

# Complexity Aversion when *Seeking Alpha*\*

Tarik Umar<sup>†</sup>  
Rice University

December 31, 2017

## Abstract

I provide evidence that news' complexity and sentiment affect investor attention to news and market reactions. Using field data with randomization from *Seeking Alpha*, I find standard-deviation increases in headline length (negativity) lead to 12%-fewer (2%-more) views. Effects are greater for less-sophisticated investors. Studying company-earnings-press-release headlines, and using the length of firms' legal names to instrument headline lengths, I find standard-deviation increases in length lead to 7%-lower turnover, 60-basis-points-tighter-intraday-price ranges, and 15-basis-points-return underreactions, which reverse in three months. Complexity matters more for less-surprising news released on quieter days to less-sophisticated investors.

JEL: G14, G41, D83, M41

---

\*I would like to thank my committee of advisors from the University of Chicago's Booth School of Business: Lubos Pastor (co-chair), Kelly Shue (co-chair), Amir Sufi, and Richard Thaler. This paper will be presented at the AFA meetings in January 2018 and the Q-Group meetings in May 2018. This paper was awarded "best PhD paper" at the 2016 Colorado Finance Summit and was a finalist for the BlackRock Applied Research Award. This work has benefited from feedback at the following seminars: Boston College, Chicago Booth, Dartmouth, Rice University, Securities and Exchange Commission, Texas A&M, University of California San Diego, University of Oklahoma, University of Oregon, University of Virginia, Vanderbilt, Washington University-St Louis, and Wharton. Paper featured in *Wall Street Journal* piece titled "You Probably Tend Not To Read Articles With Headlines That Are This Long, But Maybe You Should" on November 13, 2017. Also featured by Matt Levine on *Bloomberg View* on March 10, 2017. The following individuals connected with *Seeking Alpha* provided me with valuable resources: Eli Hoffmann (CEO), Selig Davis (Vice President of Audience and Mobile), and Yoni Madar (Mobile Product Analyst).

<sup>†</sup>Assistant Professor of Finance (tarik.umar@rice.edu)

# 1 Introduction

Investors must allocate their limited-attention across news (Kahneman, 1973). As the information investors sift through is often textual, the textual attributes of news (e.g., complexity, and sentiment) may meaningfully affect how investors allocate their attention and thus individual and market outcomes. Consistent with attributes mattering, prior studies find associations between textual information and market behavior. For example, more complex corporate disclosures are correlated with less trading by individual investors and greater-post-filing volatility. The investor-attention literature finds investors underreact to subtle, hard-to-find news. And, the sentiment literature finds investors underreact to negative sentiment in news.<sup>1</sup> Identifying causally whether textual attributes matter remains a challenge because omitted variables, related to the event, may influence market outcomes, investor attention, and the news' complexity and sentiment (Bloomfield, 2008). Furthermore, characterizing how textual attributes matter is challenged by high correlations between attributes, such as a positive correlation between negative sentiment and complexity. In this paper, I show that textual attributes affect investor attention to news using unique field data with randomization and isolate changes in the complexity of information from other attributes using an instrumental variable. Using the instrumental variable, I then show that the headlines of earnings announcements explain market responses to news, including a temporary price underreaction.

To identify how textual attributes affect investor attention, I need a measure of investor attention to a particular text, variation in the text unrelated to the event reported, and random assignment of the varied texts to investors. To address these challenges, I analyze field data with randomization from *Seeking Alpha*, a crowdsourced-investment-research firm. On January 3, 2016, *Seeking Alpha* began allowing authors of stock reports to propose two plausible titles for the same report. The editor assigned to review the stock report can provide a third title. The two-to-three plausible titles are then randomly sampled on investors who signed up to receive real-time-alert emails about the topic company (“title testing”). Each investor receives an email with only one randomly assigned title, and the emails are otherwise identical. None of the body of the stock report is included in the email. Investors must click a link in the email to read the full stock report. My chief measure of attention to a title is the number of investors who click the

---

<sup>1</sup>For models of investor attention allocation see Gabaix and Laibson (2005); Gabaix et al. (2006); Peng (2005); Peng and Xiong (2006); Kacperczyk et al. (2009); Bordalo et al. (2013). For work on complexity and market outcomes see Li (2008); You and Zhang (2009); Miller (2010); Loughran and McDonald (2011); Green and Jame (2013); Lawrence (2013); Loughran and McDonald (2014); Hwang and Kim (2016); Loughran and McDonald (2017). For work on sentiment and market outcomes see Barberis et al. (1998); Antweiler and Frank (2004); Baker and Wurgler (2006); Tetlock (2007); Tetlock et al. (2008); Engelberg (2008); Dougal et al. (2012); Hillert et al. (2014); Chen et al. (2014); Da et al. (2015); Hartzmark and Shue (2015). For work on underreactions to news see Giglio and Shue (2014) for underreactions to absence of news, Da et al. (2014) for underreactions to slow-moving information, Tetlock (2011) and Cohen et al. (2015) for reactions to stale news and subtle changes in filings, Loh (2010) for underreactions to stock recommendations, and DellaVigna and Pollet (2009), Niessner (2014) and Hirshleifer et al. (2009) for underreactions to news released on days investors are distracted, like Fridays, around holidays, or days with many other announcements. However, Michaely et al. (2016) find a selection effect explains the Friday effect. For evidence of limited attention of market makers see Corwin and Coughenour (2008); Chakrabarty and Moulton (2012)

link in the email. I use the number of clicks within 30 minutes of the email being sent to focus on attentive and active investors. I also measure the number of investors who scrolled to the end of the report by title. These direct measures of attention to a specific text differ from measures of attention used in prior studies, including Google search volume, extreme returns, abnormal trading, and advertising expenses.<sup>2</sup>

The *Seeking Alpha*-title-testing data allow me to employ a stock-report fixed effect to examine, within-a-report, how differences in title characteristics lead to differences in attention to news. The setting uniquely holds fixed the event discussed, as well as the topic firm, author, and date. The randomized assignment of titles to investors tracking the company ensures omitted characteristics of the recipient investors and the news event are not driving the relation between textual attributes and attention. In short, the title-testing data reveals what investor attention to the same news story would have been if a different title were chosen.

Using the stock-report fixed effect, I find that investors are significantly more attracted to short and simple titles. A one-standard-deviation increase in title length leads to 12% fewer page views. To help appreciate the magnitude, consider that, within a firm, investors pay 8% more attention to stock reports released on days when the VIX is a standard-deviation higher. I find a similar negative relation between attention and both average word length and number of words, suggesting that investors prefer shorter and simpler titles. The magnitudes are surprisingly large given that the subjects are investors in a high-stakes setting and have indicated an interest in the news by signing up to receive alert emails on the company. Also, the subjects are highly-engaged investors, who respond within 30 minutes of an email alert. One might expect that the magnitudes are even larger for less-engaged and less-interested investors. I find similar magnitudes when comparing pairs of titles with very similar word usage and character length and when examining very short titles that would easily fit on a mobile phone's screen. While conventional wisdom tends to recommend shorter and simpler writing, the authors of these stock reports appear to routinely violate the convention by proposing longer titles, which predictably garner less attention.

I also find that negative titles attract investor attention, and positive titles repel attention. Using Bill McDonald's lexicons adapted for financial texts, I measure sentiment by counting the number of negative and positive words in titles (Loughran and McDonald, 2011). A standard-deviation increase in a title's negativity predicts a 2% increase in page views. The magnitude of the effect of sentiment on attention is likely attenuated as sentiment is difficult to measure, especially for short titles. However, the effect decreases by 20% if I control for title complexity. This result shows that controlling for complexity is important in studies exploring the relation between sentiment and market outcomes. Also, the negative attention-sentiment relation suggests investors are more attracted to dis-confirming news, assuming *Seeking Alpha* investors are mostly long-oriented investors.

---

<sup>2</sup>Da et al. (2011), Bank et al. (2011), and Drake2012 uses Google Search Volume. Barber and Odean (2008) uses abnormal trading volume. Gervais et al. (2001); Barber and Odean (2008); Hou et al. (2009) use abnormal returns. Grullon et al. (2004); Chemmanur and Yan (2009); Fang and Peress (2009); Lou (2014); Ahern and Sosyura (2015) use advertising expense. Yuan (2008, 2015) and Li and Yu (2012) use new market highs.

The attraction to short, simple, and negative titles is stronger when the investors tracking a company are less-sophisticated. I do not have direct data describing the individuals receiving email alerts. However, the public comment section of stock reports reveals individual identities, allowing me to derive measures that characterize the sophistication of investors tracking the firm. The average sophistication of investors who write stock reports is likely to be higher than those who only read reports. Also, more sophisticated investors are likely to be more numerical and write longer comments. Consistent with this reasoning, I find that commenters who self-report having an MBA and a occupation in finance tend to write more numerical and detailed comments than those that do not. I find the sensitivities of attention to title length and sentiment increase with the fraction of comments from investors who have never written a stock report and decreases with the numerical intensity of comments, the average length of comments, and the proportion of commenters with MBAs and in finance jobs.

I also find that investors who access reports with longer titles are more thorough. The data include the number of investors who scroll to the end of the report by title received. The underlying stock report is the same so that differences in the propensity to read to the end must be due to a selection effect. I find that investors who do access stock reports with longer titles are significantly more likely to read the full report. These findings are consistent with more-sophisticated investors being less complexity averse, which may help explain prior findings showing that cognitive abilities are related to financial outcomes.<sup>3</sup> Even academics appear complexity averse, because an author's papers on *Social Science Research Network* with twice the title length receive 11% fewer views, 13% fewer downloads, and 4% fewer citations.

Although the title-testing data with randomization holds the context fixed, I do not write the sampled titles. Instead, analysts craft the titles. Therefore, longer and more positive titles might be more informative, reducing the need to read the stock report. Also, I have not isolated changes in the complexity of titles from changes in other attributes of titles.

To resolve this concern, I use the length of a company's legal name to instrument for title length. *Seeking Alpha* titles nearly always include a company's name, and companies with longer names in an industry tend to also have longer titles. The length of a company's legal name likely satisfies the exclusion restriction as legal names are chosen in the past and thus do not provide investors with new information. For example, company-name length is unrelated to the earnings surprise and tone of news. However, I do find that company-name length is negatively correlated with firm size, but after controlling for a polynomial of firm size, company-name length is not meaningfully related to other firm characteristics. Also, adding firm controls beyond firm size does not alter regression results, consistent with name length providing plausibly exogenous variation in title length conditional on firm size.

Using the instrument, I find a negative effect of title length on page views. I also find a positive effect of title length on the read-to-end rate. The instrumental-variables results suggest investors are complexity averse and not simply reading longer titles less because longer titles are

---

<sup>3</sup>See Feng and Seasholes (2005); Grinblatt et al. (2008); Agarwal et al. (2009); Grinblatt et al. (2009, 2011).

more informative. The magnitudes are not significantly different from those determined using the title-testing data, offering a nice test of the instrument before applying the instrument to the titles of company-earnings-press releases. I find similar results if I split title length into length associated with content and length associated with the firm's name. If the information story were true, then only length associated with content should be negatively related to attention. Instead, I find that each character associated with the name and content have nearly identical negative effects on attention, controlling for the other.

The evidence of complexity aversion for investors on *Seeking Alpha* motivates examining whether, in aggregate, investors underreact to company-issued news releases with longer titles. One could argue that while the behavior demonstrated by *Seeking Alpha* investors likely affects individual trading activity, notably of less-sophisticated investors, the behavior may not matter for financial markets (Barber and Odean, 2008; Barber et al., 2009; Huddart et al., 2009; Loewenstein et al., 2016; Engelberg and Parsons, 2016). Headline length is unlikely to affect meaningfully the attention of sophisticated and algorithmic investors. However, if headline length does matter for investor attention, the studies highlighted in the first paragraph suggest headlines may affect firm and market outcomes.

I assemble a database of approximately 480,000 company-issued-earnings releases distributed via *PR Newswire* and *Business Wire* during 1988 to 2016. Note that I am no longer using *Seeking Alpha* data but rather earnings-press-release data. To isolate variation in title length that is unrelated to the event reported, I instrument title length with the length of the company's legal name.

I find a significant negative effect of title length on abnormal turnover. A standard-deviation increase in a title's length predicts 7% less announcement turnover and 9% fewer trades on the Nasdaq exchange. The magnitudes are comparable to the effects title length has on the attention of *Seeking Alpha* investors and thus, support the external relevance of the *Seeking Alpha* results. Also, the negative relations occur on the announcement day but not before, consistent with title length reducing initial attention to news. The effects are robust to splits on time and firm size, and directionally stronger for smaller firms.

Title length also negatively affects abnormal-announcement volatility, captured by the day's intraday-price range. A standard-deviation increase in a title's length predicts a 60-basis-points-tighter-intraday-price range. Less attention to complex news appears to reduce volatility, which contrasts with the interpretation of prior findings that complex 10-Ks correlate with greater post-filing volatility. The length of 10-K filings may be a measure of both complexity and event type.

