The Role of Accounting Information in an Era of Fake News

Betty Liu[*]
Tippie College of Business
University of Iowa
yunchen-liu@uiowa.edu

Austin Moss
Tippie College of Business
University of Iowa
austin-moss@uiowa.edu

September 2021

Abstract

We investigate the role of accounting information in deterring the production of financial fake news and attenuating its market impact. Specifically, for firms that have issued more management forecasts, provided a 10-K with higher readability, or released major accounting information more recently, we find an 8-14% reduction in the probability of being targeted by a fake article on Seeking Alpha and an 8-18% reduction in the number of fake articles. We also find a reduction in abnormal trade volume and idiosyncratic return volatility following the publication of a fake article for these firms. Furthermore, analyses using a bunching identification strategy find that fake news production peaks prior to earnings announcements and drops upon announcement, providing evidence that impending accounting information releases induce fake news authors to avoid publishing post-disclosure when the accounting information environment is relatively stronger. Overall, our results are consistent with accounting information reducing the production of and the market reaction to fake news, providing evidence of an *ex ante* and *ex post* role of accounting information in safeguarding firms from financial misinformation.

JEL Classification: D82, D83, G14, M41

Keywords: Accounting Information, Fake News, Seeking Alpha, Financial Misinformation

---

[*] Corresponding author: Betty Liu. Postal address: W252 Pappajohn Business Building, Iowa City, Iowa 52242

We thank Robert Holthausen (editor) and an anonymous referee for their insightful suggestions. We are grateful for helpful discussions with Dan Collins, Cristi Gleason, Dennis Hamilton, Christian Hutzler, Paul Hribar, Andrew Imdieke, David Volant, and Clare Wang as well as helpful comments from workshop participants at the University of Iowa and the Midwest Accounting Research Conference. We appreciate Shimon Kogan for providing us with information about the fake news classification methodology as well as Seeking Alpha for their assistance in collecting article data. The authors have no conflicts of interest to disclose.

"As a public entity in a highly digital world, we have been and in the future may be the subject of so-called "fake news," a type of yellow journalism constructed to look legitimate while consisting of intentional misinformation and misrepresentations. […] While utilizing all available tools to defend the Company and its assets against fake news, there is limited regulatory control, making fake news an ongoing concern for any public company."

– Carvana Co. Prospectus, 5/23/2019

1. Introduction

Fake news—defined as false or misleading information with the intent to deceive—is a significant threat to efficient capital markets. In 2018, "Rota Fortunae" penned a Seeking Alpha article about Farmland Partners, Inc., alleging that "310% of 2017 earnings could be made-up" and that the firm bears "significant risk of insolvency." Despite the company refuting these claims as "false and materially misleading," the article caused investors to sell the firm's stock in a panic, resulting in an approximately 40% drop in share price (Farmland Partners, Inc., 2018).[1] Kogan, Moskowitz, and Neissner (2020) document broad empirical evidence on the magnitude and speed of investor reaction to fake news consistent with this anecdote, finding that market participants react as strongly to fake financial news as real financial news. In response to the risk posed by fake news, managers are discussing fake news during conference calls (Plymouth Industrial REIT, 2019), disseminating press releases in response to fake news (Regen BioPharma, 2019), and reporting fake news as a material risk in risk factor disclosures (Carvana, 2019). In light of this threat, we examine whether accounting information mitigates the impact of fake financial news in the capital markets.[2]

---

[1] Farmland Partners later filed a lawsuit against Rota Fortunae and his co-conspirators, who had taken a short position in the firm prior to article publication, for manipulating the stock price for profit. After three years of court proceedings, Farmland Partners eventually won the case, attesting to the difficulty of recouping the costs from a single fake article even if the firm presses charges.

[2] While accounting information can encompass information produced both internally (e.g., voluntary disclosures and mandatory SEC filings) and externally (e.g., analysts, media, and other information intermediaries), the term "accounting information" as used in this paper refers specifically to accounting information created by the firm.

