the Cross-Section of Stock Returns." *The Journal of Finance* 64 (5): 2023–52.

Jame, R., R. Johnston, S. Markov, and M. C. Wolfe. 2016. "The Value of Crowdsourced Earnings Forecasts." *Journal of Accounting Research* 54 (4): 1077–1110. doi:10.1111/1475-679X.12121.

Johnston, R. 2013. "Does Analyst Stock Ownership Affect Reporting Behavior?" *Review of Pacific Basin Financial Markets and Policies* 16 (2). doi:10.1142/S0219091513500082.

Lin, H., and M. McNichols. 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics* 25: 101-127.

Loughran, T., and B. Mcdonald. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *The Journal of Finance* 66 (1): 35–65. doi:10.1111/j.1540-6261.2010.01625.x.

Marley, R., and M. Mellon. 2015. "The Effects of Current and Expanded Analyst Ownership Disclosure on Nonprofessional Investors' Judgments and Decision-Making." Working paper, University of Tampa and the University of South Florida.

Michaely, R., and K. Womack. 1999. Conflict of interest and the credibility of underwriter analyst recommendations. *The Review of Financial Studies* 12 (4): 653-686.

Pasquariello, P., and Y. Wang. 2016. Speculation with Information Disclosure. Working paper, University of Michigan.

Securities and Exchange Commission (SEC). 2016. Analyzing Analyst Recommendations. Investor Publications. Accessed on September 2, 2016 at: https://www.sec.gov/investor/pubs/analysts.htm

Seeking Alpha. 2016. About seeking alpha. Accessed on September 2, 2016 at: http://seekingalpha.com/page/about_us

Taha, A., and J. Petrocelli. 2014. "Sending mixed messages: Investor interpretations of disclosures of analyst stock ownership." *Psychology, Public Policy, and Law* 20(1): 68-77.

Tetlock, P. C. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." *The Journal of Finance* 62 (3): 1139–68. doi:10.1111/j.1540-6261.2007.01232.x.

———. 2010. "Does Public Financial News Resolve Asymmetric Information?" *Review of Financial Studies* 23 (9): 3520–57. doi:10.1093/rfs/hhq052.

Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy. 2008. "More than Words: Quantifying Language to Measure Firms' Fundamentals." *The Journal of Finance* 63 (3): 1437–67.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48:817–38.

## Appendix A: Variable Definitions

| Variable | Definition |
|---|---|
| $AbRet_{i,t}$ | The firm's return measured on day *t* or over days [*t* to *t+k*] adjusted by a matching size and market-to-book portfolio return over the same period. If the article was published after-hours, on a weekend, or holiday, day *t* equals the first trading day following the article's release. (winsorized) |
| *Long* | The percentage of articles about firm *i* on day *t* in which the author discloses a long position. |
| *Short* | The percentage of articles about firm *i* on day *t* in which the author discloses a short position. |
| *Position* | Takes value -1 if *Short* exceeds *Long* on a given day, 1 if *Long* exceeds *Short* on a given day, and 0 if *Long* and *Short* both equal 0. On days where *Long* = *Short* and both *Long* and *Short* are non-zero, *Position* is undefined. |
| *NegPct* | The percentage of non-negated words for all SA articles on day *t* that are classified as having negative sentiment by Loughran and McDonald (2011). (winsorized) |
| *PosPct* | The percentage of non-negated words for all SA articles on day *t* that are classified as having positive sentiment by Loughran and McDonald (2011). (winsorized) |
| *lWordCount* | The natural log of the total number of words appearing in SA articles about a firm on a given day. (winsorized) |
| $ComNegPct_{i,t}$ | The percentage of non-negated words appearing in comments posted between day *t* and *t+k* about the SA article classified as having negative sentiment by Loughran and McDonald (2011). |
| $ComPosPct_{i,t}$ | The percentage of non-negated words appearing in comments posted between day *t* and *t+k* about the SA article classified as having positive sentiment by Loughran and McDonald (2011). |
| *Upgrades* | The number of analysts revising recommendations upward on day *t*, or in the days between article publication and first trading day if different. (winsorized) |
| *Downgrades* | The number of analysts revising recommendations downward on day *t*, or in the days between article publication and first trading day if different. (winsorized) |
| *PosES* | Indicator equaling 1 if the firm announces earnings exceeding the most recent consensus estimate according to IBES on day *t*, or in the days between article publication and first trading day if different. (winsorized) |
| *NegES* | Indicator equaling 1 if the firm announces earnings below the most recent consensus estimate according to IBES on day *t*, or in the days between article publication and first trading day if different. (winsorized) |
| *Volatility* | The sum of squared daily returns in the calendar month preceding day *t*. (winsorized) |

| | |
|---|---|
| $DJNegPct_{i,t}$ | The percentage of non-negated words in all Dow Jones news content published on day *t*, or in the days between article publication and first trading day if different, classified as having positive sentiment by Loughran and McDonald (2011). (winsorized) |
| $DJPosPct_{i,t}$ | The percentage of non-negated words in all Dow Jones news content published on day *t*, or in the days between article publication and first trading day if different, classified as having negative sentiment by Loughran and McDonald (2011). (winsorized) |
| *IDJ* | An indicator equaling 1 if there is no Dow Jones content about the firm published on day *t*, or in the days between article publication and first trading day if different. |
| *FirstDisc* | Indicator taking value of 1 the first time an author discloses a given position about a firm. |

## Table 1: Sample Attrition

| | |
|---|---|
| Seeking Alpha Articles Downloaded as of July 7, 2015 | 487,197 |
| Articles missing primary ticker designation | (280,219) |
| Articles missing position disclosure | (58,378) |
| Articles with ambiguous position disclosure | (246) |
| Articles with Successfully Coded Disclosures | 148,354 |
| Articles not linked to CRSP | (33,585) |
| Articles missing comments | (90) |
| Articles missing other controls | (559) |
| Total Articles in Sample | 114,120 |
| Unique Firm-Trading Day Combinations | 94,565 |
| Missing returns for any period | (13,976) |
| Starting Sample for Analyses | 80,589 |

**Table 2: Author Characteristics**

*Panel A: Author characteristics as described by SeekingAlpha*

| | # of Authors | % of Total |
|---|---|---|
| Total Authors | 13,680 | |
| Authors with independent blogs | 3,711 | 27.13% |
| Anonymous Authors | 3,358 | 24.55% |
| Financial Professionals | 4,891 | 35.75% |
| Students (Young Investors) | 1,056 | 7.72% |
| Company Executives/C-Level | 748 | 5.47% |
| | | |
| Monthly Average Payment to Authors in 2016 | $33.30 | |