I also find evidence of return underreactions for earnings news with longer titles. A standard-deviation increase in a title's length predicts a 15-basis-points return underreaction for negative news and a negligible underreaction to positive news. In other words, I find that stock prices do not fall as much for negative news with non-informatively longer titles. I determine negative news by comparing actual earnings to analyst expectations. The complexity effect may matter

more for negative news than for positive news if shorting constraints are binding. I provide statistical evidence of a reversal in the next three months, which is consistent with the instrument – company-name length – capturing non-informative variation in title length.

The market effects of complexity varies with how surprising the event is, analyst coverage of the firm, the sophistication of the investor base, and whether the release occurs in the morning or after the market closes. First, larger earnings surprises might grab the interest of investors and the media more than quiet news, reducing the relative importance of textual attributes for attention to the news. Similarly, analysts likely amplify the content of an earnings release, weakening the importance of textual attributes. Consistent with this logic complexity matters less for bigger earnings surprises and for firms with analyst coverage. Second, more-sophisticated investors should be less sensitive to complexity per the *Seeking Alpha* results. Consistent with this logic, I find that complexity matters less when more shares are held by institutional investors. Third, I show that *Seeking Alpha* investors are 20% more likely to read stock reports released in the morning relative to reports released in the afternoon. Consistent with this pattern in attention among *Seeking Alpha* investors, I find that underreactions to earnings announcements due to the complexity of headlines is greater for earnings announcements released after the market’s close.

The effect of complexity on volatile and busy days is ambiguous. On the one hand, on volatile and busy days, investors are more time constrained and inclined to skip complex news (Hirshleifer et al., 2009). On the other hand, textual attributes may matter less for attention on busy days if the benefits to reading news are relatively greater or the media lowers information processing costs by allocating more attention to financial news. Consistent with the latter reasoning, *Seeking Alpha* investors are more likely to read a report released on busy-news and high-VIX days. Also, I find the market effects of title length are greater on low-volatility and slow-news days. My results expand on the evidence of an “Ostrich Effect” provided by Loewenstein et al. (2016), who find that investors are less likely to login to their brokerage account on high-VIX days. Unlike ostriches, investor demand for news increases on high-VIX days, but like ostriches, investors seem to hide from negative returns by not logging into their accounts.

Given the effects titles have on attention to news, I next examine whether firms strategically choose titles. The results thus far suggest firms reporting positive news should write shorter titles. Instead, firm’s write significantly longer titles for positive news. However, firms reporting positive news do tend to use more abbreviations, such as “Q3” for “third quarter.” I find evidence consistent with this reasoning as firms are more likely to abbreviate such phrases for positive earnings surprises. A firm could also adjust the firm’s name in the headline by dropping parts of the name like “Inc.” or “Limited,” but I find no relation between the earnings surprise and the length of a firm’s name in the headline. Also, firm’s seem to be increasingly ignoring the advice to write shorter and simpler titles as press-release titles have been getting longer since 1988.

In addition to the aforementioned contributions, these results inform debates about how to

write an optimal headline. Some recommend short titles, while others recommend detailed titles.<sup>4</sup> Prior efforts to identify optimal title lengths lack identification and outcomes vary widely (*Facebook* headlines: 40 characters, *LinkedIn* blog post headlines: 80-120 characters).<sup>5</sup> While title testing is becoming more popular, this paper is the first academic study to analyze title-testing data in any field.

## 2 *Seeking Alpha* and Title-Testing

Founded in 2004, *Seeking Alpha* has become the leading-crowdsourced-investment-research firm, owned by David Jackson and venture capital firms Benchmark, Accel, and DAG Ventures. The platform is highly active, with 4 million registered investors and 85 million page views per month.<sup>6</sup> The community includes over 10,000 contributing analysts writing stock reports and 280,000 commenters. Stock reports cover a broad range of firms, including more than 4,000 small- and mid-cap stocks in the past year across a variety of sectors. Motivations for investors to write stock reports include, receiving quick feedback on theses from thousands of other investors and building public reputations. *Seeking Alpha*'s primary revenue source is online advertising.

The audience includes money managers, sell-side analysts, investment bankers, financial advisors, business leaders, entrepreneurs, and retail investors. Over 20% of the audience are financial professionals. Readers tend to be highly active investors. Over 50% of readers purchased stocks in the trailing 30 days. Almost 90% of unique visitors own securities. The readers also tend to be wealthy, with the highest percentage of readers, among any major finance website, managing portfolios with assets greater than \$50,000, \$100,000, \$250,000, \$500,000, and \$1,000,000.

*Seeking Alpha* hires editors to make sure stock reports are well written and not repetitive of prior published reports. The community generates 600 submissions a day and after editorial review, approximately 200 reports are published. The editors provide guaranteed monetary rewards for content: \$35 basic plus \$0.01 per page view, \$150 guaranteed minimum for must-read, and \$500 guaranteed minimum for top-idea. Analysts build a reputation and get feedback via comments to stock reports. The content produced is valuable. Chen et al. (2014) finds the sentiment of the *Seeking Alpha* stock reports and comments predict future returns. Only a small proportion of articles appear to be undisclosed promotions paid for by companies, and these are usually small drug companies (Moskowitz et al., 2017).

Starting January 3, 2016, *Seeking Alpha* began "title testing." An analyst may propose two titles for a stock report, and the editor assigned to review the stock report may propose a third title. The three titles are sampled with random assignment on investors signed up for alerts on the topic company. Figure 1 illustrates three example real-time alert emails sent during the title-

---

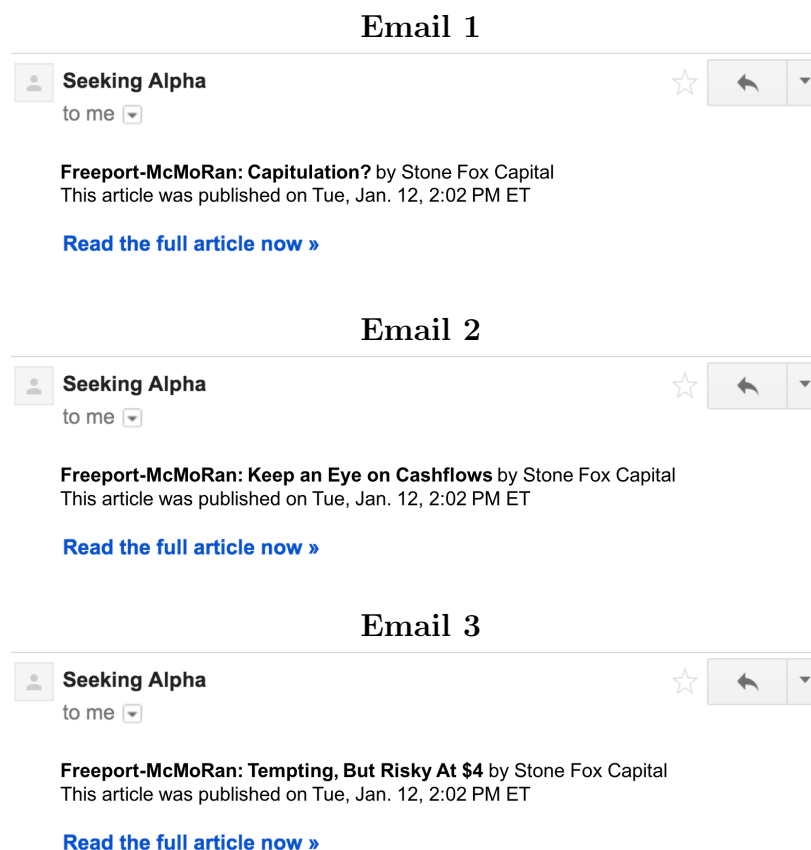
<sup>4</sup>Advice to write short headlines: <https://www.nngroup.com/articles/worlds-best-headlines-bbc-news/>. Advice to write long headlines: <https://www.poynter.org/2014/top-8-secrets-of-how-to-write-an-upworthy-headline/255886/>.

<sup>5</sup>See <https://blog.bufferapp.com/optimal-length-social-media>.

<sup>6</sup>See [http://seekingalpha.com/page/who\\_reads\\_sa](http://seekingalpha.com/page/who_reads_sa).

testing program for a single report. Each email alert contains the stock report’s title in bold, the author’s name, a time stamp, and a link to the full stock report. Notice that all elements other than the title are the same across alert emails and that none of the body of the stock report is included in the alert email. Also, there is only one title (news article) per email. In this example, the proposed titles randomly assigned to investors are “Freeport-McMoRan: Capitulation?,” “Freeport-McMoRan: Keep an Eye on Cashflows,” and “Freeport-McMoRan: Tempting, But Risky At \$4.” These alerts were sent to 8,373, 8,274 and 8,289 investors respectively. Within 30 minutes, the titles received 220, 154, and 133 clicks, respectively. *Seeking Alpha*’s title-testing algorithm chooses the title with the most page views after an interval of time as the winning title.

**Figure 1:** Example of email alerts sent during *Seeking Alpha*’s title-testing program. Each of the three alert emails contains one of the three proposed titles for a specific stock report written by Stone Fox Capital on Freeport-McMoRan. The emails were sent to 8,373, 8,274 and 8,289 investors tracking Freeport-McMoRan, respectively, with random assignment of titles to investors. The chief measure of attention is the number of investors who click the “Read the full article now” link to view the body of the stock report within 30 minutes of the alert email being sent. The titles received 220, 154, and 133 views within 30 minutes, respectively.



*Seeking Alpha* runs title testing for over 100 stock reports per day. *Seeking Alpha* does not run title testing on stock reports that editors chose as must-read or top-idea, because these are embargoed for paying subscribers for 24 hours, and the population of paying subscribers is too small at the moment. The ultimate audience of the stock report is global. *Seeking Alpha* emails



the report in a “Daily Investing Ideas” email to over 500,000 subscribers and publishes the report on major news feeds.

Whereas press releases are primarily distributed before the market opens or following the market’s close, the distribution of *Seeking Alpha* email alerts peaks during market hours between 10AM and 1PM EST. The mean click rate in 30 minutes on email alerts is 1.5%. The click rate is meaningfully higher for smaller and less popular companies.

As of December 2016, the *Wall Street Journal* and the *Financial Times* have not implemented title testing, making the *Seeking Alpha* data unique for studying how investors respond to the textual attributes of financial information.

### 3 Data Summary Statistics

**Table 1: Summary Statistics**

**Title-testing data from *Seeking Alpha*, January 4, 2016, to December 2, 2016**

Variable	Mean	SD	p25	p50	p75	N
<b>By Company</b>						
Market Capitalization (million)	11333	34497	352	1653	7525	1856
Log Market Capitalization	21.2	2.1	19.7	21.2	22.7	1856
Number of Real Time Alert Subscribers	4429	12254	620	1441	3797	1856
Length of Company’s Legal Name	21.3	7.8	16	21	27	1856
Log Length of Company’s Legal Name	3.0	0.4	2.8	3.0	3.3	1856
<b>By Title</b>						
Title Length (characters)	48.7	17.3	36	46	59	22629
Log Title Length	3.8	0.4	3.6	3.8	4.1	22629
Fraction of Negative Words in Title (%)	4.4	9.7	0	0	0	22629
Fraction of Positive Words in Title (%)	6.9	11.5	0	0	13	22629
Fraction of Positive-Negative Words in Title (%)	2.6	15.8	0	0	11	22629
% Yield (Views/Email in 30 minutes)	1.5	1.6	0.5	1.0	1.9	22629
<b>By Day</b>						
VIX (%)	16.1	4.0	13.4	14.7	18.2	232

**Press-release data from *PR Newswire* and *Business Wire*, 1988-2016**

Variable	Mean	SD	p25	p50	p75	N
<b>By Company</b>						
Market Capitalization (million)	1177	6278	33	120	498	15335
Log Market Capitalization	18.7	2.0	17	19	20	15335
Length of Company’s Legal Name	20.4	6.5	15	21	25	15335
Log Length of Company’s Legal Name	3.0	0.4	2.7	3.0	3.2	15335
<b>By Press Release</b>						
Title Length (characters)	66.9	24.1	50	62	78	480718
Log Title Length	4.1	0.3	3.9	4.1	4.4	480718
Log Turnover(t) x 100	-5.4	1.6	-6.4	-5.3	-4.3	480718
Log Intraday Price Range(t) x 100	6.5	7.4	2.5	4.6	8.2	480718
Log Return(t) x 100	-0.17	7.4	-3.0	0.0	3.1	480718

#### 3.1 *Seeking Alpha* Title-Testing Data

The title-testing program began January 3, 2016 and is ongoing. My sample period ends December 2, 2016, and includes 18,572 unique reports with 41,525 titles covering 3,573 unique securities.

I match each stock report with CRSP data. I discard observations that do not match with CRSP, including all stock reports covering OTC stocks. For this paper, I exclude stock reports about macroeconomic events, ETFs, mutual funds, REITS, currency, retirement, and commodities. An attraction to shorter, simpler, and more negative headlines also holds for investors following these other assets. I limit the sample to stock reports released on non-holiday weekdays. This approach reduces the sample to 9,944 unique stock reports with 22,623 titles for 1,856 companies.