Recent developments in the theoretical strategic communications literature suggest that false price signals are less effective when larger proportions of investors are informed (Schmidt, 2020). By providing a verifiable source of private information about operating profitability, investment opportunities, and, ultimately, fundamental value (i.e., the valuation role of accounting information), accounting information is thus particularly well-suited to counter fake news in the financial markets. For example, as one of the primary forms of voluntary disclosure, management forecasts not only offer private forward-looking information about future earnings (e.g., Hirst, Koonce, and Venkataraman, 2008) but also help investors by clarifying complexities in business transactions or reporting standards (e.g., Guay, Samuels, and Taylor, 2016) as well as signaling the quality of the manager's investment decisions (e.g., Goodman, Neamtiu, Shroff, and White, 2013). These voluntary disclosures are verified *ex post* via subsequent mandatory disclosures, complementing the *ex ante* credibility of the forecasts (i.e., the Confirmation Hypothesis) (e.g., Ball, Jayaraman, and Shivakumar, 2012; Li and Yang, 2016). In addition to their verification role as mandatory disclosures, 10-Ks contain audited financial statements and narrative disclosures that aid investors in understanding the business entity. Specifically, 10-Ks provide disaggregated line items with differential weights in forecasting future profitability (e.g., Fairfield, Sweeney, and Yohn, 1996), segment disclosures detailing profits attributed to major operating or geographical divisions (e.g., Berger and Hann, 2003), as well as management discussion and analyses that preempt or explain changes to business ecosystems (e.g., Ball, Hoberg, and Maksimovic, 2015). To the extent that these disclosures help investors learn the true asset value of the firm, accounting information reduces investor susceptibility to false price signals, disincentivizing the production of fake news.

Hence, in the context of accounting information providing a source of verifiable firm information to the market in an era of fake news, we investigate two research questions: Does accounting information deter the production of fake news? Does accounting information affect the market reaction to the fake news that is produced? In other words, we are interested in the ability of accounting information to *ex ante* disincentivize the production of fake news and to *ex post* mitigate the market reaction to fake news conditional on its publication.

We use Seeking Alpha, an independent investor research website, as our setting to study how accounting information impacts financial fake news. Seeking Alpha is conducive to studying our research questions for several reasons. First, though many authors' identities are hidden under pseudonyms, Seeking Alpha articles are read by 15.2 million visitors every month and elicit sizable market reactions (e.g., Hu, 2019; Kogan et al., 2020; Seeking Alpha, 2020). These factors provide an opportunity for self-interested authors to manipulate market opinions by writing fake news and largely avoid the reputation costs of doing so.[3] Second, Seeking Alpha publishes articles on the universe of firms, affording us a broad cross section of firms to study and increasing the external validity of our paper. Third, Seeking Alpha's webpage for each public firm allows for easy retrieval of the firm's filings with the Security and Exchange Commission (SEC), earnings call transcripts, and press releases. The saliency and accessibility of accounting information to Seeking Alpha authors and readers facilitates our ability to identify the effect of accounting information on fake news. Lastly, Kogan et al. (2020) develop a methodology to identify a large sample of Seeking Alpha articles written with the intent of

[3] Interestingly, Rota Fortunae (from the previously discussed Farmland Partners case) remained anonymous for almost two years of court proceedings and was found to be the subject of another lawsuit with similar allegations of promoting a "short-and-distort" scheme from a different firm, attesting to the difficulty of imposing reputation costs on authors who publish fake Seeking Alpha articles under a pseudonym.

deceiving readers and validate their classification approach on a sample of known fraudulent Seeking Alpha articles that led to SEC enforcement.[4]

We examine our research questions using three measures of accounting information: (1) management forecast frequency, (2) 10-K readability, and (3) information staleness (i.e., time since the last major accounting information release). We choose these proxies as they are particularly salient to fake news authors relative to other common measures of the accounting information environment.[5] Our goal in examining three distinct measures that capture voluntary, mandatory, and temporal aspects of accounting information is to provide convergent validity and general evidence on the role of accounting information in combating fake news.