*Panel B: Author characteristics harvested from biographies*

| Characteristic | Total (n=111,372) | Position = -1 | Position = 0 | Position = 1 |
|---|---|---|---|---|
| Authored by individual | 57.27% | 60.86% | 54.16% | 64.61% |
| Authored by company | 16.05% | 9.05% | 19.24% | 8.81% |
| Anonymously authored (alias) | 26.69% | 30.08% | 26.59% | 26.57% |
| | 100% | 100% | 100% | 100% |
| | | | | |
| Includes "Analyst" in bio | 14.35% | 16.19% | 14.58% | 13.61% |
| References blog or website (other than LinkedIn) | 41.93% | 41.13% | 42.10% | 41.59% |
| Mentions “MBA” in bio | 8.83% | 17.06% | 8.79% | 8.12% |
| Mentions “CFA” in bio | 6.71% | 7.19% | 7.35% | 5.06% |
| | | | | |
| Followers at time author page was downloaded (mean) | 4,486 | 3,385 | 4,494 | 4,532 |

Table 2 reports descriptive statistics for author characteristics. Panel A reports information provided by SeekingAlpha and Panel B reports information pertaining to 111,372 of our 114,120 sample articles. For Panel B, we downloaded each author’s bio from SeekingAlpha (http://seekingalpha.com/author/...) and used hand-coding or textual analysis to harvest select information. We manually coded each author name as an individual, a company, or an alias. Remaining information was systematically extracted from each author page.

## Table 3: Descriptive Statistics

| | | | | | | | *POSITION=-1* | | *POSITION=0* | | *POSITION=1* | | *Tests of Differences* | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Variable* | n | Mean | Std. Dev | 25% | 50% | 75% | Mean | 50% | Mean | 50% | Mean | 50% | -1 vs. 0 | -1 vs 1 | 0 vs. 1 |
| $AbRet_{i,[t,t+1]}$* | 80,589 | 0.070 | 3.895 | -1.430 | -0.016 | 1.430 | -1.270 | -0.706 | -0.007 | -0.051 | 0.372 | 0.107 | 0.000 | 0.000 | 0.000 |
| $AbRet_{i,[t+3,t+60]}$* | 80,589 | -1.417 | 17.011 | -9.924 | -1.227 | 6.840 | -2.508 | -1.900 | -1.360 | -1.170 | -1.427 | -1.309 | 0.011 | 0.018 | 0.614 |
| $AbRet_{i,[t-2]}$* | 80,589 | 0.011 | 2.621 | -0.978 | -0.036 | 0.940 | -0.068 | -0.179 | 0.019 | -0.033 | 0.000 | -0.034 | 0.244 | 0.370 | 0.359 |
| $AbRet_{i,[t-1]}$* | 80,589 | 0.030 | 3.056 | -1.011 | -0.029 | 0.984 | -0.172 | -0.207 | 0.023 | -0.027 | 0.070 | -0.017 | 0.023 | 0.006 | 0.055 |
| $AbRet_{i,[t-60,t-3]}$* | 80,589 | -0.152 | 19.207 | -9.831 | -1.006 | 7.828 | 2.274 | 0.041 | -0.324 | -0.854 | -0.014 | -1.406 | 0.000 | 0.000 | 0.042 |
| *Short* | 80,589 | 0.027 | 0.156 | 0.000 | 0.000 | 0.000 | 0.896 | 1.000 | 0.000 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| *Long* | 80,589 | 0.272 | 0.430 | 0.000 | 0.000 | 0.667 | 0.002 | 0.000 | 0.000 | 0.000 | 0.907 | 1.000 | 0.000 | 0.000 | 0.000 |
| *Position* | 80,433 | 0.270 | 0.505 | 0.000 | 0.000 | 1.000 | | | | | | | | | |
| *NegPct** | 80,589 | 1.329 | 0.872 | 0.688 | 1.176 | 1.791 | 2.083 | 1.935 | 1.322 | 1.157 | 1.270 | 1.154 | 0.000 | 0.000 | 0.000 |
| *PosPct** | 80,589 | 1.467 | 0.746 | 0.938 | 1.378 | 1.903 | 1.015 | 0.937 | 1.489 | 1.394 | 1.464 | 1.390 | 0.000 | 0.000 | 0.000 |
| *lWordCount* | 80,589 | 6.744 | 0.631 | 6.349 | 6.727 | 7.115 | 7.066 | 7.069 | 6.655 | 6.676 | 6.905 | 6.856 | 0.000 | 0.000 | 0.000 |
| $ComPosPct_{i,[t,t+1]}$* | 80,589 | 1.035 | 1.196 | 0.000 | 0.915 | 1.529 | 0.897 | 0.910 | 0.978 | 0.741 | 1.176 | 1.130 | 0.000 | 0.000 | 0.000 |
| $ComNegPct_{i,[t,t+1]}$* | 80,589 | 1.009 | 1.082 | 0.000 | 0.912 | 1.639 | 1.502 | 1.566 | 0.943 | 0.671 | 1.105 | 1.125 | 0.000 | 0.000 | 0.000 |
| $ComPosPct_{i,[t+3,t+60]}$* | 80,589 | 0.597 | 1.210 | 0.000 | 0.000 | 0.884 | 0.706 | 0.431 | 0.523 | 0.000 | 0.751 | 0.000 | 0.000 | 0.042 | 0.000 |
| $ComNegPct_{i,[t+3,t+60]}$* | 80,589 | 0.600 | 1.112 | 0.000 | 0.000 | 0.985 | 1.180 | 0.980 | 0.506 | 0.000 | 0.750 | 0.000 | 0.000 | 0.000 | 0.000 |
| *Upgrades* | 80,589 | 0.016 | 0.123 | 0.000 | 0.000 | 0.000 | 0.020 | 0.000 | 0.016 | 0.000 | 0.015 | 0.000 | 0.107 | 0.072 | 0.498 |
| *Downgrades* | 80,589 | 0.037 | 0.187 | 0.000 | 0.000 | 0.000 | 0.044 | 0.000 | 0.038 | 0.000 | 0.033 | 0.000 | 0.191 | 0.017 | 0.001 |
| *PosES* | 80,589 | 0.026 | 0.158 | 0.000 | 0.000 | 0.000 | 0.027 | 0.000 | 0.029 | 0.000 | 0.020 | 0.000 | 0.556 | 0.056 | 0.000 |
| *NegES* | 80,589 | 0.011 | 0.104 | 0.000 | 0.000 | 0.000 | 0.011 | 0.000 | 0.012 | 0.000 | 0.010 | 0.000 | 0.511 | 0.685 | 0.003 |
| *Volatility* | 80,589 | 0.019 | 0.033 | 0.003 | 0.008 | 0.019 | 0.036 | 0.017 | 0.017 | 0.007 | 0.020 | 0.009 | 0.000 | 0.000 | 0.000 |
| *DJPosPct** | 80,589 | 0.565 | 0.930 | 0.000 | 0.000 | 0.977 | 0.580 | 0.000 | 0.566 | 0.000 | 0.559 | 0.000 | 0.494 | 0.299 | 0.289 |
| *DJNegPct** | 80,589 | 0.442 | 0.684 | 0.000 | 0.000 | 0.856 | 0.448 | 0.000 | 0.442 | 0.000 | 0.437 | 0.000 | 0.692 | 0.445 | 0.288 |
| *IDJ* | 80,589 | 0.489 | 0.499 | 0.000 | 0.000 | 1.000 | 0.479 | 0.000 | 0.488 | 0.000 | 0.490 | 0.000 | 0.358 | 0.294 | 0.667 |