For each stock report, the data include the titles associated with the stock report, the number of emails sent by title, the timestamp of the emails (all three emails sent at same time), and an identifier for whether the title was from the analyst (original or alternative) or from the editor. The page-view data by title capture the number of clicks on a specific title in the 30-minutes following an email alert. For January and February 2016, I have data on page views for all 30-minute intervals in the 24 hours following the time an email alert is sent. For this limited period, I also have data on the number of email alert subscribers who read the full stock report, measured as the number who scrolled to the end of the report within 24 hours.

Each of the two-or-three titles is not equally likely to have the most page views ex-ante. Appendix A.10 Table 20 shows the analyst's alternative title is less likely to have the most page views and is slightly longer and more positive when compared to the analyst's original title. The editor's title is generally shorter and more positive in tone than the analyst's original title and is less likely to receive the most page views, even though editors are likely more experienced at writing titles. I include dummy variables reflecting whether a title is the author's original, author's alternative, or editor's title in all regressions using *Seeking Alpha* data. The results in this paper hold if I discard the editor's title.

I supplement the data by gathering the body of stock reports and all of the comments. The comment data help gauge the sophistication of investors paying attention to the stock. I also collect data on the author (analyst), including the number of years as an author, the number of followers, the number of published stock reports, whether self-reports as a finance professional, and whether self-reports having an MBA. I gather similar data for those investors that comment on stock reports.

In Table 1, the summary statistics show that the median firm has a market capitalization of \$1.6 billion and has 1,441 investors signed up for alert emails. The median title has 1% of email recipients click in the first 30 minutes. The table also shows that the sentiment of titles tends to be more positive than negative as at the 75-th percentile 13% of a title is positive and 0% is negative. The sentiment statistics also suggest that the sentiment of titles is difficult to measure, which may attenuate my results.

## 3.2 Earnings-Press-Release Data

Using *Seeking Alpha*-title-testing data, I show investors exhibit strong attractions to short, simple, and negative titles. I cannot use the *Seeking Alpha*-title-testing data to identify the effect of title

length on market reactions to news. Instead, I transition to a new empirical setting and examine the effect of the lengths of headlines of company-issued-earnings-press releases on market reactions to the releases. I focus on earnings announcements as I can measure the earnings surprises and the announcements occur regularly for all firms, are prescheduled, and provide important information to investors.

I collect press-release titles from *PR Newswire* and *Business Wire* for the years 1988 to 2016. These two newswires are the primary means of distributing news for public companies. I match earnings releases with market data from CRSP based on the date and company’s ticker symbol used in the press release. I exclude over-the-counter stocks as these stocks are not included in CRSP. I match earnings announcements with I/B/E/S estimates to capture analyst expectations. I keep firms without analyst coverage. I also match earnings announcements with the latest Compustat last-twelve-months firm characteristics released at least 6 months prior to the current earnings announcement.

The final sample contains 480,718 earnings announcements from 15,335 firms for the years 1988 to 2016. The 25th percentile firm has 10 releases, and the 75th percentile firm has 45 releases. Approximately 60% of earnings announcements have analyst earnings estimate data available from I/B/E/S. The median company has a market capitalization of 120 million, which is considerably smaller than the median firm in the *Seeking Alpha* data of \$1.6 billion.

## 4 Empirical Results

In all regressions, unless otherwise noted, variables are standardized for ease of comparability.

### 4.1 Complexity and Attention

I measure the complexity of titles by calculating the title’s length in characters, number of words, word length, and word usage frequencies. Title length is a function of the number of words and the average word length. Longer words tend to be more difficult to understand, and more words require greater synthesis to interpret. Less common words may also be more challenging to understand. I measure a word’s usage frequency by counting the number of times the word appears in *Seeking Alpha* headlines from 2006 to 2015, which predates the title-testing data.<sup>7</sup> I then divide by the total number of words in titles for that period. I exclude company names when calculating usage frequencies.

Table 2 shows the following empirical specification: regression of page views by title on measures of complexity by title. Page views are measured over the 30 minutes following an alert email, capturing variation in the attention of highly engaged investors who indicated an interest

---

<sup>7</sup>The results are similar if word frequency is calculated using the frequency of words in 10-K filings from 1994 to 2014.

in the firm’s news by signing up for email alerts:

$$\text{Views/Emails}_{i,j} \sim \beta \text{Title Complexity}_{i,j} + \alpha_j + \epsilon_{i,j}$$

Because the three titles  $i$  for stock report  $j$  are randomly assigned,  $\beta$  captures the average effect of textual attributes on attention. The  $\alpha_j$  is the stock report fixed effect, which uniquely holds fixed the event as well as the firm, author, event, and date. Random assignment of titles to investors makes the distribution of omitted characteristics of the audience similar across titles in expectation. I cluster standard errors by stock report, and the significance of the results is robust to other forms of clustering. Throughout the paper I standardize regression variables for ease of comparing magnitudes.

To illustrate variation in title length within a stock report holding the context fixed, consider the following two titles.

- Bank of America is an attractive long term investment
- Bank of America is an attractive investment if your horizon is longer than a year

The first title received 195 page views in 30 minutes. The second title received 159 page views in 30 minutes.

Table 2 regression (1) shows a significant positive relation between title length and attention. However, regression (2), which includes the stock report fixed effect, shows a significant negative relation between title length and page views. The flip of the sign shows that the unconditional relation is biased by omitted variables related to the event, author, date, or firm. For example, for a larger company, the fraction of investors that access a stock report is smaller and the titles tend to be shorter, which together would create a positive bias.

Using the stock-report fixed effect, a standard-deviation increase in title length predicts 12%-fewer page views. This magnitude compares to the relation between the VIX and activity on *Seeking Alpha*. Appendix A.9 Table 19 shows that, within a firm, investors pay 8% more attention to stock reports released on days when the VIX is a standard-deviation higher. Also, this magnitude characterizes the behavior of investors who have indicated an interest in the news by signing up for alert emails. Less-interested investors may have a higher sensitivity to title length. The within- $R^2$  is 6%, suggesting title length is an important explanatory variable for attention.<sup>8</sup>

Table 2 regression (3) shows a negative relation between attention and both the average length of words in the title and the number of words. This result suggests that investors prefer simpler titles with fewer words and shorter words. Regression (4) shows a negative relation between the frequency of the least common word (relative to all past titles from 2006 to 2015) and attention. The negative relation suggests less-common words attract more attention. These results suggest unusual information attracts attention, but measures of length repel attention.

---

<sup>8</sup>I conduct 10,000 bootstrap out-of-sample tests to evaluate the out-of-sample explanatory power of this relation. The model is estimated using 50% of the stock reports, regressing within-stock-report page views on within-stock-report variation in title length. Testing the model shows 95% of  $R^2$  estimates are between 5% to 7%.

The negative relation is not driven only by titles that are very different in length or content. Figure 2 shows a clear negative relation in levels between within-stock-report page views and within-stock-report title length. Appendix A.1 and A.2 confirm that the negative relation can be identified even when the sample is restricted to titles that are at most 3 characters different in length and when the sample is restricted to titles with low- and high-word overlap. These findings speak to the robustness of the relations and suggest that differences in information content are not driving the results.

**Table 2:** Regressions of investor attention on title complexity, using *Seeking Alpha*-title-testing data. The attention measure is the fraction of email alerts sent that investors click on, calculated by randomly assigned title. Title length is measured in characters. Title length is decomposed into average word length and number of words. The frequency of the least common word in the title (excludes firm name) is measured relative to word usage in *Seeking Alpha* titles from 2006 to 2015. Regressions (2) to (4) include an article fixed effect, which holds the event, firm, author, and date fixed. All variables are standardized, and standard errors are clustered by article.

	Page Views/Emails			
	(1)	(2)	(3)	(4)
Title Length	0.04***	-0.11***		
	(0.01)	(0.00)		
Average Word Length			-0.05***	-0.05***
			(0.00)	(0.00)
Number of Words			-0.11***	-0.11***
			(0.00)	(0.00)
Frequency of Least Common Word in Title				-0.01***
				(0.00)
Article FE	No	Yes	Yes	Yes
Adjusted R <sup>2</sup>	0.03	0.91	0.91	0.91
Within R <sup>2</sup>	.	0.06	0.06	0.06
Num. Articles	9947	9947	9947	9947
Observations	22629	22629	22629	22629

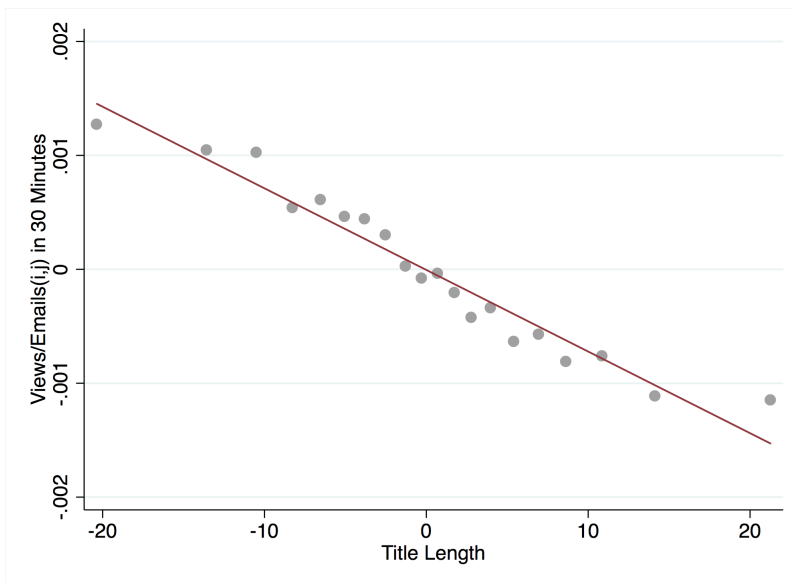
## 4.2 Sentiment and Attention

This section examines, using the *Seeking Alpha* title-testing data, whether the sentiment of headlines matters for investor attention. No prior studies have cleanly documented a relation between title sentiment and attention holding the event, firm, author, and date fixed. On the one hand, more positive titles may garner more attention due to confirmation bias, assuming investors on *Seeking Alpha* are mostly holding long positions. On the other hand, investors may be less willing to pay the costs of engaging with positive news. Consistent with the latter reasoning, I find a positive relation between the negativity of titles and attention.

To measure the sentiment of titles, I count the number of positive and negative words in a title using Bill McDonald’s lexicons (lists) of positive and negative words (Loughran and McDonald, 2011). The net sentiment of a title is the number of positive words less the number of negative words divided by the number of words in the title. I exclude the firm’s name from these calculations.

Table 3 regression (1) does not include an article fixed effect and suggests the raw correlation between net sentiment and attention is not significantly different from zero. In contrast, using the

**Figure 2:** Binned scatter plot illustrating the relation between title length in characters and readership (views/emails) in the first 30 minutes following the time alert emails are sent. Title length and yield are demeaned at the stock report level, so that “0” denotes the average yield and average title length for a single stock report. The bin scatterplot divides the sample by title length into equal-sized groups. The mean yield is then determined for each group and plotted as a point.



article fixed effect, regression (2) shows that positive sentiment reduces attention. This difference suggests omitted variables related to the firm, date, and event bias the results. For example, larger firms tend to have less positive sentiment and lower readership yields, inducing a positive relation between net sentiment and attention. The magnitude of the coefficient suggests that a standard-deviation increase in a title’s net sentiment predicts a 2% increase in page views. The magnitude is likely attenuated as measuring sentiment is difficult, especially for short titles. In regression (5), I decompose net sentiment into the fraction of positive and negative words. Positive words are negatively related to attention, while negative words are positively related to attention.

Because measures of complexity and sentiment are correlated, I control for title length in regression (3). Controlling for length reduces the magnitude of the coefficient on *Net Sentiment*, as estimated in regression (2), by 20%, suggesting that studies of sentiment in finance should account for textual complexity.

### 4.3 Cognitive Abilities and Attention

I now examine whether the sensitivities of investor attention to title length and sentiment are stronger for less-sophisticated investors. I would prefer to use a direct measure of the sophistication of investors signed up for alert emails. However, I do not have these data. Instead, I take advantage of individual data revealed in the actively used comment sections of stock reports to characterize the sophistication of the topic company’s followers. About 60% of comments are

**Table 3:** Regressions showing relation between title sentiment and investor attention, using *Seeking Alpha*-title-testing data. Title sentiment is measured using the number of words in the title (excluding firm name) that appear in the Bill and McDonald lexicons of positive and negative words. The number of positive and negative words is then scaled by the number of words in the title (excluding firm name). Regressions (2) to (4) include a stock report fixed effect. All variables are standardized, and standard errors are clustered by stock report.

	Views/Emails				
	(1)	(2)	(3)	(4)	(5)
Net Sentiment	0.002 (0.008)	-0.021*** (0.003)	-0.017*** (0.003)	-0.015*** (0.00)	
Net Sentiment x Numerical Comments				0.008** (0.004)	
Positive Sentiment					-0.017*** (0.003)
Negative Sentiment					0.007** (0.003)
Title Length			-0.106*** (0.003)	-0.106*** (0.004)	-0.105*** (0.004)
Article FE	No	Yes	Yes	Yes	Yes
Adjusted R <sup>2</sup>	0.02	0.90	0.91	0.91	0.91
Within R <sup>2</sup>	0.00	0.00	0.07	0.07	0.07
Num. Articles	9947	9947	9947	9538	9947
Observations	22629	22629	22629	21731	22629

made on the day of the report’s publication, and 20% of comments are made on the following day. Comments from investors who have contributed a stock report on *Seeking Alpha* previously may be more sophisticated on average than investors who have never contributed a stock report. Also, more-sophisticated investors may be more numerical and write longer comments. More numerical comments have more digits relative to total characters. Lastly, investors who self-report holding a finance job and having an MBA or CFA are likely more sophisticated, and I find these investors write longer and more numerical comments.