Our first research question investigates the effect of accounting information on the production of fake news. We use a logit model to test the probability that a fake article is written about the firm. We hypothesize and find results consistent with accounting information reducing the production of fake news. A one-standard-deviation increase in management forecast frequency or 10-K readability or a one-standard-deviation decrease in information staleness is associated with an 8-14% decrease in the likelihood that fake news is written. As complementary evidence, we use Poisson pseudo-maximum likelihood estimation to test the effect of accounting information on the number of fake articles written about a firm in a quarter and find qualitatively similar inferences to the logit model. To help mitigate omitted variable concerns, we find that our results are largely unchanged in additional analyses both within subsamples of firms with

[4] In 2014 and 2017, the SEC levied enforcement actions against various companies and individuals for fraudulently commissioning authors on Seeking Alpha to write several hundred optimistic, self-promoting articles under the guise of independent analyses (SEC, 2014; 2017).

[5] Other common measures of accounting information often require explicit estimation using statistical analyses (e.g., earnings persistence, abnormal accruals, conservatism, etc.). We view these measures as being less salient to fake news authors and less likely to affect the publication of fake news articles.

similar information environments and with firm-year fixed effects. Hence, we provide robust evidence on accounting information deterring the production of fake news.

Our second research question examines the effect of accounting information on the market reaction to fake news. We hypothesize and find results mostly consistent with accounting information attenuating both abnormal trade volume and idiosyncratic return volatility following fake article publication. A one-standard-deviation increase in management forecast frequency or 10-K readability or a one-standard-deviation decrease in information staleness is associated with a 3-6% decrease in abnormal trade volume and a 16-23% decrease in idiosyncratic return volatility to fake news. Overall, we find evidence consistent with accounting information mitigating the market reaction to fake news.

As our final set of analyses, we use a bunching identification strategy to address concerns that alternative sources of financial information other than accounting information drive the reduction of fake news publication in our results. Conceptually similar to discontinuities in earnings distributions (e.g., Burgstahler and Dichev, 1997), bunching is an empirical methodology developed in the economics literature to ascribe behavioral distortions to a discontinuous change in incentives at certain thresholds (Kleven, 2016). If the distribution of observed outcomes exhibits a "bunching" of outcomes on the preferred side of threshold and a missing mass of avoided outcomes on the other, the anomalous pattern is attributed to the discontinuity in incentives at the threshold. Because prior literature documents significant informativeness linked to earnings announcements (e.g., Beaver, 1968; Atiase and Bamber, 1994), we use these accounting events as a large discontinuous increase in informed investors to study the publication timing preferences of fake news authors in a narrow event window.

Overall, we find bunching distributions consistent with accounting information deterring fake news: (1) fake news authors prefer to publish fake articles outside of the earnings announcement window and (2) when fake news authors do publish fake articles near an earnings announcement, they strongly prefer to publish them prior to the release of accounting information relative to afterwards.

Our paper makes several contributions to the literature. First, we provide evidence of an additional role of accounting information in an era of fake news—to disincentivize fake news authors from spreading intentionally misleading information. We document results consistent with accounting information *ex ante* deterring the production of fake news and *ex post* reducing the market reaction to fake news. Second, we contribute to the limited empirical literature on the effects of public misinformation on the stock market. Historically, empirical studies of known stock market manipulations, such as “pump-and-dumps”, have been scarce, due to the small number of occurrences enforced by regulators and the difficulty in identifying unenforced market manipulations (e.g., De Franco, Lu, and Vasvari, 2007; Leuz, Meyer, Muhn, Soltes, and Hackethal, 2017; Weiner, Weber, and Hsu, 2017).[6] More recently, researchers have extended this literature by investigating the effects of potentially exploitative behavior on fast-growing investor websites, such as Seeking Alpha (e.g., Hu, 2019; Kogan et al., 2020). As investor websites are mainstays of the contemporary financial environment, evidence of accounting information moderating the negative effects of the fake news disseminated on these websites is especially meaningful, speaking directly to the concern raised by Blankespoor, deHaan, and

[6] Pump-and-dumps are market manipulation schemes in which one party takes a position in a stock, disseminates false positive information about that firm to boost its stock price temporarily, and profits by liquidating the position while the stock price is inflated.

Marinovic (2020) regarding the lack of traditional oversight on these online platforms and their potential for misinformation.