Table 3 presents descriptive statistics for primary variables used this study. Each observation represents a unique firm-trading day combination. All variables are defined in Appendix A. Variables marked with "*" are multiplied by 100 for presentation purposes.

## Table 4: Correlation Matrix

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) | (17) | (18) | (19) | (20) | (21) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) $AbRet_{i,[t,t+1]}$ | | 0.01 | *-0.01* | -0.01 | *0.00* | -0.04 | 0.04 | -0.06 | 0.04 | 0.01 | -0.03 | *-0.01* | -0.02 | *0.00* | 0.01 | 0.04 | -0.05 | -0.01 | *-0.01* | 0.01 | *0.00* |
| (2) $AbRet_{i,[t+3,t+60]}$ | 0.01 | | 0.01 | 0.01 | 0.02 | -0.01 | *-0.01* | -0.02 | 0.02 | -0.02 | -0.03 | -0.03 | -0.04 | *0.00* | *0.00* | 0.01 | *-0.01* | -0.08 | 0.04 | 0.04 | 0.04 |
| (3) $AbRet_{i,[t-2]}$ | *0.00* | 0.01 | | 0.01 | 0.02 | -0.01 | *0.00* | -0.05 | 0.04 | *0.00* | -0.03 | -0.01 | -0.02 | *0.00* | *0.00* | 0.01 | *0.00* | -0.02 | -0.01 | *0.00* | *0.00* |
| (4) $AbRet_{i,[t-1]}$ | *-0.01* | *0.00* | 0.02 | | 0.01 | -0.02 | *0.01* | -0.07 | 0.04 | -0.01 | -0.04 | -0.02 | -0.02 | *0.00* | *-0.01* | 0.02 | -0.01 | -0.02 | -0.01 | *0.00* | *0.00* |
| (5) $AbRet_{i,[t-60,t-3]}$ | -0.01 | 0.01 | 0.02 | 0.01 | | 0.01 | -0.01 | -0.14 | 0.08 | -0.02 | -0.08 | -0.03 | -0.06 | *-0.01* | *0.00* | 0.02 | -0.01 | -0.06 | *0.00* | 0.02 | 0.03 |
| (6) *Short* | -0.06 | -0.01 | *-0.01* | -0.01 | 0.02 | | -0.09 | 0.14 | -0.11 | *0.00* | 0.10 | 0.07 | 0.11 | 0.02 | 0.02 | 0.01 | *0.00* | 0.11 | 0.01 | 0.01 | *0.00* |
| (7) *Long* | 0.06 | *0.00* | *0.00* | 0.01 | *0.01* | -0.10 | | -0.03 | 0.01 | 0.12 | 0.08 | 0.13 | 0.13 | *-0.01* | -0.01 | -0.02 | -0.01 | 0.05 | -0.02 | -0.02 | -0.01 |
| (8) *NegPct* | -0.07 | -0.02 | -0.05 | -0.08 | -0.12 | 0.15 | -0.05 | | -0.11 | *0.01* | 0.18 | 0.05 | 0.10 | 0.04 | 0.02 | 0.04 | 0.05 | 0.15 | 0.12 | 0.08 | 0.06 |
| (9) *PosPct* | 0.04 | 0.01 | 0.04 | 0.04 | 0.06 | -0.11 | *0.00* | -0.13 | | 0.05 | -0.08 | -0.02 | -0.06 | *0.00* | *0.00* | 0.05 | *0.00* | -0.08 | -0.01 | 0.01 | *0.00* |
| (10) $ComPosPct_{i,[t,t+1]}$ | 0.01 | -0.01 | *0.00* | *0.00* | *0.00* | -0.02 | 0.07 | -0.04 | 0.06 | | 0.48 | 0.28 | 0.25 | 0.02 | 0.01 | -0.02 | -0.01 | -0.02 | 0.06 | 0.06 | 0.06 |
| (11) $ComNegPct_{i,[t,t+1]}$ | -0.04 | -0.02 | -0.03 | -0.04 | -0.07 | 0.08 | 0.04 | 0.18 | -0.07 | 0.23 | | 0.28 | 0.33 | 0.03 | 0.02 | -0.03 | *0.01* | 0.07 | 0.11 | 0.08 | 0.09 |
| (12) $ComPosPct_{i,[t+3,t+60]}$ | *0.00* | -0.01 | *0.00* | -0.01 | *-0.01* | 0.02 | 0.07 | -0.01 | 0.01 | 0.14 | 0.12 | | 0.66 | *0.00* | -0.01 | -0.05 | -0.03 | 0.06 | -0.02 | -0.02 | -0.02 |
| (13) $ComNegPct_{i,[t+3,t+60]}$ | -0.02 | -0.04 | -0.01 | -0.02 | -0.04 | 0.09 | 0.07 | 0.10 | -0.06 | 0.11 | 0.23 | 0.28 | | *0.00* | -0.01 | -0.06 | -0.02 | 0.10 | -0.01 | -0.02 | -0.02 |
| (14) *Upgrades* | *0.00* | *0.00* | *0.00* | -0.01 | -0.01 | *0.00* | -0.01 | 0.04 | *-0.01* | 0.01 | 0.02 | *0.00* | *0.00* | | 0.08 | 0.06 | 0.02 | 0.04 | 0.10 | 0.08 | 0.07 |
| (15) *Downgrades* | *0.01* | *0.00* | *0.00* | -0.01 | *0.00* | *0.00* | -0.02 | 0.02 | *0.00* | *0.01* | 0.02 | -0.01 | -0.01 | 0.07 | | 0.07 | 0.03 | 0.02 | 0.10 | 0.09 | 0.10 |
| (16) *PosES* | 0.05 | 0.01 | *0.01* | 0.01 | 0.01 | *0.00* | -0.03 | 0.05 | 0.05 | -0.02 | -0.02 | -0.04 | -0.05 | 0.05 | 0.06 | | -0.02 | *0.00* | 0.16 | 0.17 | 0.12 |
| (17) *NegES* | -0.07 | *-0.01* | *0.00* | -0.01 | -0.01 | *0.00* | -0.01 | 0.05 | *0.00* | -0.01 | 0.01 | -0.02 | -0.02 | 0.01 | 0.03 | -0.02 | | 0.01 | 0.10 | 0.09 | 0.07 |
| (18) *Volatility* | *0.00* | -0.04 | *0.00* | *0.00* | 0.03 | 0.09 | 0.03 | 0.13 | -0.08 | -0.01 | 0.07 | 0.01 | 0.06 | 0.04 | 0.02 | -0.01 | 0.01 | | -0.05 | -0.07 | -0.07 |
| (19) *DJPosPct* | -0.03 | 0.03 | -0.02 | -0.02 | -0.03 | *-0.01* | -0.04 | 0.14 | -0.02 | 0.02 | 0.10 | -0.02 | *0.00* | 0.10 | 0.09 | 0.11 | 0.08 | *0.00* | | 0.82 | 0.75 |
| (20) *DJNegPct* | 0.01 | 0.03 | *0.00* | *0.00* | 0.01 | *-0.01* | -0.03 | 0.05 | 0.03 | 0.02 | 0.04 | -0.02 | -0.02 | 0.07 | 0.08 | 0.13 | 0.07 | -0.04 | 0.51 | | 0.76 |
| (21) *IDJ* | *-0.01* | 0.04 | *-0.01* | -0.01 | 0.02 | -0.01 | -0.03 | 0.07 | *-0.01* | 0.04 | 0.07 | -0.02 | -0.02 | 0.07 | 0.10 | 0.12 | 0.07 | -0.03 | 0.62 | 0.66 | |