Table 4 regression (3) shows that when more comments come from investors who have never contributed a stock report, the negative sensitivity to title length is stronger. Regressions (4) and (5) show that the sensitivity to title length appears to be weaker when comments are more numerical and longer. Regressions (6) and (7) show a weaker sensitivity when more of the audience appear to be finance professionals and investors who self-report having an MBA or CFA. I also look at variation in the sensitivity with firm size and popularity since less-sophisticated investors are more likely to be aware of larger and more popular firms ( $\rho = 0.7$  between size and popularity) (Barber and Odean, 2008). Regressions (1) and (2) show that when the company is larger and more popular, the audience seems to be more sensitive to title length. These results suggest complexity aversion is lower for more-sophisticated investors. Similarly, Table 3 regression (4) shows that more-sophisticated investors appear to be less sensitive to the sentiment of titles.

Another way to gauge whether sophisticated investors have less complexity aversion is to examine the reading intensity of those who click on a relatively longer title. The underlying report is the same, and because of randomization, characteristics of the audience receiving each title are similar. Thus, any differences in read-to-end rates (number who scroll to bottom of the

**Table 4:** Regressions showing how the sensitivity of investor attention to title complexity varies in the cross-section of investor sophistication, using *Seeking Alpha*-title-testing data. Title length is measured in characters. Title length is interacted with firm size, number of followers (number of *Seeking Alpha* investors signed up to receive email alerts for the topic company), the fraction of article comments from non-analysts (*Seeking Alpha* investors who have never written a stock report), the average length of comments, the fraction of characters in comments that are digits (“numerical comments”), the fraction of comments from users who self-report having a finance job, and the fraction of users who self-report having an MBA or CFA. Each regression includes an article fixed effect, which holds the event, firm, author, and date fixed. All variables are standardized, and the standard errors are clustered by stock report.

	Page Views/Emails						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Title Length	-0.11*** (0.00)	-0.11*** (0.00)	-0.11*** (0.00)	-0.11*** (0.00)	-0.11*** (0.00)	-0.11*** (0.00)	-0.11*** (0.00)
Title Length x Market Cap		-0.02*** (0.00)					
Title Length x Followers			-0.02*** (0.00)				
Title Length x Fraction Comments Non-Analyst				-0.01*** (0.00)			
Title Length x Avg. Comment Length					0.01* (0.00)		
Title Length x Numerical Comments						0.01** (0.00)	
Title Length x Fraction Comments from Finance Professionals							0.03*** (0.01)
Title Length x Fraction Comments form MBA or CFA							0.02*** (0.00)
Article FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R <sup>2</sup>	0.91	0.91	0.91	0.91	0.91	0.91	0.91
Within R <sup>2</sup>	0.07	0.07	0.07	0.07	0.07	0.07	0.07
Num. Articles	9538	9538	9538	9538	9538	9538	9538
Observations	21725	21731	21731	21731	21731	21725	21725

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

article) predicted by title length must be due to a difference in the composition of readers who viewed the title, a selection effect. Table 5 regression (2) shows the relation between title length and the number of investors who read the full stock report in the 24 hours following the alert email, controlling for the number of investors who click on the stock report. The data is only available for January and February of 2016. The coefficient on title length is positive and highly significant. A standard-deviation increase in title length predicts a 4% higher read-to-end rate. This finding is consistent with more-sophisticated investors being less sensitive to title length and generally more thorough in information acquisition.

#### 4.4 Instrument for Title Length

At this point, I still cannot formally conclude title complexity affects attention. Complexity is correlated with other attributes, such as sentiment. Also, more complex titles might provide more information, reducing the need to click to read the stock report. More informative titles may also lead to better matches between an investor’s interests and the content of the report. In the previous tests, I did split the sample to try to mitigate information differences. Again, I showed that the negative relation with title length has a similar magnitude when estimated using



**Table 5:** Regressions showing the relation between the read-to-end rate (the fraction of investors who read the full stock report conditional on viewing the report) and title length, using *Seeking Alpha*-title-testing data. Regression (2) includes an article fixed effect, which holds the underlying stock report the same. Due to data availability, the read-to-end rate sample period is January 4, 2016, to February 29, 2016. All variables are standardized, and the standard errors are clustered by stock report.

	Read-to-End Rate	
	(1)	(2)
Title Length	0.21*** (0.02)	0.17*** (0.01)
Article FE	No	Yes
Adjusted R <sup>2</sup>	0.04	0.72
Within R <sup>2</sup>	0.04	0.05
Number of Articles	1687	1687
Observations	4008	4008

titles that are 3 or fewer characters different and pairs of titles with 90% word overlap.

To shut down the information story, I instrument title length with the length of a company’s legal name (not the length of the name in the title). The relevance condition is satisfied as *Seeking Alpha* headlines generally include the topic company’s name and longer names predict longer titles. The length of a company’s name is a significant component of a title’s length, making up 21% of a title’s length on average. A monotonicity condition must hold – all companies with longer names have on average longer titles because of their longer names. Some company names are defiers in the sense that the companies have long legal names but commonly used abbreviations (e.g., “International Business Machines” uses “IBM”). Excluding any titles with abbreviated names does not alter the relations.

The exclusion restriction is almost surely satisfied. A company’s legal name is chosen in the past, and thus the choice is unrelated to the new information discussed in a stock report. Although a wide variety of firm names is possible, the variety of plausible lengths is smaller. Company-name length may be correlated with the company’s industry; for example, “ABC pharmaceuticals” is longer than “XYZ energy.” I include SIC4-industry-by-year fixed effects in regressions to control for differences in name length across industries and industry trends.

I examine in Table 6 the correlations between company-name length and firm characteristics. I use Compustat’s 1988 to 2016 annual files. I match the Compustat data with data from CRSP and ownership data from Thomson. Regression (1) shows a significant negative relation between a firm’s book value of assets and company-name length. Therefore, I orthogonalize company-name length to firm size by regressing name length on a third-degree polynomial of log assets and log debt. I control for SIC4-by-year industry fixed effects and fixed effects for the date of the financial release. I take the residuals from the regression as “Adjusted Name Length.” Regressions (2) through (8) show that adjusted-company-name length is unrelated to other firm characteristics including market capitalization, institutional ownership (shares held by 13-F filers), revenues, gross margins, the company’s market beta, and age. Also, regression (6) confirms that the length of the company’s name is unrelated to earnings surprises. One can see the within- $R^2$  goes to 0.00 in these regressions. The lack of a meaningful relation between company-name length and

a variety of firm characteristics suggests company-name length provides variation in title length that is unrelated to characteristics of the firm, conditional on a company’s size and industry.<sup>9</sup>

**Table 6:** Using the Compustat 1988 to 2016 annual files, this table shows that log company-name length after adjusting for firm size is mostly unrelated to firm characteristics. “Adj Name Length” is determined by taking the residuals from a regression of log company name length on a cubic polynomial of log assets and log debt, while controlling for SIC4-by-year industry fixed effects and fixed effects for the date financials are released. Institutional ownership is the fraction of shares held by 13-F filers as of December 31 from Thomson. Earnings surprise is the actual EPS versus the median analyst estimate scaled by the stock price. Market beta is estimated using the past 5 years of monthly returns. The sample only includes firms with a December 31st year end. Standard errors are clustered by firm and the date financials are released. All variables are winsorized at the 1% level and standardized.

	Name Length (1)	(2)	(3)	(4)	Adj Name Length			
					(5)	(6)	(7)	(8)
Total Assets	-0.14*** (0.00)							
Market Capitalization		-0.02 (0.02)						
Institutional Ownership			-0.00 (0.02)					
Revenue				-0.01 (0.02)	-0.04 (0.03)			
COGS					0.03 (0.03)			
Earnings Surprise						0.00 (0.05)		
Market Beta							0.01 (0.01)	
Age (since IPO)								-0.02 (0.02)
Adjusted R <sup>2</sup>	0.07	-0.13	-0.12	-0.13	-0.14	-0.12	-0.11	-0.11
Within R <sup>2</sup>	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Observations	146020	146020	146020	136043	136043	79302	125043	49966

Table 7 regression (1) shows the first-stage regression of title length on the length of a company’s legal name. The coefficient on company-name length is positive and has a F-statistic of 192, exceeding the threshold of 10 recommended by Stock and Yogo (2005). The quadratic term is negative and significant, suggesting that the relation increases at a decreasing rate. In other words, companies with very long legal names are less likely to use their full names, but these companies still tend to have longer titles because of their longer names. I use both terms as instruments in all instrumental variables regressions. Overall, the first stage result suggests company-name length is highly relevant for title length. Regression (2) regresses the length of the topic company’s name as it appears in the title on the length of the legal name. Again, there is a strong positive relation and the  $R^2$  is 10%. I can either instrument the length of the full title or instrument the length of the firm’s name as it appears in the title and control for the length of other content. Either approach yields similar results, but to avoid the step of carving out a firm’s name from the title, which requires some processing, I instrument the length of the full title.

<sup>9</sup>For further validation of the instrument, see Appendix A.13.

Table 7 regression (4) shows the instrumental-variable (IV) relation between title length and attention. One cannot use the article fixed effect because I am exploiting variation in the length of a company’s legal name. The coefficient of -0.16 is not significantly different from the OLS coefficient of -0.10 in regression (3).

Table 7 regressions (5) and (6) examine the relation between title length and the read-to-end rate measure. The IV coefficient of 0.52 from regression (6) is not significantly different from the coefficient of 0.14 from regression (4). This result shows that the observed selection effect is driven by non-informative variation in title length, suggesting that investors who are less sensitive to title length are more thorough and sophisticated. The instrumental-variable results rule out the competing story that the negative relation between title length and attention occurs simply because longer titles are more informative.

**Table 7:** Regressions showing the relation between instrumented-title length and investor attention, using *Seeking Alpha*-title-testing data. The instrumental variable is the character length of the firm’s legal name. Regression (1) is the first-stage of title length regressed on company-name length. Regression (2) is an alternative first-stage, in which the length of the firm’s name as the name appears in the title is regressed on the length of the company’s legal name. Regression (3) and (4) show that title length affects the number of investors who click to view an email. Regression (4) is the instrumental variable regression. Regressions (5) and (6) show that title length affects the read-to-end rate, or the number of investors who read the full report conditional on viewing the report. Regression (6) is the instrumental regression, showing that investors who click on longer titles are also more thorough. All regressions include the number of emails sent by title, the firm’s market capitalization, and fixed effects for the industry (SIC4) and date. I only have read-to-end rate data for January and February 2016. Standard errors are clustered by stock report and all variables are standardized.

	(1) Title Length OLS	(2) Length Name in Title OLS	(3) Views/ Emails OLS	(4) Views/ Emails IV	(5) Read-to- End Rate OLS	(6) Read-to- End Rate IV
Length of Company’s Legal Name	0.15*** (0.01)	0.39*** (0.01)				
Length of Company’s Legal Name <sup>2</sup>	-0.04*** (0.01)	-0.09*** (0.01)				
Title Length			-0.10*** (0.01)	-0.16*** (0.06)	0.14*** (0.02)	0.52** (0.27)
Market Capitalization			-0.16*** (0.01)	-0.16** (0.01)	-0.26*** (0.04)	-0.24*** (0.05)
Number of Email Alert Subscribers			-0.56*** (0.01)	-0.56** (0.01)	-0.17*** (0.05)	-0.14*** (0.05)
% Adjusted R <sup>2</sup>	0.12	0.37	0.58	.	0.30	.
% Within R <sup>2</sup>	0.01	0.10	0.33	.	0.12	.
Number of Articles	9944	9944	9944	9944	1686	1686
Observations	22623	22623	22623	22623	4006	4006

Another test is to split title length into length associated with content and length associated with the firm’s name. To do so, I remove words in the title that overlap with the firm’s name, the first letters of each word in a firm’s name, and the ticker symbol. If the information story were true, then only length associated with content should be negatively related to attention. If the complexity aversion story were true, then length associated with both content and the firm’s name should be negatively related to attention. Table 8 regression (5) shows that both

length associated with content and with firm name are negatively related to attention. In fact, a character change in length due to content or name has an equivalent negative effect, which strongly suggests that pure character length repels attention.