Lastly, we contribute to the broader scientific literature investigating the proliferation and social impact of fake news. Lazer et al. (2018) note the relative scarcity of research on the effects of fake news and call for interdisciplinary research on how to safeguard the public from its effects. The authors highlight two broad interventions that may diminish the dissemination and influence of fake news but for which further research is needed: (1) preventing exposure to fake news and (2) empowering individuals to evaluate the fake news they encounter. We provide some evidence on both interventions in the context of financial markets. Specifically, we find that accounting information (1) decreases the exposure of market participants to financial fake news by disincentivizing its production and (2) empowers investors in evaluating financial fake news as evidenced by reduced market reactions.[7] Further research is needed to evaluate the generalizability of our results to fake news settings outside of financial markets.

2. Theoretical Background and Hypothesis Development

We reference the theoretical literature on strategic communication to motivate our investigation of the effect of accounting information on fake news. Specifically, we refer to the stream of literature on why market participants react to *ex ante* unverifiable information, as is the case with Seeking Alpha articles published under pseudonyms (Dyer and Kim, 2021).

[7] Other than a brief discussion in Section 5, we note that we largely do not speak to the effect of accounting information on total fake news production in equilibrium and leave in-depth investigation on whether authors produce fewer fake articles in aggregate or merely shift their choice of firms to target from those with robust information environments to those with weaker information environments to future research.

Historically, analytical models of strategic communication in financial markets have examined an information sender's reputation as a source of credibility when information is costly to verify (e.g., Benabou and Laroque, 1992; Stocken, 2000; Van Bommel, 2003). However, theories relying on reputation in repeated games to induce credible information sharing cannot rationalize why markets react to *ex ante* unverifiable information (i.e., rumors) from anonymous senders.[8]

Schmidt (2020) proposes a cheap talk model in which *ex ante* unverifiable information from an unknown source can be credible in equilibrium. Rather than relying on reputation as the mechanism to incentivize truthful communication, Schmidt rationalizes market reactions to rumors by allowing investor investment horizons to be occasionally misaligned with the horizon of the investor's private information. For example, an investor has private information about long-term price movements but has a short-term investment horizon. In anticipation, market participants rationalize that, as this investor's investment horizon shortens, his incentive to send truthful information increases in order to accelerate price discovery and maximize the capitalization of his private information into stock price.[9] This truth-telling incentive allows for a credible information sharing equilibrium to develop absent a reputation concern for the information sender. More importantly, for the context of our paper, Schmidt's model also sheds light on how the truth-telling preference of a sender changes as the fraction of informed investors changes. In the model, informed investors disregard false messages and trade according to their own information, resulting in an equilibrium in which a sender's preference for sending untruthful messages decreases as the proportion of informed investors increases.

[8] There is substantial documentation of market reaction to unverified merger or takeover rumors (Greenberg, 2007; Jia, Shu, and Zhao, 2020).

[9] Schmidt extends his original model to a generalized case in which market participants do not need to know the horizon of a specific sender as long as there exists investors with a continuum of investment horizons.

Although Schmidt's model is agnostic on how investors learn the true asset value of the stock, longstanding theoretical and empirical literatures endorse accounting information as one of the primary mechanisms for informing investors about fundamental value (i.e., the valuation role of accounting information). Noisy rational expectations models support the usefulness of accounting disclosures in increasing the precision of investor beliefs about future cash flows or earnings by decreasing information asymmetry or investor uncertainty (e.g., Diamond, 1985; Dye, 1985; Verrecchia, 2001). In addition, ample empirical literature documents significant informativeness associated with accounting information events, such as earnings announcements (e.g., Beaver, 1968; Landsman and Maydew, 2002; Collins, Li, and Xie, 2009), 10-K releases (e.g., Stice, 1991; Griffin, 2003; You and Zhang, 2009), and management forecasts (e.g., Jennings, 1987; Yang, 2012; Twedt, 2016).

Hence, the role of accounting information in facilitating an increase in the proportion of informed investors in conjunction with the inferences drawn from Schmidt (2020) leads to the following hypotheses stated in the alternative form:

H_1: Accounting information deters the production of fake news.

H_2: Accounting information decreases the market reaction to fake news.