Table 4 presents correlation coefficients among variables. Spearman (Pearson) correlations are presented above (below) the diagonal. Italicized correlations are insignificantly different from 0 ($p>0.05$). All variables are defined in Appendix A.

## Table 5: Test of H1

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *Variables* | ALL | ALL | ALL | NO DOW-JONES | NO DOW-JONES OR EARNINGS SURPRISE |
| *Short* | -1.076*** | -1.070*** | -1.092*** | -1.354*** | -1.364*** |
| | (-8.92) | (-8.69) | (-8.84) | (-7.90) | (-7.82) |
| *Long* | 0.431*** | 0.436*** | 0.431*** | 0.544*** | 0.542*** |
| | (10.03) | (10.40) | (10.35) | (7.67) | (7.78) |
| *NegPct* | -22.321*** | -24.395*** | -23.184*** | -22.818*** | -22.726*** |
| | (-10.52) | (-12.20) | (-11.64) | (-7.15) | (-7.19) |
| *PosPct* | 14.528*** | 14.465*** | 14.025*** | 13.872*** | 13.185*** |
| | (8.03) | (8.12) | (7.88) | (5.59) | (5.28) |
| *lWordCount* | 0.022 | 0.021 | 0.027 | 0.091** | 0.098** |
| | (0.78) | (0.73) | (0.95) | (2.17) | (2.34) |
| $ComPosPct_{i,[t,t+1]}$ | 2.797** | 2.857** | 2.857** | 2.710 | 2.760* |
| | (2.39) | (2.45) | (2.44) | (1.64) | (1.68) |
| $ComNegPct_{i,[t,t+1]}$ | -10.564*** | -10.807*** | -10.114*** | -8.739*** | -7.863*** |
| | (-6.88) | (-6.95) | (-6.54) | (-4.60) | (-4.14) |
| *Upgrades* | | -0.090 | -0.040 | -0.428 | -0.479 |
| | | (-0.47) | (-0.21) | (-1.14) | (-1.45) |
| *Downgrades* | | 0.129 | 0.163* | 0.236 | 0.460*** |
| | | (1.34) | (1.70) | (1.44) | (2.68) |
| *Volatility* | | 1.253 | 1.150 | -0.674 | -0.563 |
| | | (1.26) | (1.16) | (-0.58) | (-0.50) |
| $AbRet_{i,[t-60,t-3]}$ | | -0.475*** | -0.483*** | -0.318* | -0.295* |
| | | (-3.59) | (-3.63) | (-1.87) | (-1.77) |
| $AbRet_{i,[t-2]}$ | | -1.394* | -1.429* | -1.718* | -1.807* |
| | | (-1.70) | (-1.74) | (-1.80) | (-1.87) |
| $AbRet_{i,[t-1]}$ | | -2.082** | -2.145*** | -1.892** | -1.828** |
| | | (-2.61) | (-2.69) | (-2.14) | (-2.08) |
| *PosES* | | 1.286*** | 1.313*** | 0.809** | |
| | | (9.42) | (9.81) | (2.26) | |
| *NegES* | | -2.357*** | -2.301*** | -2.312*** | |
| | | (-12.79) | (-12.45) | (-5.00) | |
| *DJPosPct* | | | -12.938*** | | |
| | | | (-6.51) | | |
| *DJNegPct* | | | 17.152*** | | |
| | | | (7.26) | | |
| *IDJ* | | | -0.129*** | | |
| | | | (-3.23) | | |
| n | 80,589 | 80,589 | 80,589 | 41,053 | 40,574 |
| Adjusted $R^2$ | 0.013 | 0.020 | 0.021 | 0.020 | 0.018 |

Table 5 presents results from estimating [1]. The dependent variable is $AbRet_{i,[t,t+1]}$, multiplied by 100 to facilitate exposition. Columns 1 through 3 include all observations, and Column 4 (5) excludes observations with concurrently issued Dow-Jones news content (Dow Jones news content or earnings surprises). All variables are defined in Appendix A. All estimations include year-month fixed effects. *** (**, *) denotes two-tailed significance at the $p<0.01$ ($p<0.05$, $p<0.10$) level (two-tailed) assessed using *t*-statistics (in parentheses) derived from White (1980) standard errors clustered by year-month to correct for cross-sectional correlation in abnormal returns.