**Table 8:** This table helps rule out the information story - that longer titles have more information, reducing the need to read the article. Assuming the information story is true, we would expect to see that the number of a headline’s characters devoted to content is negatively related to attention and that the number of a headline’s characters devoted to a firm’s name (non-informative length) has no correlation with attention. To calculate characters in a headline related to content, I remove all words in the headline that are related to the firm’s name, acronym, and ticker symbol. Using *Seeking Alpha*-title-testing data, regression (5) shows that both the number of characters related to content and the number of characters related to the firm’s name are equally-negatively related to attention to the news. This result is not consistent with the information story but is consistent with the story that investors are complexity averse. All regressions include an article fixed effect, which holds the event, firm, author, and date fixed. Thus, variation in length of name comes from different presentations of the firm’s name in headlines within an article. All measures of length are in number of characters and not standardized. All variables are standardized. Standard errors are clustered by article.

	(1)	(2)	(3)	(4)	(5)
	Length of Content in Title	Views/ Emails	Views/ Emails	Views/ Emails	Views/ Emails
Title Length		-0.006*** (0.000)			
Length of Content in Title			-0.006*** (0.000)		-0.006*** (0.000)
Length of Firm Name in Title	-0.60*** (0.04)			-0.003*** (0.001)	-0.006*** (0.001)
Article FE	Yes	Yes	Yes	Yes	Yes
Adjusted R <sup>2</sup>	0.46	0.91	0.91	0.90	0.91
Within R <sup>2</sup>	0.02	0.06	0.05	0.00	0.06
Num. Articles	9944	9944	9944	9944	9944
Observations	22622	22622	22622	22622	22622

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

*Seeking Alpha* authors do not seem to be aware of the negative effect title length has on attention. Appendix Figure 3 shows, from 2009 to 2015, *Seeking Alpha* titles have been steadily increasing in length.

## 4.5 Title Length Affects Market Behavior

Although I have provided evidence that longer-instrumented titles reduce investor attention using *Seeking Alpha* data, I have not shown this behavior matters for market outcomes. One could argue that while headline attributes matter for the attention of *Seeking Alpha* investors, the results may not hold externally. This section shows that title length does seem to matter for market behavior.

I gathered 480-thousand earnings-press-release titles distributed by 15-thousand firms via *PR Newswire* and *Business Wire* during the period 1988 to 2016. The advantage of studying earnings announcements is that I know how surprising the news is relative to analyst expectations. Also, earnings events are pre-scheduled, regularly occurring, and important. I designate the first trading period following the announcement as  $t$ . Although each press release contains the exact date of

the release, I do not have the precise time of day the release is distributed. However, starting in 2003, companies have strong incentives to file an 8-K with the press release as an exhibit before an earnings conference call, which are typically held shortly after the press release is distributed. Thus, starting in 2003, I do have timestamps of the 8-K filings, which may approximate the timing of the press release. I find that determining the first trading period following the announcement, date  $t$ , is reliably accomplished by choosing the day of the date in the press release or the following business day, whichever has more volume.

Since I am no longer using *Seeking Alpha*-title-testing data, but rather company-issued press releases, I need variation in title length that is unrelated to the earnings surprise. I use the instrument – the length of a company’s legal name – to identify whether earnings-release-title length affects market outcomes. The first-stage regression using company-name length and press-release-title length is highly significant. Appendix A.4 Table 14 regressions (1)-(3) show the sensitivity of title length to the length of a company’s legal name is positive and highly significant. The F-statistic for the hypothesis that the instrument’s coefficient is zero is 1,338, exceeding the threshold of 10 recommended by Stock and Yogo (2005). A regression in levels would show a coefficient of 0.68, which suggests that a firm with a one character longer name has on average 0.7 more characters in the title. The coefficient may be less than one if companies do not include full legal names in press-release titles. For example, firms may not include “Limited” or “LLC.” I do not find that firms with longer names have meaningfully less content in titles, and controlling for the length of the non-name-related content of titles has no effect on results. The exclusion restriction is likely satisfied as the length of a company’s legal name is chosen in the past and is empirically uncorrelated with earnings surprise measures.<sup>10</sup>

I first examine the relation between instrumented-title length and turnover around earnings announcements. The *Seeking Alpha* results suggest longer titles receive less attention, and turnover should be positively related to attention. Turnover is the volume traded divided by the number of shares outstanding. To calculate abnormal turnover, I estimate a firm’s turnover sensitivities to the turnover of all stocks in CRSP. I use these estimated sensitivities to remove market-related changes in turnover. I regress abnormal turnover on announcement of earnings on the instrumented-title length, the firm’s market capitalization, and lags from  $t-2$  to  $t-10$  of turnover, intraday price ranges, volatility, and returns. Table 9 regression (1) shows a significant negative relation between instrumented-title length and turnover on the day of the earnings announcement. There is no such relation on days prior to the announcement. A standard-deviation increase in title length predicts 7%-less announcement turnover, which is 4% of a standard-deviation change in turnover on announcement. The 7%-magnitude is comparable to the 12% reduction in attention to *Seeking Alpha* stock reports documented previously.

Turnover is a function of the number of trades and the average size of trades. If longer titles

---

<sup>10</sup>Similar to DellaVigna and Pollet (2009), I measure a firm’s earnings surprise as the difference between reported EPS and the median analyst forecast of EPS from  $t-45$  to  $t-3$ , where  $t$  is the announcement date. I normalize the difference in actual and forecasted EPS by the stock price at  $t-3$ . Scaling by price reveals the magnitude of the surprise. A \$0.10 surprise is bigger for a \$1 stock than a \$10 stock.

reduce attention to news, then the number of trades should be lower. For about half of the firms, CRSP provides data on the number of trades on the NASDAQ exchange. Examining Table 9 regression (2), I find that a standard-deviation increase in length predicts about a 7% decline in the number of trades, consistent with length reducing attention.

Volatility and attention may be positively correlated.<sup>11</sup> I measure abnormal volatility in two ways. First, I use the the abnormal intraday price range, which I calculate by regressing a firm's intraday price range (day's high price minus day's low price) on the intraday price range of the CRSP value-weighted market index and using the betas to adjust a firm's daily spreads. Regression (3) in Table 9 shows that instrumented title length is negatively related to the abnormal intraday price range on earnings announcement days. Again, this relation does not hold on the days prior to the announcement or following the announcement. A standard-deviation increase in title length predicts a 60-basis-points decrease in the announcement day's trading range, which is 9.5% of a standard-deviation change in intraday price range on announcement. Second, I measure abnormal volatility as the square of risk-adjusted returns, calculated using the Fama-French Five Factor model plus the momentum factor. I estimate rolling factor betas for every firm in CRSP using daily data for the preceding year. A standard-deviation increase in title length predicts a 4% reduction in the abnormal volatility. These volatility results suggest less investor attention to news reduces market volatility on announcement.

While less attention to news due to longer titles leads to less turnover and less intraday volatility, it is not clear that longer titles should result in a price underreaction. On the one hand, algorithmic traders and sophisticated investors are likely unaffected by title length. On the other hand, the price implications of earnings announcements are still challenging to understand for even sophisticated investors who often publicly debate the values of firms. I calculate a firm's abnormal returns around the announcement date using rolling betas from the Fama-French Five Factor model plus momentum factor. Table 9 regression (5) suggests that there is a price underreaction to longer titles for negative news and less so for positive news. I define positive news as a strictly positive earnings surprise relative to median analyst estimates from I/B/E/S. The observation counts are lower because not all firms have analyst coverage. The coefficient on title length is 0.118, suggesting that for negative news the price does not fall as much. For positive news, the coefficient is -0.026 (-0.144+0.118), suggesting that the price does not rise as much, though the magnitude is negligible. A standard-deviation increase in title length predicts a 15-basis-points underreaction to negative news. Note that a standard-deviation change in title length is 24 characters. The relation is robust to including a variety of firm controls and controlling for the earnings surprises as shown in Appendix A.5 Table 15. The result is also robust to time splits and firm-size splits, though slightly larger for smaller firms.

Since the instrument captures non-informative variation in title length, the underreactions should reverse. Table 9 regressions (6) and (7) provide evidence of such a reversal. These two

---

<sup>11</sup>See Andrei and Hasler (2015). Also, the attention of less-sophisticated investors proxied for by activity on stock message boards is correlated with greater volatility (Antweiler and Frank, 2004).

regressions are in percentage terms rather than using standardized variables. Regression (6) shows the announcement effect in percentage terms. Regression (7) shows the 3-month cumulative return starting at  $t+1$ . The coefficient on title length is negative for negative news consistent with the delay due to less attention on announcement reversing completely.

## 4.6 Variation in the Effect of Length on Market Behavior

I now examine whether the market effects vary with proxies for constraints on investor attention. On the one hand, on busy-news days or high-VIX days, investors may be more time constrained and likely to skip news with more complex titles. This reasoning is supported by results from Hirshleifer et al. (2009), finding investors underreact more to earnings news on days with more competing earnings news. On the other hand, I find that investors are more focused on the markets on high-VIX days, and complex news may become more valuable to read. Appendix A.9 Table 19 shows that investor attention to *Seeking Alpha* stock reports is significantly higher when the VIX is higher. Regarding the market effects, I find evidence consistent with the latter reasoning. Table 10 regressions (1) and (2) show that investors are more sensitive to title length for news released on slow-news days (mostly Fridays) and low-VIX days. A standard deviation decrease in the VIX or decrease in the number of earnings announcements that day increases the complexity effect on turnover by 50%.

I also examine whether the market effects are stronger when there is less amplification of the news by the analyst community. If a company's earnings announcements are discussed promptly with investors, then textual attributes should matter less for awareness of the news. Consistent with this logic, Table 10 regression (6) shows that the market effects are weaker for companies with analyst coverage. Also, larger earnings surprises should gain more media attention and reduce the importance of textual attributes on attention. Consistent with this logic, regression (3) shows a reduction in the effect of title length on turnover when earnings surprises are bigger.

I next examine whether market effects are stronger in the afternoon than in the morning. The motivation is that Appendix A.7 shows *Seeking Alpha* investors appear more focused on news in the morning than in the afternoon. Consistent with the results thus far, Appendix A.8 shows investors are also more sensitive to title length in the afternoon than in the morning. Regarding the market effect, Table 19 regression (4) shows that investors may be more sensitive to title length in the afternoon.

Market effects should also be weaker when the investor base looks more sophisticated. More sophisticated investors should be less affected by textual attributes, as supported by the *Seeking Alpha* title-testing results. Consistent with this logic, regression (5) shows that the market effect is weaker when more of the shares outstanding are held by institutional investors required to file 13-F disclosures.

**Table 9:** Regressions showing the relation between instrumented-earnings-press-release-title length and market outcomes on announcement of earnings. Title length is instrumented with the length of the company’s legal name. The dependent variables in regressions (1) is abnormal turnover, or volume scaled by shares outstanding, taking out market volume exposure, (2) is abnormal number of trades on Nasdaq exchange scaled by shares outstanding, taking on market trade count exposure, (3) is the abnormal intraday stock price range (Day’s High Price - Day’s Low Price), taking out exposure to market intraday range, (4) is volatility, or the abnormal return squared, and (5) is the abnormal return, adjusted for the Fama-French Five Factor model plus momentum. The dependent variables in regressions (6) and (7) are in percentage terms, whereas all other variables are standardized. The dummy for positive news is one for positive EPS surprises, measured relative to median analyst expectations from  $t - 3$  to  $t - 45$ . All regressions include SIC4-by-year fixed effects to account for industry trends and an announcement day fixed effect. Other controls include lags of turnover, the intraday spread, absolute returns, and returns for t-2 to t-10. Variables are winsorized at the 1% level and then standardized. Standard errors are clustered by firm.

	Turnover (t) IV (1)	Trade Count (t) IV (2)	Intraday Range (t) IV (3)	Volatility (t) IV (4)	Return (t) IV (5)	Return (t) IV (6)	Return (t+1,t+90) IV (7)
Title Length (t)	-0.043*** (0.01)	-0.038*** (0.01)	-0.095*** (0.02)	-0.042*** (0.01)	0.118*** (0.03)	0.009** (0.002)	-0.012** (0.006)
Title Length x Positive Surprise (t)					-0.144*** (0.04)	-0.010*** (0.003)	0.010 (0.008)
Positive Surprise (t)					0.476*** (0.01)	0.033*** (0.001)	0.020*** (0.001)
Market Cap (t-3)	-0.053*** (0.00)	0.000 (0.00)	-0.142*** (0.00)	-0.037*** (0.00)	-0.030*** (0.00)	-0.002*** (0.000)	-0.006*** (0.001)
SIC-4 x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	417876	218360	417876	417876	263947	263947	261942

## 4.7 Are Firm’s Strategic with Headlines?

Despite evidence that longer titles receive less attention, Appendix Figure 3 shows that earnings-release titles are getting longer over time rather than shorter. This observation suggests firms may under-appreciate the effect title length has on investor attention to news.

Table 11 examines how the earnings surprise relates to title attributes within a firm. One would expect a firm promoting positive news to write concise titles and possibly abbreviate commonly used phrases like “first quarter” with “Q1.” Firms trying to minimize attention to negative earnings should instead write long titles with many non-informative words and tend not to use abbreviations.