3. Sample Selection and Variable Measurement

3.1 Data Sources and Sample Selection

We obtain data from Seeking Alpha for all articles written from 2006 through 2018. We gather the article's text, author, publication date, and the primary stock tickers associated with the firms discussed in the article.[10] We eliminate articles without a primary stock ticker and articles written by Seeking Alpha employees. These restrictions eliminate news updates and conference call transcripts as well as articles about the economy or other general topics not linked to a specific company. To ensure that the linguistic software used to classify fake news has sufficient content, we require articles to have greater than 100 words. We drop articles that are not classified as fake or non-fake using the methodology discussed in Section 3.2. In addition, we require non-missing financial data from Compustat and CRSP and obtain analyst data from IBES. Our final sample includes 125,475 articles across 37,864 firm-quarters. Table 1 provides details of our sample selection process.

3.2 Identifying Fake News Articles

We follow the fake news classification method detailed in Kogan et al. (2020) to identify articles as "fake" or "non-fake" using the Linguistic Inquiry Word Count (LIWC2015) model from Pennebaker et al. (2015). This algorithm, built upon linguistic and psychometric research, detects the intent to deceive in written text and calculates an authenticity score using a proprietary formula. The linguistics literature documents that individuals who are being dishonest use less self-reference words, shorter sentences, less insight words, less specific

[10] If an article lists multiple primary stock tickers, the article appears as multiple observations in our sample, with one observation linked to each stock ticker.

information about time and space, and more discrepancy verbs (Pennebaker, 2011).[11] Kogan et al. (2020) obtain 171 paid-for fake articles and 334 non-fake articles all written by the same set of authors on Seeking Alpha. The authors use this cleanly-identified sample to map the LIWC-based authenticity score into the conditional probability of being fake. The authenticity cutoffs that Kogan et al. (2020) use for classification are conservative in nature, achieving a type II error (i.e, incorrectly classifying a fake article as non-fake) of less than 10% and a type I error (i.e, incorrectly classifying a non-fake article as fake) of less than 1%. The Central Intelligence Agency and Federal Bureau of Investigations use similar linguistic methods to measure the authenticity of written text and speech, providing application-based validity for this methodology. In our sample, the proportion of fake articles to the total number of fake and non-fake articles is 2.5%, quantitatively similar to the 2.8% identified in Kogan et al. (2020).

3.3 Measures of Accounting Information

We examine our research questions using 3 measures of accounting information: (1) management forecast frequency, (2) 10-K readability, and (3) information staleness. We discuss the specific measurement of each below.

3.3.1 Management Forecast Frequency

Management forecasts serve as important voluntary disclosures that reduce information asymmetry in the market (e.g., Verrecchia, 2001; Healy and Palepu, 2001; Beyer, Cohen, Lys, and Walther, 2010). Beyer et al. (2010) show that management forecasts provide 55% of the firm's accounting-based information in explaining stock returns. In addition, prior literature

[11] *Realize*, *understand*, and *think* are examples of insight words. Discrepancy verbs, such as *could*, assert that an event might have occurred but possibly did not.

documents many specific avenues in which the provision of management forecasts informs investors, including their assistance in projecting bottom line earnings as well as other key line items from the income statement (Lansford, Lev, and Tucker, 2007), explaining complex financial statements (Guay et al., 2016), and signaling the manager's corporate investment efficiency (Goodman et al., 2014). To the extent that management forecasts provide detailed forward-looking information about anticipated earnings, sales projections, and potential growth, fake articles that portray exaggerated future firm conditions are less likely to sway investors. We measure *Management Forecast Frequency* as the natural logarithm of one plus the number of management forecasts a firm has issued within the past year of the Seeking Alpha article publication date.

3.3.2 10-K Readability

Our second proxy, the linguistic readability of the firm's 10-K, captures a salient element of mandatory accounting information quality. Though the 10-K contains mandatory disclosures crafted to follow standards set forth by the Financial Accounting Statement Board and vetted by legal and audit teams, there is nevertheless considerable variation in the writing style and length of 10-Ks (e.g., Li, 2008; Bonsall, Leone, Miller, and Rennekamp, 2017). According to the disclosure processing cost framework presented in Blankespoor et al. (2020), the lower information acquisition and integration costs associated with clearer textual disclosures allow investors to incorporate more information from the disclosure in their valuation and investment decisions. In support of this framework, empirical evidence finds that more readable disclosures increase trading on information (Bloomfield, 2002; Miller, 2010) as well as individual investors' understanding of financial disclosures (Lawrence, 2013). In the context of our paper, if investors can more easily glean narrative information from the firm's annual reports about its operating environment such as product line synergies, peer competition, and risk factors, fake news that inaccurately portray details about firm operations is less persuasive. We measure *10-K Readability* as the Bog Index from Bonsall et al. (2017) multiplied by -1 for ease of interpretation.