## Table 6: Test of H2 and H3

| | | | | | Tests of Differences (*p*-values) | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | | | | |
| *Variables* | *Position=0* | *\|Position\|=1* | *Position=-1* | *Position=1* | (1) v. (2) | (1) v. (3) | (1) v. (4) | (3) v. (4) |
| *NegPct* | -20.785*** | -28.118*** | -39.166*** | -28.465*** | 0.00 | 0.04 | 0.04 | 0.33 |
| | (-9.04) | (-8.29) | (-3.74) | (-7.83) | | | | |
| *PosPct* | 12.684*** | 16.844*** | 67.276*** | 14.021*** | 0.01 | 0.01 | 0.78 | 0.01 |
| | (7.21) | (3.71) | (3.04) | (3.18) | | | | |
| *lWordCount* | 0.013 | 0.095* | -0.459** | 0.087** | 0.82 | 0.01 | 0.20 | 0.00 |
| | (0.35) | (1.87) | (-2.50) | (2.08) | | | | |
| $ComPosPct_{i,[t,t+1]}$ | 2.090 | 5.075* | 8.124 | 4.750* | 0.10 | 0.64 | 0.38 | 0.78 |
| | (1.52) | (1.92) | (0.62) | (1.86) | | | | |
| $ComNegPct_{i,[t,t+1]}$ | -7.790*** | -15.713*** | -1.273 | -17.197*** | 0.00 | 0.56 | 0.00 | 0.16 |
| | (-4.81) | (-5.04) | (-0.11) | (-5.34) | | | | |
| *Upgrades* | -0.097 | 0.055 | 0.442 | -0.033 | 0.73 | 0.68 | 0.87 | 0.72 |
| | (-0.43) | (0.16) | (0.33) | (-0.10) | | | | |
| *Downgrades* | 0.313*** | -0.185 | -0.497 | -0.135 | 0.04 | 0.23 | 0.07 | 0.58 |
| | (3.00) | (-0.83) | (-0.73) | (-0.61) | | | | |
| *PosES* | 1.311*** | 1.277*** | 0.859 | 1.264*** | 0.68 | 0.67 | 0.87 | 0.71 |
| | (8.77) | (4.80) | (0.81) | (4.72) | | | | |
| *NegES* | -2.608*** | -1.515*** | -2.704 | -1.428*** | 0.03 | 0.96 | 0.02 | 0.53 |
| | (-11.01) | (-3.70) | (-1.33) | (-3.47) | | | | |
| *Volatility* | -0.810 | 4.046*** | -11.294*** | 7.542*** | 0.02 | 0.00 | 0.00 | 0.00 |
| | (-0.68) | (2.79) | (-3.81) | (4.63) | | | | |
| $AbRet_{i,[t-60,t-3]}$ | -0.428*** | -0.608*** | -2.012*** | -0.419* | 0.25 | 0.00 | 0.98 | 0.00 |
| | (-2.70) | (-2.74) | (-4.27) | (-1.83) | | | | |
| $AbRet_{i,[t-2]}$ | -1.140 | -1.963 | -12.907*** | -0.318 | 0.58 | 0.01 | 0.66 | 0.00 |
| | (-1.11) | (-1.41) | (-3.24) | (-0.20) | | | | |
| $AbRet_{i,[t-1]}$ | -2.266** | -1.893 | -8.238** | -1.186 | 0.84 | 0.10 | 0.55 | 0.04 |
| | (-2.09) | (-1.42) | (-2.49) | (-0.85) | | | | |
| *DJPosPct* | -14.662*** | -9.236** | -22.579 | -9.452** | 0.25 | 0.60 | 0.26 | 0.38 |
| | (-5.88) | (-2.44) | (-1.49) | (-2.55) | | | | |
| *DJNegPct* | 18.251*** | 14.217*** | 39.336 | 10.293** | 0.30 | 0.41 | 0.18 | 0.28 |
| | (6.53) | (2.95) | (1.46) | (2.13) | | | | |
| *IDJ* | -0.030 | -0.320*** | 0.520 | -0.420*** | 0.00 | 0.13 | 0.00 | 0.01 |
| | (-0.63) | (-4.02) | (1.45) | (-4.90) | | | | |
| *Short* | | -0.890*** | | | | | | |
| | | (-4.51) | | | | | | |
| *Long* | | 0.608*** | | | | | | |
| | | (3.84) | | | | | | |
| n | 54,042 | 26,391 | 2,330 | 24,061 | | | | |
| Adjusted $R^2$ | 0.018 | 0.027 | 0.043 | 0.019 | | | | |

Table 6 presents results from estimating [1] separately by position and tests of differences in coefficients across the four columns. The dependent variable is $AbRet_{i,[t,t+1]}$, multiplied by 100 to facilitate exposition. All variables are defined in Appendix A. All estimations include year-month fixed effects. *** (**, *) denotes significance at the $p<0.01$ ($p<0.05$, $p<0.10$) level (one-tailed where a prediction is made and two-tailed otherwise) assessed using *t*-statistics (in parentheses) derived from White (1980) standard errors clustered by year-month to correct for cross-sectional correlation in abnormal returns.

**Table 7: Early Morning SA Articles**

| *Variable* | (1) ALL | (2) ALL | (3) ALL | (4) ALL | (5) NO DOW-JONES | (6) NO DOW-JONES OR EARNINGS SURPRISE |
|---|---|---|---|---|---|---|
| *Short* | -1.322*** | -1.139*** | -1.101*** | -1.114*** | -1.292*** | -1.355*** |
| | (-5.80) | (-4.96) | (-4.80) | (-4.85) | (-4.01) | (-4.24) |
| *Long* | 0.355*** | 0.356*** | 0.368*** | 0.363*** | 0.564*** | 0.567*** |
| | (6.62) | (6.75) | (7.09) | (7.08) | (5.96) | (5.92) |
| *NegPct* | | -8.406** | -9.674*** | -8.520*** | -14.856*** | -15.061*** |
| | | (-2.18) | (-2.66) | (-2.33) | (-2.64) | (-2.71) |
| *PosPct* | | 15.054*** | 13.677*** | 13.439*** | 19.025*** | 18.699*** |
| | | (4.76) | (4.42) | (4.29) | (4.03) | (3.86) |
| *lWordCount* | | 0.030 | 0.029 | 0.026 | 0.014 | 0.010 |
| | | (0.54) | (0.55) | (0.49) | (0.20) | (0.13) |
| $ComPosPct_{i,[t,t+1]}$ | | 2.120 | 1.906 | 1.944 | 0.841 | 0.596 |
| | | (1.17) | (1.05) | (1.07) | (0.31) | (0.22) |
| $ComNegPct_{i,[t,t+1]}$ | | -8.741*** | -7.990*** | -7.300*** | -5.553* | -5.898** |
| | | (-3.86) | (-3.51) | (-3.22) | (-1.98) | (-2.15) |
| *Upgrades* | | | 0.139 | 0.192 | -0.100 | 0.011 |
| | | | (0.51) | (0.71) | (-0.20) | (0.02) |
| *Downgrades* | | | -0.079 | -0.045 | -0.400 | -0.272 |
| | | | (-0.55) | (-0.31) | (-1.48) | (-1.00) |
| *PosES* | | | 0.778*** | 0.812*** | 0.555 | |
| | | | (3.71) | (3.85) | (0.96) | |
| *NegES* | | | -1.872*** | -1.810*** | -1.122* | |
| | | | (-6.90) | (-6.63) | (-1.92) | |
| *Volatility* | | | -1.454 | -1.650 | 0.214 | 0.352 |
| | | | (-1.06) | (-1.20) | (0.11) | (0.18) |
| $AbRet_{i,[t-60,t-3]}$ | | | -0.246 | -0.245 | -0.396 | -0.357 |
| | | | (-1.10) | (-1.09) | (-1.47) | (-1.31) |
| $AbRet_{i,[t-2]}$ | | | -0.164 | -0.215 | 0.512 | 0.295 |
| | | | (-0.11) | (-0.15) | (0.22) | (0.13) |
| $AbRet_{i,[t-1]}$ | | | -1.857* | -1.912* | -2.611* | -2.596* |
| | | | (-1.77) | (-1.84) | (-1.68) | (-1.68) |
| *DJPosPct* | | | | -11.375*** | | |
| | | | | (-3.95) | | |
| *DJNegPct* | | | | 10.097** | | |
| | | | | (2.40) | | |
| *IDJ* | | | | -0.091 | | |
| | | | | (-1.33) | | |
| n | 22,064 | 22,064 | 22,064 | 22,064 | 11,138 | 10,989 |
| Adjusted $R^2$ | 0.009 | 0.012 | 0.017 | 0.018 | 0.021 | 0.022 |