Regression (1) shows a highly significant and meaningful positive relation between the earnings surprise and title length. Firms seem to increase title length for positive news, rather than shorten title length. This behavior is not strategic, as I found using *Seeking Alpha* data that longer and more positive titles receive less attention. While keeping negative words out of titles for negative news is consistent with results using *Seeking Alpha* data, since negative words grab investor attention, titles tend to be shorter which seems to increase attention to the news. Regression (2) shows that even though firms write longer titles for positive surprises, firms write shorter press releases. This result is consistent with positive news being less complex overall. Regression (3)



shows firms use positive words in titles to communicate positive earnings surprises.

There are a variety of parts of the title that firms reporting positive news may abbreviate. Regression (4) shows that firms do not adjust the length of their name in titles in relation to the earnings surprise. A firm with positive news may exclude “Inc.” from the firm’s name, for example, to shorten the overall title. Regressions (5) to (7) suggest firms do abbreviate other content. Regression (5) shows that firms are significantly more likely to abbreviate phrases like “first quarter” with “Q1” for positive news.<sup>12</sup> Regression (6) finds firms are significantly more likely to use the word “reports” for positive news, which is shorter than the word “announces.” Regression (7) finds firms are less likely to include a year (e.g. “2012”) in the title for a positive earnings surprise. Altogether, regressions (5) to (7) show firms are more likely to use abbreviations, use shorter words, and exclude unnecessary words for positive news. Nevertheless, positive titles tend to be longer. This finding suggests firms do try to minimize the non-informative content of titles when firms want to include more positive information in titles.

## 5 Conclusion

The results of this paper are of interest to a variety of literatures in finance examining investor behavior, market reactions to news, and strategic disclosures. Also, as no prior study provided casual-field evidence that non-informative title length repels attention, the results are also broadly interesting to other fields, such as marketing, communications, and psychology.

There are a number of advantages to using investors as subjects to examine the effects of textual attributes on attention. The investors I examine operate in high-stakes markets, likely with substantial money invested. Also, the investors I examine have a keen interest in the firm’s news they receive. Nevertheless, textual attributes affect their attention to news, and commonly enough to affect market behavior.

The *Seeking Alpha* title-testing data with randomization cleanly show that investors tend to skip over stock reports with longer and more positive titles. I rule out the possibility that longer titles are simply more informative, using company-name length as an instrumental variable for title length. The cognitive abilities of investors matter as less-sophisticated investors are more put-off by complex and positive titles. The large effect title length has on attention suggests that investors are complexity averse since titles are relatively short pieces of text.

The aversion to title length measured in the *Seeking Alpha* data motivates testing whether investors in aggregate underreact to news with longer titles. One would hope that markets with sophisticated investors and algorithmic traders would not be impacted by title length. Instead, using company-name length as an instrumental variable for title length, I find that longer-earnings-announcement titles lead to less trading, smaller intraday-price ranges, and return underreactions. The results are stronger on slow-news days, on low-VIX days, and in the afternoons, which

---

<sup>12</sup>Other phrases are “second quarter,” “third quarter,” and “fourth quarter.” Abbreviations are variations of “qtr” or “q1.”

contrasts with prior findings that suggest investors are more attention constrained during busy periods. Perhaps, while investors are more time constrained on busy days, investors are also more focused on the markets and the value of reading complex news increases. The effects are also stronger when the investor base is less sophisticated.

Firms seem to be unaware of these findings. Titles are getting longer over time, and firms write longer titles for positive news. However, firms are more likely to abbreviate common phrases for positive news. While short soundbites distributed through mediums like Twitter are becoming more popular, firms are on a different trend.

**Table 10:** Regressions showing how the effect of press-release-title length on earnings-announcement-market outcomes varies with the VIX, news intensity, analyst coverage, timing of the release (morning vs afternoon), fraction of shares held by 13F holders, and the absolute earnings surprise. The VIX is the log of the ratio of the VIX on day  $t$  to the average VIX of 17.8. News intensity is the z-score of news count in a given calendar year. Analyst coverage is 1 if analysts provided earnings estimates ahead of the event and zero otherwise. Afternoon is 1 if the earnings release came out after market close at 4PM EST. The fraction of 13F holders captures the fraction of total shares outstanding held by more sophisticated investors. The absolute earnings surprise captures how surprising the earnings release is relative to median analyst expectations of EPS from  $t - 45$  to  $t - 3$ . The instrument for title length is the length of the company's name in CRSP. Earnings-announcement-press-release data are from *PR Newswire* and *Business Wire*. Market data is from CRSP. Regressions include SIC4-by-year and date fixed effects. Other controls include a polynomial of firm market capitalization on day  $t - 1$  and 10 days of lags of turnover, intraday spread, squared returns, and returns. All variables are winsorized at the 1% level. Standard errors are clustered by firm and date, and all variables are standardized.

	Turnover (t)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Title Length	-0.045*** (0.014)	-0.044*** (0.014)	-0.015 (0.015)	-0.011 (0.011)	-0.041** (0.018)	-0.061*** (0.021)	-0.043*** (0.014)
Title Length x VIX	0.028*** (0.007)						
Title Length x News Intensity		0.021** (0.009)					
Title Length x Abs. Earnings Surprise			0.026* (0.014)				
Title Length x Afternoon				-0.061** (0.028)			
Title Length x Fraction 13F Holders					0.024* (0.013)		
Title Length x Analyst Coverage						0.037* (0.020)	
Title Length x Market Cap (t-3)							-0.014 (0.017)
Absolute Surprise			0.005*** (0.001)				
Fraction 13F Holders					0.049*** (0.004)		
Analyst Coverage						0.086*** (0.003)	
Market Capitalization	-0.054*** (0.002)	-0.054*** (0.002)	-0.054*** (0.003)	-0.050*** (0.003)	-0.071*** (0.003)	-0.066*** (0.002)	-0.052*** (0.002)
SIC-4 x Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date x Afternoon FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	421560	421560	263947	421560	421560	421560	421560

**Table 11:** Regressions showing how the earnings surprise relates to title length, press-release length, and usage of abbreviations in earnings-release titles. Earnings surprise is the actual earnings per share less median estimated earnings, of those made from  $t - 45$  to  $t - 3$ . The surprise is scaled by the stock price at  $t - 3$  and winsorized at the 1% level. All regressions include a firm and release-date fixed effect. Standard errors are clustered by firm, and all variables are standardized.

	Title Length (1)	Log Press Release Word Count (2)	Title Sentiment (3)	Log Length of Name in Title (4)	Qtr=1 Quarter=0 (5)	Reports=1 Announces=0 (6)	Year excluded=1 Year included=0 (7)
Earnings Surprise	0.09*** (0.01)	-0.03*** (0.01)	0.18*** (0.01)	0.00 (0.01)	0.02* (0.01)	0.03*** (0.01)	0.02** (0.00)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R <sup>2</sup>	0.48	0.95	0.39	0.89	0.65	0.80	0.68
Within R <sup>2</sup>	0.00	0.00	0.00	0.01	0.00	0.00	0.00
Number of Firms	8213	8213	11601	11601	11243	11469	11603
Observations	263947	124530	263947	261432	232278	245276	263975

## References

- Agarwal, Sumit, John C Driscoll, Xavier Gabaix, and David Laibson, 2009, The age of reason: Financial decisions over the life cycle and implications for regulation, *Brookings Papers on Economic Activity* 51–117.
- Ahern, Kenneth, and Denis Sosyura, 2015, Rumor has it: Sensationalism in financial media, *Review of Financial Studies* 28, 2050–2093.
- Andrei, Daniel, and Michael Hasler, 2015, Investor attention and stock market volatility, *Review of Financial Studies* 28, 33–72.
- Antweiler, Werner, and Murray Frank, 2004, Is all that talk just noise? The information content of internet stock message boards, *Journal of Finance* 59, 1259–1294.
- Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *The Journal of Finance* 61, 1645–1680.
- Bank, Matthias, Martin Larch, and Georg Peter, 2011, Google search volume and its influence on liquidity and returns of German stocks, *Financial Markets and Portfolio Management* 25, 239–264.
- Barber, Brad, Terrance Odean, and Ning Zhu, 2009, Do retail trades move markets?, *Review of Financial Studies* 22, 151–186.
- Barber, Brad M, and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785–818.
- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.
- Bloomfield, Robert, 2008, Discussion of: Annual report readability, current earnings, and earnings persistence, *Journal of Accounting and Economics* 45, 221–247.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2013, Salience and asset prices, *American Economic Review* 103, 623–628.
- Chakrabarty, Bidisha, and Pamela Moulton, 2012, Earnings announcements and attention constraints: The role of market design, *Journal of Accounting and Economics* 53, 612–634.
- Chemmanur, Thomas, and An Yan, 2009, Product market advertising and new equity issues, *Journal of Financial Economics* 92, 40–65.
- Chen, Hailiang, Prabuddha De, Yu Hu, and Byoung Hyoun Hwang, 2014, Wisdom of crowds: The value of stock opinions transmitted through social media, *Review of Financial Studies* 27, 1367–1403.

- Cohen, Lauren, Christopher Malloy, and Quoc Nguyen, 2015, Lazy prices.
- Corwin, Shane, and Jay Coughenour, 2008, Limited attention and the allocation in securities trading of effort, *Journal of Finance* 63, 3031–3067.
- Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In search of attention, *Journal of Finance* 66, 1461–1499.
- Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2015, The sum of all FEARS investor sentiment and asset prices, *Review of Financial Studies* 28, 1–32.
- Da, Zhi, Umit G. Gurun, and Mitch Warachka, 2014, Frog in the pan: Continuous information and momentum, *Review of Financial Studies* 27, 2171–2218.
- DellaVigna, Stefano, and Joshua Pollet, 2009, Investor inattention and Friday earnings announcements, *The Journal of Finance* 64, 709–749.
- Dougal, Casey, Joseph Engelberg, Diego Garcia, and Christopher Parsons, 2012, Journalists and the stock market, *Review of Financial Studies* 25, 640–679.
- Engelberg, Joseph, 2008, Costly information processing: Evidence from earnings announcements.
- Engelberg, Joseph, and Christopher Parsons, 2016, Worrying about the stock market: Evidence from hospital admissions, *The Journal of Finance* 71, 1227–1250.
- Fang, Lily, and Joel Peress, 2009, Media coverage and the cross-section of stock returns, *The Journal of Finance* 64, 2023–2052.
- Feng, Lei, and Mark Seasholes, 2005, Do investor sophistication and trading experience eliminate behavioral biases in financial markets?, *Review of Finance* 9, 305–351.
- Gabaix, Xavier, and David Laibson, 2005, Bounded rationality and directed cognition.
- Gabaix, Xavier, David Laibson, Guillermo Moloche, and Stephen Weinberg, 2006, Costly information acquisition: Experimental analysis of a boundedly rational model, *American Economic Review* 96, 1043–1068.
- Gervais, Simon, Ron Kaniel, and Dan Mingelgrin, 2001, The high-volume return premium, *The Journal of Finance* 56, 877–919.
- Giglio, Stefano, and Kelly Shue, 2014, No news is news: Do markets underreact to nothing?, *Review of Financial Studies* 27, 3389–3440.
- Green, T. Clifton, and Russell Jame, 2013, Company name fluency, investor recognition, and firm value, *Journal of Financial Economics* 109, 813–834.
- Grinblatt, Mark, S Ikaheimo, and Matti Keloharju, 2008, Are mutual fund fees competitive? What IQ-related behavior tells us.

- Grinblatt, Mark, Matti Keloharju, and Juhani Linnainmaa, 2009, Do smart investors outperform dumb investors?
- Grinblatt, Mark, Matti Keloharju, and Juhani Linnainmaa, 2011, IQ and stock market participation, *Journal of Finance* 66, 2121–2164.
- Grullon, Gustavo, George Kanatas, and James P. Weston, 2004, Advertising, breadth of ownership, and liquidity, *Review of Financial Studies* 17, 439–461.
- Hartzmark, Samuel, and Kelly Shue, 2015, A tough act to follow: Contrast effects in financial markets.
- Hillert, Alexander, Heiko Jacobs, and Sebastian Müller, 2014, Media makes momentum, *Review of Financial Studies* 27, 3467–3501.
- Hirshleifer, David, Sonya Seongyeon Lim, and Siew Hong Teoh, 2009, Driven to distraction: Extraneous events and underreaction to earnings news, *Journal of Finance* 64, 2289–2325.
- Hou, Kewei, Lin Peng, and Wei Xiong, 2009, A tale of two anomalies: The implications of investor attention for price and earnings momentum.
- Huddart, Steven, Mark Lang, and Michelle Yetman, 2009, Volume and price patterns around stock’s 52-week highs and lows: Theory and evidence, *Management Science* 55, 16–31.
- Hwang, Byoung-hyoun, and Hugh Hoikwang Kim, 2016, It pays to write well.
- Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2009, Attention allocation over the business cycle.
- Kahneman, Daniel, 1973, *Attention and Effort* (Prentice-Hall, Inc.).
- Lawrence, Alastair, 2013, Individual investors and financial disclosure, *Journal of Accounting and Economics* 56, 130–147.
- Li, Feng, 2008, Annual report readability, current earnings, and earnings persistence, *Journal of Accounting and Economics* 45, 221–247.
- Li, Jun, and Jianfeng Yu, 2012, Investor attention, psychological anchors, and stock return predictability, *Journal of Financial Economics* 104, 401–419.
- Loewenstein, George, Duane Seppi, Nachum Sicherman, and Stephen Utkus, 2016, Financial attention, *Review of Financial Studies* 29, 863–897.
- Loh, Roger, 2010, Investor inattention and the underreaction to recommendations, *Financial Management* 39, 1123–1252.
- Lou, Dong, 2014, Attracting investor attention through advertising, *Review of Financial Studies* 27, 1797–1829.