3.3.3 Information Staleness

Our last proxy captures a temporal aspect of accounting information, in which the information in accounting disclosures becomes less value relevant as time passes. This construct exists more commonly in the analyst literature: existing analyst reports become "stale" as new information and analyst reports are created (e.g., O'Brien, 1988). We expect that accounting information from the firm also suffers from this decay in informativeness. To the extent that accounting information loses value relevance as time passes, creating more uncertainty about the true asset value of the firm, fake news authors may take advantage of the increased uncertainty to spread misinformation. We measure *Information Staleness* as the natural logarithm of one plus the number of days between the Seeking Alpha article publication date and the last major accounting information release from the firm (i.e., 10-K or 10-Q filing, management forecast, or earnings announcement). We note that one particular strength of this proxy is that it measures a time-varying construct within-firm, providing within-firm-quarter evidence on the effect of accounting information on fake news.

4. Empirical Analyses and Results

4.1 The Content of Fake News

In Appendix B, we provide two Seeking Alpha articles that demonstrate how accounting information can be used by investors to uncover fake news. The first article, shown in Exhibit A, is a fake Seeking Alpha article that was later prosecuted by the SEC in 2014, and the second article, shown in Exhibit B, disputes claims made in the first by referring to the firm's financial statements. In the first article, the author provides analyses of the firm's future profitability and share price based on its two major pharmaceutical products, resembling other non-fake articles with fundamental analyses. Interestingly, the author downplays the management forecast of 10-15% long term market share as "conservative" (highlighted under "Best Case"), redirecting instead to a potential future skyrocketing market share of 30%. Although the management forecast itself was unsuccessful in deterring fake news in this specific anecdote, the fact that the author chose to downplay the validity of the management forecast shows that he is not only aware of these accounting information disclosures himself but also aware of investors using management forecasts in judging the veracity of the claims in Seeking Alpha articles.

In the second article, a different author discredits the bullish sentiment caused by the fake news surrounding Galena Biopharma by providing internet links to and screen captures of the company's 10-Qs, 10-Ks, and press releases (highlighted throughout the article). The stock price fell by 20% after the publication of this article, partially offsetting the inflated stock price from fake news (SCAC, 2014). In addition to correcting the market, this article demonstrates that the author uses accounting information to verify the news surrounding Galena Biopharma and that he believes he can convince general market participants by referring to the firm's accounting information in his article. These examples provide useful anecdotes into what information a fake article may contain as well as how market participants can counteract the misinformation in fake news using accounting information.

Next, we use textual analysis to provide more data on the content of these fake news articles. Specifically, we use Latent Dirichlet Allocation (LDA), a linguistic machine learning method used to identify latent topics in a corpus of text, to analyze topical areas of content for all articles in our sample (see IA1 and IA2 in the Internet Appendix for implementation details of LDA). We find that articles are written about topics such as accounting information and forecasts, specific industry news, legal matters, macroeconomic conditions, and more. Table 2 contains the identified article topics. For the purpose of displaying descriptive statistics, we compute the probabilities of an article containing each topic and assign articles to the topic with the highest probability. We tabulate the number of articles in our sample assigned to each topic as well as the percentage of articles that are classified as fake within the articles in that topic. Our evidence demonstrates that a non-trivial number of Seeking Alpha articles are primarily about accounting information as 31% of articles have accounting information (Topic 5 and Topic 25) as their most prominent topic. Importantly, we note that in untabulated analysis we find that 86% of articles are assigned a positive probability of including an accounting topic, demonstrating that even articles focused on non-accounting topics often contain accounting content. We view this evidence as support for our usage of the broad sample of Seeking Alpha articles, as accounting information has the potential to deter specific pieces of fake news even within non-accounting articles.