Table 7 presents results from estimating [1] using a sample of articles published between 9:30 am and 1:00 pm on trading days. The dependent variable is the two-day abnormal return, where the day 0 return is defined as (closing price – opening price) / opening price. The two-day abnormal return is multiplied by 100 to facilitate exposition. Columns 1 through 3 impose no additional sample screens, and Column 4 (5) excludes observations with concurrently issued Dow-Jones news content (Dow Jones news content or earnings surprises). All variables are defined in Appendix A. All estimations include year-month fixed effects. *** (**, *) denotes two-tailed significance at the $p<0.01$ ($p<0.05$, $p<0.10$) level (two-tailed) assessed using *t*-statistics (in parentheses) derived from White (1980) standard errors clustered by year-month to correct for cross-sectional correlation in abnormal returns.

## Table 8: Article Length and Author Positions

| Variables | (1) ALL | (2) NO DOWJONES | (3) NO DOWJONES OR EARNINGS SURPRISE |
|---|---|---|---|
| *Short* | -0.941*** | -1.123*** | -1.129*** |
| | (-7.76) | (-6.28) | (-6.26) |
| *Long* | 0.430*** | 0.532*** | 0.530*** |
| | (10.08) | (7.22) | (7.33) |
| *lWordCount* | -0.001 | 0.042 | 0.052 |
| | (-0.02) | (0.87) | (1.08) |
| *Short x lWordCount* | -0.656*** | -0.683*** | -0.691*** |
| | (-4.21) | (-3.46) | (-3.44) |
| *Long x lWordCount* | 0.173*** | 0.247*** | 0.238** |
| | (3.03) | (2.69) | (2.61) |
| *NegPct* | -23.228*** | -22.873*** | -22.769*** |
| | (-11.70) | (-7.23) | (-7.26) |
| *PosPct* | 14.169*** | 14.177*** | 13.481*** |
| | (7.99) | (5.72) | (5.41) |
| $ComPosPct_{i,[t,t+1]}$ | 2.769** | 2.651 | 2.704 |
| | (2.36) | (1.60) | (1.64) |
| $ComNegPct_{i,[t,t+1]}$ | -9.943*** | -8.609*** | -7.738*** |
| | (-6.46) | (-4.52) | (-4.07) |
| *Upgrades* | -0.042 | -0.433 | -0.485 |
| | (-0.22) | (-1.15) | (-1.47) |
| *Downgrades* | 0.165* | 0.243 | 0.467*** |
| | (1.73) | (1.48) | (2.71) |
| *PosES* | 1.317*** | 0.803** | |
| | (9.80) | (2.24) | |
| *NegES* | -2.301*** | -2.324*** | |
| | (-12.50) | (-5.05) | |
| *Volatility* | 1.225 | -0.619 | -0.509 |
| | (1.24) | (-0.54) | (-0.45) |
| $AbRet_{i,[t-60,t-3]}$ | -0.481*** | -0.317* | -0.294* |
| | (-3.62) | (-1.87) | (-1.77) |
| $AbRet_{i,[t-2]}$ | -1.450* | -1.765* | -1.853* |
| | (-1.76) | (-1.84) | (-1.92) |
| $AbRet_{i,[t-1]}$ | -2.159*** | -1.925** | -1.859** |
| | (-2.71) | (-2.18) | (-2.12) |
| *DJPosPct* | -12.938*** | | |
| | (-6.50) | | |
| *DJNegPct* | 17.082*** | | |
| | (7.20) | | |
| *IDJ* | -0.130*** | | |
| | (-3.27) | | |
| n | 80,589 | 41,053 | 40,574 |
| Adjusted $R^2$ | 0.022 | 0.020 | 0.019 |

Table 8 presents results from estimating [1] after incorporate article length into the analyses. Column 1 includes all observations, and Column 2 (3) excludes firm-days with contemporaneously issued Dow-Jones content (contemporaneously issued Dow-Jones content or earnings-surprise announcements). All variables are defined in Appendix A. All estimations include year-month fixed effects. *** (**, *) denotes significance at the *p<0.01* (*p<0.05, p<0.10*) level (one-tailed where a prediction is made and two-tailed otherwise two-tailed) assessed using *t*-statistics (in parentheses) derived from White (1980) standard errors clustered by year-month to correct for cross-sectional correlation in abnormal returns.