- Loughran, Tim, and Bill McDonald, 2011, When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks, *Journal of Finance* 66, 35–65.
- Loughran, Tim, and Bill McDonald, 2014, Measuring readability in financial disclosures, *Journal of Finance* 69, 1643–1671.
- Loughran, Tim, and Bill McDonald, 2017, The use of EDGAR filings by investors, *Journal of Behavioral Finance* 18, 231–248.
- Michaely, Roni, Amir Rubin, and Alexander Vadrashko, 2016, Are Friday announcements special? Overcoming selection bias, *Journal of Financial Economics* 122, 65–85.
- Miller, Brian, 2010, The effects of reporting complexity on small and large investor trading, *Accounting Review* 85, 2107–2143.
- Moskowitz, Tobias J, Shimon Kogan, and Marina Niessner, 2017, Fake News in Financial Markets, *Working Paper* .
- Niessner, Marina, 2014, Strategic disclosure timing and insider trading.
- Peng, Lin, 2005, Learning with information capacity constraints, *The Journal of Financial and Quantitative Analysis* 40, 307–330.
- Peng, Lin, and Wei Xiong, 2006, Investor attention, overconfidence and category learning, *Journal of Financial Economics* 80, 563–602.
- Stock, James, and Motohiro Yogo, 2005, Testing for weak instruments in linear IV regression, in *Identification and Inference for Econometric Models*, 80–108 (Cambridge University Press).
- Tetlock, Paul, 2007, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance* 62, 1139–1168.
- Tetlock, Paul, Maytal Saar-Tsechansky, and Sofus MacSkassy, 2008, More than words: Quantifying language to measure firms’ fundamentals, *Journal of Finance* 63, 1437–1467.
- Tetlock, Paul C., 2011, All the news that’s fit to reprint: Do investors react to stale information?, *Review of Financial Studies* 24, 1481–1512.
- You, Haifeng, and Xiao-jun Zhang, 2009, Financial reporting complexity and investor underreaction to 10-k information, *Review of Accounting Studies* 14, 559–586.
- Yuan, Yu, 2008, Attention and trading.
- Yuan, Yu, 2015, Market-wide attention, trading, and stock returns, *Journal of Financial Economics* 116, 548–564.



# A Appendix

## A.1 Small Differences in Title Length

**Table 12:** Regressions showing that the negative relation between title length and attention is identifiable even when titles have very similar character lengths, using *Seeking Alpha*-title-testing data. I restrict the sample to articles with only two titles. I then restrict the sample to pairs of titles that have a character difference less than or equal to the max specified in the table. For example, regression (2) only includes titles that are two-or-fewer-characters different in length. Spaces count in title length. All regressions include an article fixed effect, which holds the event, firm, author, and date fixed. Standard errors are clustered by stock report, and all variables are standardized.

	Views/Emails			
	(1)	(2)	(3)	(4)
Title Length	-0.11 (0.28)	-0.16 (0.13)	-0.18** (0.08)	-0.18*** (0.06)
Max Character Difference in Length	1	2	3	4
Article FE	Yes	Yes	Yes	Yes
Adjusted R <sup>2</sup>	0.91	0.91	0.91	0.91
Within R <sup>2</sup>	0.00	0.00	0.00	0.01
Num. Articles	577	914	1266	1632
Observations	1154	1828	2532	3264

## A.2 Very Similar, Very Different Titles

**Table 13:** Regressions showing that the negative relation between title length and attention is identifiable even when the word overlap of titles for a report is high, medium, or low, using *Seeking Alpha*-title-testing data. Title length is measured in characters and includes space characters. I restrict the sample to articles with only two titles. I then restrict the sample by word overlap. Word overlap is measured as the fraction of title length that is shared across the two titles compared. High overlap implies that enough words overlap so that at least 90% of the characters are the same across the two titles. Low overlap titles have less than 10% of characters the same. All regressions include an article fixed effect, which holds the event, firm, author, and date fixed. Standard errors are clustered by stock report. All variables are standardized.

	Views/Emails		
	(1)	(2)	(3)
Title Length	-0.12*** (0.05)	-0.11*** (0.01)	-0.11*** (0.02)
Word Overlap	High	Middle	Low
Article FE	Yes	Yes	Yes
Adjusted R <sup>2</sup>	0.92	0.92	0.90
Within R <sup>2</sup>	0.02	0.06	0.08
Num. Articles	344	6268	410
Observations	688	12536	820

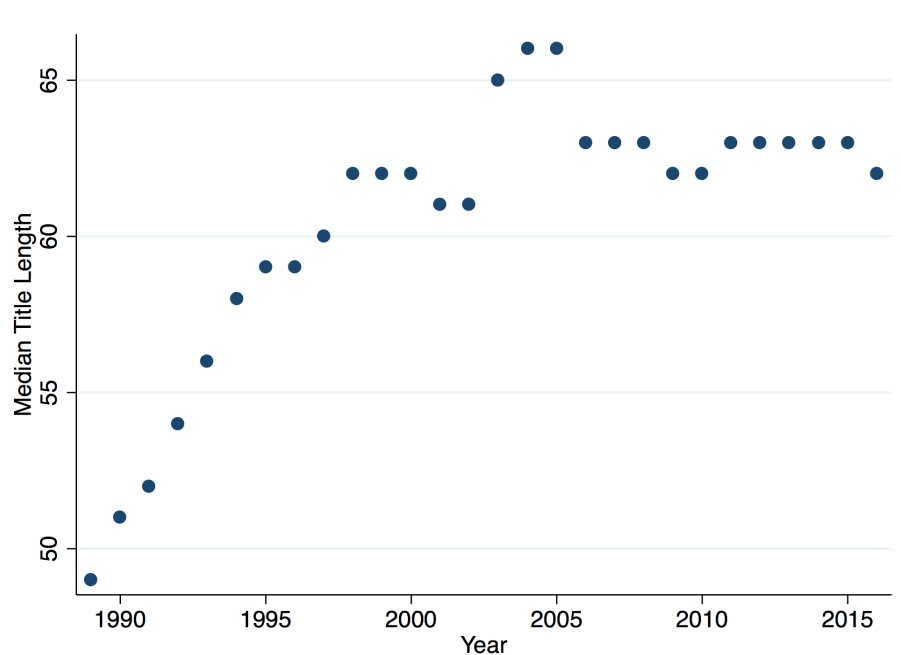
### A.3 Time Series of Title Lengths

**Figure 3:** Trends in the length of titles.

Figure A: 10-week average title length (characters) of *Seeking Alpha*-stock reports from 2006 to 2015.



Figure B: Median title length (characters) of earnings announcements released via *PR Newswire* and *Business Wire* from 1988 to 2016.



## A.4 First Stage using Press-Release Data

**Table 14:** Regressions showing the first-stage of title length on firm name length, using earnings-releases from 2003 to 2016. Firm name length and title length are measured in characters. The F-statistic for the coefficient on firm name length is 1338, exceeding the 10 recommended by Stock and Yogo (2005). Each regression includes an SIC4-by-year and date fixed effect. Standard errors are clustered by firm. Variables related to length are measured in characters.

	Title Length			Length of Name in Title		
	(1)	(2)	(3)	4)	(5)	(6)
Length of Firm's Legal Name	0.15*** (0.01)	0.13*** (0.00)	0.15*** (0.00)	0.60*** (0.01)	0.56*** (0.01)	0.55*** (0.01)
Log Market Cap			0.08*** (0.01)			-0.07*** (0.01)
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
SIC-4 FE	No	Yes	Yes	No	Yes	Yes
Adjusted R <sup>2</sup>	0.06	0.10	0.10	0.37	0.43	0.43
Within R <sup>2</sup>	0.02	0.02	0.02	0.36	0.32	0.32
Observations	421293	421293	421293	415433	415433	415433

## A.5 Robustness: Adding Firm and Event Controls

**Table 15:** Robustness check of Table 9 results by adding firm and event controls. Book leverage is the debt to total assets. Earnings surprise is the difference between reported EPS and the median analyst forecast of EPS from  $t - 45$  to  $t - 3$ , where  $t$  is the announcement date. I normalize the difference in actual and forecasted EPS by the stock price at  $t - 3$ . I restrict the sample to only those firms with all firm and event controls to keep sample size fixed. Requiring analyst-eps estimates substantially reduces the sample size. All regressions include SIC-4 by year fixed effects to account for industry trends and an announcement day fixed effect. Other controls include 10 days of lags of turnover, the intraday spread, absolute returns, and returns. All variables are standardized.

	% Turnover (t)				
	(1)	(2)	(3)	(4)	(5)
Title Length	-0.02 (0.01)	-0.04*** (0.01)	-0.04*** (0.01)	-0.04*** (0.01)	-0.04*** (0.01)
Book Leverage			0.01** (0.00)	0.01** (0.00)	0.01** (0.00)
Revenue			0.09*** (0.03)	0.09*** (0.03)	0.10*** (0.03)
Total Assets			-0.07*** (0.01)	-0.07*** (0.04)	-0.07*** (0.04)
EBITDA Margin				0.02*** (0.00)	0.02*** (0.00)
Earnings Surprise					0.01*** (0.00)
Log Market Cap (t-3)		-0.05*** (0.00)	-0.05*** (0.00)	-0.05*** (0.00)	-0.05*** (0.00)
Observations	165967	165967	165967	165967	165967

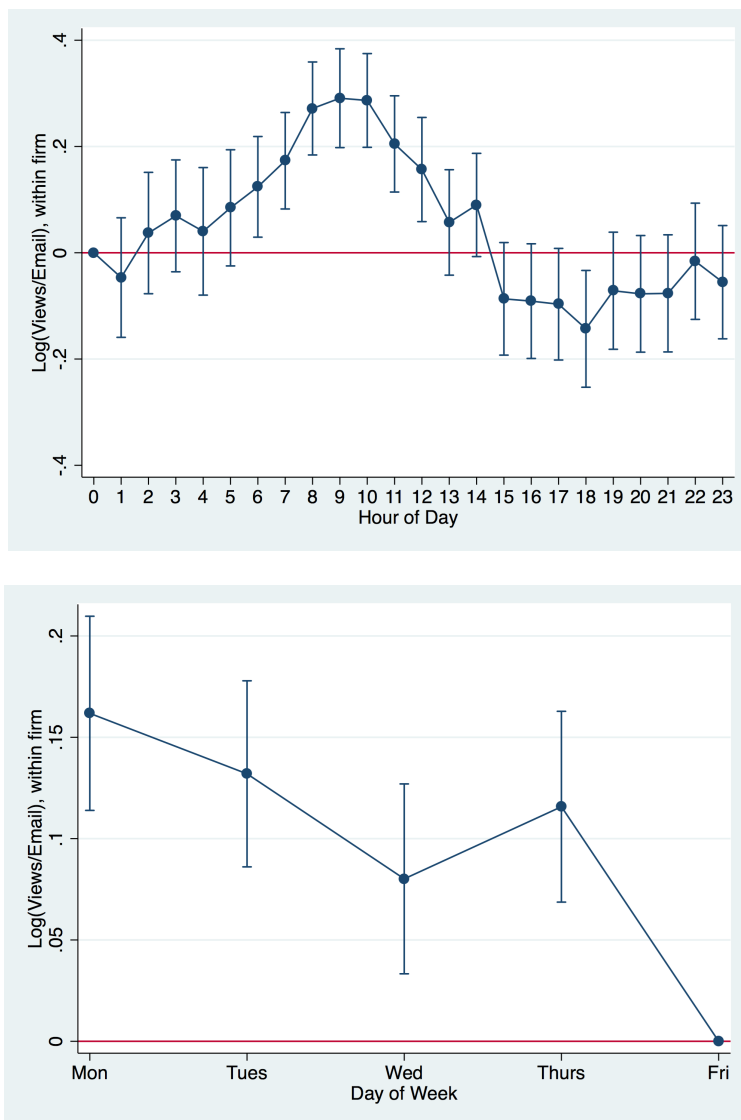
## A.6 Title Attributes Predict Stock-Report Attributes

**Table 16:** Regressions showing the relation between stock-report-title attributes and stock-report-body attributes, using *Seeking Alpha*-title-testing data. Title length and stock report length are measured in characters. Positive and negative sentiment are measured as the fraction of words that are positive or negative according to Bill McDonald’s lexicons. Firm names are excluded from the title length and sentiment scoring. The dependent variable in regression (2) is the Gunning fog index. A higher Gunning fog index predicts greater difficulty to read. The fog index depends on the number of words in sentences and the fraction of words with more than three syllables. Each regression includes a firm and date fixed effect. Standard errors are clustered by the topic firm of the stock report. All variables are standardized.