4.2 The Role of Accounting Information in Deterring Fake News Production

4.2.1 Accounting Information and the Probability of Fake News

We examine the role of accounting information in deterring the production of fake news by estimating the conditional probability that an article is fake. In accordance with our first hypothesis, we expect an increase in *Management Forecast Frequency* or *10-K Readability* or a decrease in *Information Staleness* to decrease the probability that a fake article is written. To analyze the determinants of fake versus non-fake articles, we estimate the following model at the article level:

$$Pr(\textit{Fake Article}_j) = \beta_1 \textit{ Accounting Information}_i + \sum \beta_i \textit{ Controls}_i + \sum \textit{Fixed Effects} + \varepsilon. \quad (1)$$

Fake Article is an indicator variable equal to one when the article is classified as fake and zero when non-fake. *Accounting Information* is one of three variables as defined in Section 3.3: *Management Forecast Frequency, 10-K Readability,* or *Information Staleness.* In all our regression specifications, we include a vector of control variables that influence the firm's external information environment or firm's operating environment. Appendix A contains definitions for variables used in our analyses. Accounting characteristics are measured as of the fiscal quarter end in which the earnings announcement for the quarter occurs on or before the article publication date. We also include industry and year fixed effects, unless noted otherwise, to control for unobserved heterogeneity along these two dimensions that could be correlated with both our accounting information variables and our dependent variables. Table 3 contains descriptive statistics for our primary regression variables. We note that all dependent variables are tabulated at their corresponding observation level in regressions and that all independent variables are tabulated at the article level.

Using binned scatterplots shown in Panels A, B, and C in Figure 1, we first present the conditional, nonparametric relationship between the probability of a fake article (i.e., *Fake Article*) and the number of management forecasts in the past year (i.e., *Management Forecast Frequency (Count)*), the Bog Index multiplied by -1 (i.e., *10-K Readability*), or the number of days since the most recent major accounting information release (i.e., *Information Staleness (Days)*), respectively.[12] The notes to Figure 1 provide details on how we construct these binned scatterplots.

All three panels in Figure 1 demonstrate evidence that accounting information is associated with a reduction in fake news. The dashed lines in Figure 1 show the best linear fit estimated on the underlying sample of articles using OLS regression.[13] The relationship between *Fake Article* and *Management Forecast Frequency (Count)* in Panel A is negative, suggesting that a higher quantity of management forecasts is associated with a lower prevalence of fake news. Panel B documents a negative relationship between *Fake Article* and *10-K Readability,* implying that firms providing more readable 10-Ks are less likely to be the target of fake news. Lastly, the positive slope in Panel C indicates that the probability of fake news is lower when firms have recently released major accounting information. The conditional, nonparametric representation of the relationship between *Fake Article* and *Accounting Information* in each panel also shows that the positive or negative relationships are consistent throughout the range of

[12] *Management Forecast Frequency (Count)* and *Information Staleness (Days)* are the untransformed versions of *Management Forecast Frequency* and *Information Staleness,* our main independent variables described above.

[13] All slope estimates are significantly different than zero at conventional significance levels (i.e., p-value < .05).

values for *Accounting Information*, providing reassurance that our results are not driven by a particular subset of the data distribution.

Table 4 provides the results of estimating Equation (1) using a logit regression model. We present coefficients as marginal effect estimates multiplied by 100 to interpret them as percentage changes. We discuss economic magnitudes relative to the unconditional probability that an article is fake. In Column 1, we examine whether the number of management forecasts affects the likelihood of a fake article. We find a negative and significant coefficient for *Management Forecast Frequency,* indicating that a one-standard-deviation increase in *Management Forecast Frequency* prior to the article publication date reduces the probability that an article is fake by 8%. Column 2 examines how the readability of the 10-K affects the likelihood of a fake article. The negative and significant coefficient for *10-K Readability* suggests that a one-standard-deviation increase in *10-K Readability* decreases the probability of a fake article by 10%. In Column 3, we test whether the likelihood of a fake article changes as the most recently issued accounting disclosure ages and loses informativeness. We find a positive and significant coefficient for *Information Staleness*, suggesting that the presence of more recently issued accounting information decreases the probability that a fake article is written. Specifically, the probability of a fake article decreases by 14% for a one-standard-deviation decrease in the number of days since the last major accounting information release.