## Table 9: First-time vs. Repeated Disclosures of Position

| *Variable* | (1) ALL | (2) NO DOWJONES | (3) NO DOWJONES OR EARNINGS SURPRISE |
|---|---|---|---|
| *FirstDisc* | 0.032 | 0.083 | 0.067 |
| | (0.34) | (0.64) | (0.53) |
| *Short* | -0.405** | -0.617** | -0.630** |
| | (-2.45) | (-2.55) | (-2.46) |
| *FirstDisc x Short* | -1.345*** | -1.389*** | -1.375*** |
| | (-6.13) | (-4.35) | (-4.21) |
| *Long* | 0.344*** | 0.406*** | 0.404*** |
| | (6.94) | (4.96) | (5.06) |
| *FirstDisc x Long* | 0.197*** | 0.300*** | 0.299*** |
| | (3.20) | (3.46) | (3.53) |
| *NegPct* | -27.332*** | -25.121*** | -24.670*** |
| | (-9.30) | (-5.21) | (-5.20) |
| *FirstDisc x NegPct* | 7.568* | 4.181 | 3.570 |
| | (1.89) | (0.68) | (0.59) |
| *PosPct* | 19.423*** | 18.647*** | 17.367*** |
| | (5.79) | (3.75) | (3.54) |
| *FirstDisc x PosPct* | -9.750** | -8.457 | -7.430 |
| | (-2.30) | (-1.44) | (-1.30) |
| *lWordCount* | 3.288 | 16.666*** | 18.655*** |
| | (0.83) | (2.80) | (3.29) |
| *FirstDisc x lWordCount* | -1.107 | -14.151** | -16.360** |
| | (-0.20) | (-2.02) | (-2.35) |
| *$ComPosPct_{i,[t,t+1]}$* | 1.986 | 1.155 | 0.967 |
| | (1.12) | (0.46) | (0.37) |
| *FirstDisc x $ComPosPct_{i,[t,t+1]}$* | 1.500 | 2.537 | 2.917 |
| | (0.65) | (0.83) | (0.94) |
| *$ComNegPct_{i,[t,t+1]}$* | -10.219*** | -7.329** | -6.739** |
| | (-3.77) | (-2.34) | (-2.20) |
| *FirstDict x $ComNegPct_{i,[t,t+1]}$* | 0.231 | -2.302 | -1.851 |
| | (0.07) | (-0.57) | (-0.47) |
| Controls included? | Yes | Yes | Yes |
| n | 80,589 | 41,053 | 40,574 |
| Adjusted $R^2$ | 0.022 | 0.021 | 0.019 |

Table 9 presents results from estimating [1] after including interactions between SA article-related variables and *FirstDisc*, an indicator equaling 1 the first time the author writes about the firm. All other control variables from [1] are included, but coefficient estimates are suppressed to facilitate exposition. Column 1 includes all observations, and Column 2 (3) excludes firm-days with contemporaneously issued Dow-Jones content (contemporaneously issued Dow-Jones content or earnings-surprise announcements). All variables are defined in Appendix A. All estimations include year-month fixed effects. *** (**, *) denotes significance at the $p<0.01$ ($p<0.05$, $p<0.10$) level (one-tailed where a prediction is made and two-tailed otherwise) assessed using *t*-statistics (in parentheses) derived from White (1980) standard errors clustered by year-month to correct for cross-sectional correlation in abnormal returns.

## Table 10: Author Position and Post-Publication Drift

| Variable | (1) ALL | (2) ALL | (3) ALL | (4) NO DOW-JONES | (5) NO DOW-JONES OR EARNINGS SURPRISE |
|---|---|---|---|---|---|
| *Short* | -0.969 | -0.642 | -0.558 | -0.051 | -0.079 |
| | (-1.53) | (-1.03) | (-0.90) | (-0.06) | (-0.10) |
| *Long* | -0.237 | -0.148 | -0.115 | -0.041 | -0.051 |
| | (-0.99) | (-0.62) | (-0.48) | (-0.13) | (-0.16) |
| *NegPct* | -36.826** | -28.627** | -33.067** | -52.010** | -50.966** |
| | (-2.57) | (-2.22) | (-2.59) | (-2.61) | (-2.57) |
| *PosPct* | 26.113* | 18.897 | 18.390 | 27.840 | 28.299 |
| | (1.80) | (1.42) | (1.39) | (1.51) | (1.55) |
| *lWordCount* | 0.431*** | 0.473*** | 0.425*** | 0.264 | 0.241 |
| | (2.90) | (3.25) | (2.87) | (1.06) | (0.96) |
| $ComPosPct_{i,[t,t+1]}$ | -9.865* | -11.212** | -11.682** | -10.790 | -10.106 |
| | (-1.89) | (-2.15) | (-2.25) | (-1.38) | (-1.27) |
| $ComNegPct_{i,[t,t+1]}$ | -14.470* | -10.578 | -13.696* | -28.501*** | -28.878*** |
| | (-1.68) | (-1.37) | (-1.78) | (-2.91) | (-2.92) |
| $ComPosPct_{i,[t+3,t+60]}$ | -0.749 | -0.324 | 0.653 | 12.495 | 12.703 |
| | (-0.15) | (-0.07) | (0.13) | (1.45) | (1.47) |
| $ComNegPct_{i,[t+3,t+60]}$ | -50.922*** | -46.776*** | -45.637*** | -66.354*** | -66.140*** |
| | (-6.93) | (-6.55) | (-6.37) | (-6.74) | (-6.71) |
| *Upgrades* | | 0.185 | -0.134 | -0.608 | -0.924 |
| | | (0.34) | (-0.25) | (-0.64) | (-0.99) |
| *Downgrades* | | 0.234 | -0.003 | -0.363 | -0.228 |
| | | (0.71) | (-0.01) | (-0.62) | (-0.38) |
| *PosES* | | 0.811* | 0.393 | 1.548 | |
| | | (1.98) | (0.97) | (1.43) | |
| *NegES* | | -0.516 | -0.936 | -1.985 | |
| | | (-0.92) | (-1.62) | (-1.57) | |
| *Volatility* | | -26.757*** | -25.604*** | -44.412*** | -44.319*** |
| | | (-3.18) | (-3.05) | (-4.81) | (-4.79) |
| $AbRet_{i,[t-60,t-3]}$ | | 0.670 | 0.636 | -0.270 | -0.282 |
| | | (0.45) | (0.43) | (-0.18) | (-0.19) |
| $AbRet_{i,[t-2]}$ | | 3.427 | 3.587 | -0.863 | -1.205 |
| | | (0.88) | (0.92) | (-0.17) | (-0.23) |
| $AbRet_{i,[t-1]}$ | | -1.910 | -1.785 | -5.614 | -5.629 |
| | | (-0.64) | (-0.60) | (-1.44) | (-1.46) |
| *DJPosPct* | | | 23.935** | | |
| | | | (2.51) | | |
| *DJNegPct* | | | 14.415 | | |
| | | | (1.12) | | |
| *IDJ* | | | 0.644** | | |
| | | | (2.57) | | |
| $AbRet_{i,[t,t+1]}$ | 1.104 | 1.072 | 1.294 | -0.739 | -1.272 |
| | (0.42) | (0.40) | (0.48) | (-0.21) | (-0.34) |
| N | 80,589 | 80,589 | 80,589 | 41,053 | 40,574 |
| Adjusted $R^2$ | 0.022 | 0.025 | 0.026 | 0.030 | 0.030 |

Table 10 presents results from estimating [1] using $AbRet_{i,[t+3,t+60]}$ (multipled by 100) as the dependent variable. Columns 1 to 3 include all observations, and Column 4 (5) excludes observations with concurrently issued Dow-Jones news content (Dow Jones news content or earnings surprises). All variables are defined in Appendix A. All estimations include year-month fixed effects. *** (**, *) denotes significance at the $p<0.01$ ($p<0.05$, $p<0.10$) level (one-tailed where a prediction is made and two-tailed otherwise) assessed using *t*-statistics (in parentheses) derived from White (1980) standard errors clustered by year-month to correct for cross-sectional correlation in abnormal returns.