	Stock Report Length (1)	Stock Report’s Fog Index (2)	%Positive Words Stock Report (3)	%Negative Words Stock Report (4)
Title Length	0.11*** (0.02)	0.09*** (0.02)	-0.01 (0.02)	-0.05*** (0.02)
Negative Sentiment	-0.04*** (0.02)	0.02 (0.02)	-0.01 (0.01)	0.07*** (0.01)
Positive Sentiment	-0.02 (0.02)	0.01 (0.02)	0.03** (0.01)	-0.05*** (0.02 )
Firm FE	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes
Adjusted R <sup>2</sup>	0.17	0.14	0.56	0.45
Within R <sup>2</sup>	0.01	0.01	0.01	0.02
Observations	4128	4083	4128	4128

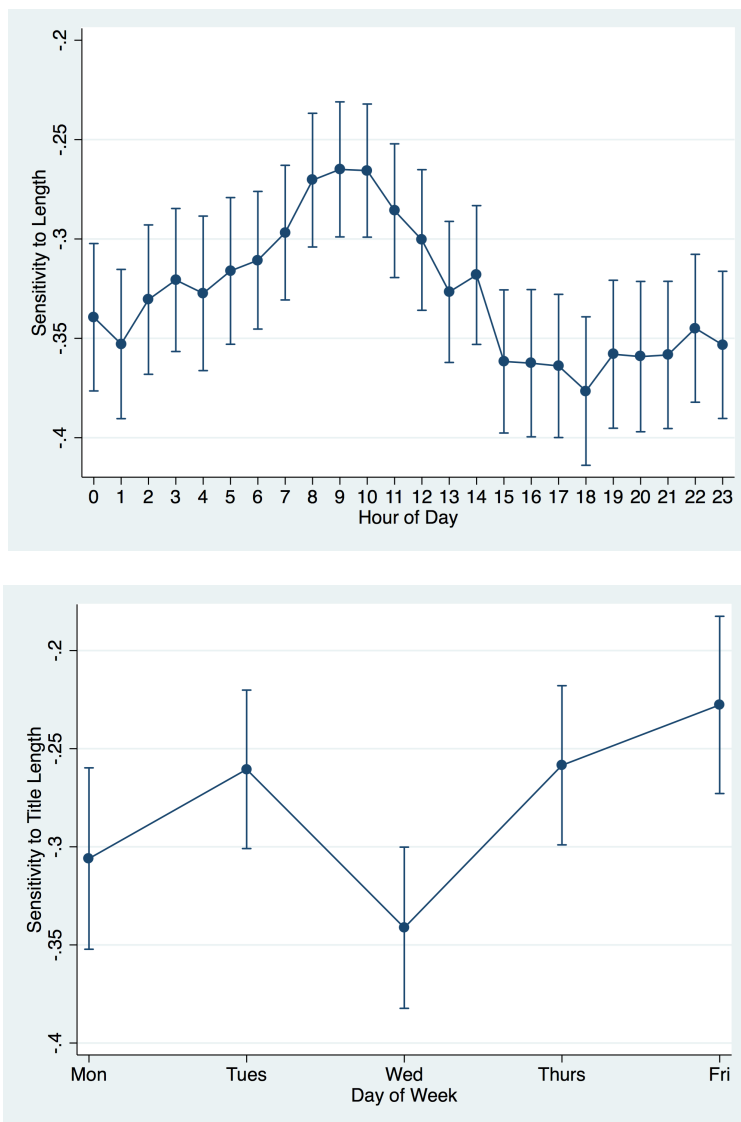
## A.7 Attention by Hour and Day of Week

**Table 17:** Figure showing the yield (log views/emails) for stock reports by hour-of-day released and day-of-week released, using *Seeking Alpha*-title-testing data. The by-hour figure (top) includes firm and date fixed effects. The weekday figure (bottom) includes firm and week fixed effects.



## A.8 Sensitivity to Title Length by Hour and Day of Week

**Table 18:** Figure showing the sensitivity of attention (log views/emails) to title length by hour-of-day released and day-of-week released, using *Seeking Alpha*-title-testing data. The by-hour figure (top) includes firm and date fixed effects. The weekday figure (bottom) includes firm and week fixed effects.





## A.9 Attention and the VIX

**Table 19:** Relation between the VIX and both the investor demand for news (email alert subscriptions) and attention to news (readership per email sent), using *Seeking Alpha* data. Market returns are the S&P 500 one-month market returns, since using the VIX derived from S&P 500 options. Each regression includes a firm fixed effect. Standard errors are clustered by firm and date.

	Total	
	Email Alerts (1)	Page Views (2)
VIX(t)	-0.01*** (0.00)	0.10*** (0.01)
Market Return (t-30,t)	-0.00 (0.00)	0.03** (0.01)
Firm Market Capitalization (t-1)	-0.01 (0.04)	-0.42*** (0.14)
Number of Emails Sent		-0.52*** (0.09)
Firm FE	Yes	Yes
Adjusted R <sup>2</sup>	0.99	0.60
Within R <sup>2</sup>	0.01	0.03
Number of Firms	1855	1855
Number of Titles	22563	22563

## A.10 Differences in Author’s and Editor’s Titles

This section examines whether the author’s original title tends to differ from the author’s alternative title and the editor’s title. I restrict the sample to only articles with three tested titles. Regression (1) examines for differences in title lengths. The author’s alternative title appears to be 4% longer than the original on average. The editor’s title tends to be 14% shorter than the original on average.

Regression (2) examines for differences in the frequency of being the title with the most views “winner.” I control for title length since title length explains differences in attention. Nevertheless, the author’s alternative title is less 7% less likely to receive the most views, and the editor’s title is 11% less likely to be the winner.

Regression (3) looks for differences in the sentiment of titles, controlling for the length of titles. I measure tone using the number of positive and negative words in the titles using the Bill McDonald lexicons. The author’s alternative title tends to have no difference in sentiment from the original. In contrast, the editor’s title tends to be more positive in tone.

**Table 20:** Regressions showing the differences between author and editor titles, use *Seeking Alpha*-title-testing data. The titles are the author’s original, author’s alternative, or editor’s. Negative sentiment is determined by counting the number of positive and negative words in a title, excluding the firm’s name, according to Bill McDonald’s lexicons of positive and negative words. Regressions include an article fixed effect, which holds fixed the event, firm, and date.

	Log Title Length (1)	1 if Title w/ Most Views (2)	Net Sentiment (3)
Author Alternative	0.06*** (0.01)	-0.04*** (0.01)	0.04*** (0.01)
Editor Title	-0.47*** (0.01)	-0.10*** (0.01)	0.08*** (0.02)
Log Title Length		-0.20*** (0.01)	0.05*** (0.01)
Article FE	Yes	Yes	Yes
Adjusted R <sup>2</sup>	0.46	-0.58	0.29
Within R <sup>2</sup>	0.11	0.05	0.00
Number of Articles	2744	2744	2744
Number of Titles	8232	8232	8232

## A.11 Robust to Screen Size Limitations

The negative relation between attention and title length might be due to longer titles not fitting on the small screens of mobile phones. To resolve this concern, I split the *Seeking Alpha* title-testing sample into groups of titles where the maximum length of the tested titles is 25 characters and can easily fit on a screen and separately, a group where the minimum length is 65 characters, so that all tested titles would not fit on a screen. The magnitudes of the effect of title length on attention is identical across groups, suggesting that screen size is not driving the results.

**Table 21:** Regressions for various sample splits using *Seeking Alpha*-title-testing data. Regression (1) only includes groups of tested titles for an article that have a maximum title length in characters less than or equal to 25. Regression (3) only includes groups of tested titles for an article that have a minimum title length greater than or equal to 65. The attention measure is the fraction of email alerts sent that investors click on, calculated by randomly assigned title. Regressions include an article fixed effect, which holds the event, firm, author, and date fixed. All variables are standardized, and standard errors are clustered by article.

	Views/Emails		
	(1)	(2)	(3)
Title Length	-0.13*** (0.05)	-0.12*** (0.01)	-0.12*** (0.04)
Short or Long Titles	Very Short ( $\leq 25$ )	Middle ( $> 33 \& < 65$ )	Long ( $\geq 65$ )
Article FE	Yes	Yes	Yes
Num. Articles	92	584	564
Observations	196	1294	1179

**Table 22:** This table examines the relation between the titles of SSRN papers of all academics with ID numbers between 1 and 700,000 on SSRN and the paper’s abstract views, downloads, and citations.

	Log Abstract Views (1)	Log Abstract Views (2)	Log Paper Downloads (3)	Log Paper Downloads (4)	Log Paper Citations (5)	Log Paper Citations (6)
Log Title Length	-0.22*** (0.01)	-0.11*** (0.00)	-0.13*** (0.01)	0.01* (0.00)	-0.04*** (0.00)	-0.02*** (0.00)
Log Abstract Views				1.22*** (0.00)		0.23*** (0.01)
Log Downloads						-0.01*** (0.00)
Log Abstract Length				0.01*** (0.00)		0.05*** (0.00)
Year-Week FE	No	Yes	Yes	Yes	Yes	Yes
Author FE	No	Yes	Yes	Yes	Yes	Yes
Adjusted R <sup>2</sup>	0.01	0.66	0.55	0.81	0.49	0.51
Within R <sup>2</sup>	.	0.00	0.00	0.58	0.00	0.04
Num. Academics	50011	50010	50010	50010	50010	50010
Observations	295390	295349	295349	295349	295349	295349
Clustering SE	Author	Author	Author	Author	Author	Author

## A.12 Title Length and Attention on SSRN

I examine non-experimentally how headline length matters for academics’ information-acquisition process. I cycled through academic ID numbers on SSRN from 0 to 700,000. Not all numbers have been assigned to individuals in this interval. The sample includes approximately 50,000 academics. For each academic, I retrieve all papers. Then, for each paper, I collect the abstract, number of abstract views, number of paper downloads, and number of citations.

Table 22 provides evidence consistent with the importance of title length. In regression (2), the dependent variable is abstract views, which is comparable to page views in *Seeking Alpha* data. The regression holds the author fixed and controls for the age of the papers. Doubling title length predicts 11% fewer abstract views.

In regression (3), the dependent variable is paper downloads. This measure is comparable to the read-to-end measure in *Seeking Alpha* data. Doubling title length predicts 13% fewer paper downloads. Regression (4) shows that conditional on viewing an abstract with a longer title, longer titles predict academics are more likely to download the paper. Also, longer abstracts predict that academics are more likely to download the paper.

In regression (5), the dependent variable switches to number of citations. About 60% of papers have 0 citations, so the dependent variable is  $\log(\text{citations} + 1)$ . Doubling title length predicts 4% fewer citations. Conditional on the number of abstract views and number of downloads, longer titles still predict fewer citations.

## A.13 Experimental Google Survey

One concern with the company-name-length instrument is that shorter company names may be more familiar or attention grabbing. To provide further support that optical-title length is off putting, I run two national surveys using Google’s survey tool. For each survey, I asked 500 adults in the United States between the ages of 35 and 55 the following question: “Which news on Apple is more interesting?” One set of 500 adults received Survey 1, whereas the other set received Survey 2. The news events are the same in both surveys; however, I varied the “Inc.” and “Incorporated,” which varies the relative optical length of the title non-informatively. In Survey 1, the first title was selected 41% of the time. In Survey 2, the first title was selected 52% of the time. The 95% confidence interval is  $\pm 3\%$ . The shortest title was overall preferred in both surveys even though the information content was held constant. The surveys were conducted at the same time, and Google randomly switched the order of headlines within each survey.

- Survey 1
  - Apple Incorporated Sees iPhone Sales Slump
  - Apple Inc. Launching Smart Home
- Survey 2
  - Apple Inc. Sees iPhone Sales Slump
  - Apple Incorporated Launching Smart Home

## A.14 Machine Learning Sentiment

Measuring the sentiment using lexicons is less suitable for titles, because titles are short and thus, 75% of titles have no matching words with the lexicons. Also, the lexicon approach equal-weights positive words like “good” and “great” and does not consider combinations of words. Therefore, I check my results using a machine-learning classifier approach.<sup>13</sup> This classifier learns “features” of titles that best predict whether the title is positive or negative. Features are words and combinations of words. The model learns the importance of features from annotated training data. *Seeking Alpha* classifies every stock report as bullish (recommending a purchase) or bearish (recommending a sale or short position). This classification of titles separates positive and negative titles well. The training data are all bullish and bearish titles from 2006 to 2015, a period that predates the title-testing program. The trained model then determines the sentiment of the titles in the title-testing sample. More specifically, the model provides a probability a title is negative. The results using this machine-learning measure of sentiment are very similar.

To illustrate variation in title sentiment within a stock report holding the context fixed, consider the following three titles. The model had a 5.3% confidence the first title is negative, a 31.6% probability that the second title is negative, and a 9.3% confidence that the third title is negative. The model learned that the word “trouble” is more indicative of a “negative” title than worried. The confidence scores are relatively low because the unconditional probability of a title being negative is low since most titles are bullish.

- Prospect Capital: About that Barclays note
- Prospect Capital: Is Barclays borrowing trouble?
- Prospect Capital: Barclays is worried, should you?

---

<sup>13</sup>I use a naïve Bayes classifier, which is a probabilistic classifier based on applying Bayes’ theorem with strong (naïve) independence assumptions between features.