## Table 11: The Predictive Power of Seeking Alpha Content by Position

| Variable | (1) Position=0 | (2) \|Position\|=1 | (3) Position=-1 | (4) Position=1 | Tests of Differences (1) v. (2) | (1) v. (3) | (1) v. (4) | (3) v. (4) |
|---|---|---|---|---|---|---|---|---|
| *NegPct* | -29.580** | -35.551* | -116.327** | -22.665 | 0.56 | 0.09 | 0.73 | 0.09 |
| | (-2.34) | (-1.75) | (-2.21) | (-1.04) | | | | |
| *PosPct* | 17.990 | 16.273 | 105.897 | 6.953 | 0.94 | 0.27 | 0.59 | 0.22 |
| | (1.22) | (0.85) | (1.31) | (0.36) | | | | |
| *lWordCount* | 0.307 | 0.452** | 0.797 | 0.578*** | 0.17 | 0.42 | 0.30 | 0.75 |
| | (1.64) | (2.10) | (1.23) | (2.85) | | | | |
| $ComPosPct_{i,[t,t+1]}$ | -13.443** | -5.236 | 44.937 | -7.959 | 0.44 | 0.35 | 0.63 | 0.40 |
| | (-2.23) | (-0.54) | (0.70) | (-0.83) | | | | |
| $ComNegPct_{i,[t,t+1]}$ | -18.167** | -0.931 | 34.511 | -3.934 | 0.28 | 0.35 | 0.38 | 0.47 |
| | (-2.16) | (-0.06) | (0.61) | (-0.27) | | | | |
| $ComPosPct_{i,[t+3,t+60]}$ | -3.101 | 9.519 | -2.877 | 11.666 | 0.22 | 1.00 | 0.18 | 0.79 |
| | (-0.53) | (1.07) | (-0.05) | (1.25) | | | | |
| $ComNegPct_{i,[t+3,t+60]}$ | -38.068*** | -63.740*** | -18.398 | -68.815*** | 0.06 | 0.66 | 0.02 | 0.26 |
| | (-4.26) | (-5.86) | (-0.40) | (-6.80) | | | | |
| *Upgrades* | -0.025 | -0.733 | -3.629 | -0.340 | 0.61 | 0.20 | 0.80 | 0.25 |
| | (-0.04) | (-0.82) | (-1.30) | (-0.35) | | | | |
| *Downgrades* | 0.320 | -0.734 | 0.268 | -0.806 | 0.14 | 0.98 | 0.10 | 0.66 |
| | (0.96) | (-1.13) | (0.11) | (-1.24) | | | | |
| *PosES* | 0.081 | 1.095 | 4.297 | 0.881 | 0.19 | 0.17 | 0.28 | 0.23 |
| | (0.18) | (1.45) | (1.36) | (1.41) | | | | |
| *NegES* | -0.536 | -1.810 | -6.448 | -1.364 | 0.33 | 0.26 | 0.53 | 0.36 |
| | (-0.94) | (-1.44) | (-1.20) | (-1.03) | | | | |
| *Volatility* | -26.374*** | -22.907** | -18.042 | -21.672** | 0.74 | 0.62 | 0.61 | 0.82 |
| | (-2.79) | (-2.41) | (-1.00) | (-2.31) | | | | |
| $AbRet_{i,[t-60,t-3]}$ | 2.222 | -1.999 | -3.563 | -1.731 | 0.00 | 0.05 | 0.00 | 0.52 |
| | (1.47) | (-1.17) | (-1.05) | (-1.06) | | | | |
| $AbRet_{i,[t-2]}$ | 4.639 | 1.688 | -27.198* | 5.571 | 0.69 | 0.04 | 0.91 | 0.06 |
| | (1.17) | (0.24) | (-1.77) | (0.72) | | | | |
| $AbRet_{i,[t-1]}$ | 1.021 | -7.243 | -12.150 | -6.668 | 0.14 | 0.26 | 0.17 | 0.62 |
| | (0.30) | (-1.48) | (-1.04) | (-1.35) | | | | |
| *DJPosPct* | 37.054*** | -1.292 | 63.734 | -6.086 | 0.09 | 0.67 | 0.02 | 0.23 |
| | (3.14) | (-0.08) | (1.03) | (-0.40) | | | | |
| *DJNegPct* | 4.764 | 31.814 | -65.127 | 39.806 | 0.24 | 0.46 | 0.17 | 0.30 |
| | (0.38) | (1.33) | (-0.68) | (1.52) | | | | |
| *IDJ* | 0.385 | 1.030** | 1.920 | 1.015** | 0.09 | 0.28 | 0.12 | 0.53 |
| | (1.49) | (2.58) | (1.31) | (2.52) | | | | |
| $AbRet_{i,[t,t+1]}$ | 0.048 | -0.047 | -0.009 | -0.061 | 0.04 | 0.61 | 0.02 | 0.63 |
| | (1.52) | (-1.20) | (-0.09) | (-1.45) | | | | |
| *Short* | | -1.671 | | | | | | |
| | | (-1.66) | | | | | | |
| *Long* | | -1.156* | | | | | | |
| | | (-1.76) | | | | | | |
| n | 54,042 | 26,391 | 2,330 | 24,061 | | | | |
| Adjusted $R^2$ | 0.026 | 0.033 | 0.067 | 0.031 | | | | |

Table 11 presents results from estimating [1] using $AbRet_{i,[t+3,t+60]}$ (multipled by 100) instead of $AbRet_{i,[t,t+1]}$ as the dependent variable, separately by position, and tests of differences in coefficients across the four columns. All variables are defined in Appendix A. All estimations include year-month fixed effects. *** (**, *) denotes significance at the $p<0.01$ ($p<0.05$, $p<0.10$) level (one-tailed where a prediction is made and two-tailed otherwise) assessed using $t$-statistics (in parentheses) derived from White (1980) standard errors clustered by year-month to correct for cross-sectional correlation in abnormal returns.