In Column 4, we include all three accounting information variables to examine whether each of our variables of interest has an incremental effect on the production of fake news. The coefficient estimates on all three accounting information variables remain significant in the expected directions without notable decreases in magnitude. Thus, we are reassured that our main independent variables capture distinct measures of accounting information and offer convergent validity for our inferences on the role of accounting information in deterring fake news.[14]

We briefly note the effects of our control variables on the publication of fake articles. We find that variables capturing firm monitors or information intermediaries largely do not offer incremental deterrence to fake news publication over our accounting information variables. Both *Analyst Coverage* and *Institutional Ownership* are insignificant or inconsistently significant. *Media Coverage* is positive and significant with an economic magnitude of 13%, comparable to our effect estimates for our accounting information variables of 8-14%. This result indicates that more media articles in the press are associated with a higher probability that a fake article is published. We conjecture that the positive coefficient is potentially caused by attention effects in that fake news authors want to write about firms with increased investor attention to maximize readership and influence of their fake articles. We also note that poor past performance (i.e., *Adj. ROA* and $Returns_{m-12,m-1}$) increases the likelihood of a fake article publication, but poor short-term news (i.e. $Returns_{t-10,t-1}$) decreases the likelihood. We do not try to hypothesize why these opposing signs occur and leave the investigation of the link between performance and fake news publication to future research.

One concern with our interpretation of the results in Table 4 is that we primarily rely on cross sectional variation to identify the association between accounting information and the production of fake news, resulting in potential omitted variable bias. We address this concern

[14] Our results in Column 4 are robust to using either a 180-day or 90-day window for measuring management forecast frequency as well as dropping industry-years with less than 50 observations. See IA3 in the Internet Appendix for tabulated results.

with a series of additional tests, as follows: (1) Partitioning our sample by whether the article contains accounting content to provide validation that our fake article sample discusses content for which accounting information can deter, (2) Adding firm-year fixed effects to our estimation of *Information Staleness* in Column 3 of Table 4, (3) Partitioning our sample by the sign of the earnings surprise of the last earnings announcement to account for firm performance as a confound, and (4) Performing a host of additional subsample analyses within firms with more similar general information environments. These tests limit the variation of both observed and unobserved variables, providing robustness for our inferences.

We first address the concern that our results are contaminated by fake non-accounting articles that cannot be influenced by accounting information. Columns 1 and 2 of Table 5 present our main specification partitioning by whether the article contains accounting content.[15] Specifically, we partition our sample based on whether articles have a positive probability of containing either of the accounting topics (Topic 5 and Topic 25) as computed by LDA (untabulated). We tabulate the results of running our main test using only articles that contain accounting content in Column 1. All three accounting information coefficients remain statistically significant in the expected direction. As a complement, Column 2 displays the results of using articles without accounting content. We find statistically insignificant coefficients for all three accounting information variables, providing falsification evidence against correlated omitted variables expected to influence the publication of fake non-accounting articles. This pair

[15] We note that our usage of the terms “main test” or “main specification” refers to the logit test including all 3 accounting measures of interest as shown in Table 4 Column 4 unless otherwise specified.

of analyses provides solace that our results are driven by articles for which accounting information is particularly relevant.

In Column 3, we test an alternative specification for *Information Staleness* using firm-year fixed effects. This more rigid specification examines the association between *Information Staleness* and *Fake Article* while controlling for factors that are constant within a firm-year, alleviating some concerns that unobserved factors drive this particular result. The addition of firm-year fixed effects necessitates two changes to this analysis compared to the specification in Table 4 Column 3. One, we limit our estimation sample to articles with at least one fake and one non-fake article in a firm-year to ensure sufficient variation in the dependent variable within each fixed effect. Two, because it is infeasible to estimate a logit model with a high number of fixed effect parameters, we estimate this specification using OLS. We continue to find a positive and significant coefficient for *Information Staleness*, providing evidence against the concern that our results in Table 4 Column 3 are driven by unobservable factors that do not exhibit meaningful variation within a firm-year. We note that, since neither *Management Forecast Frequency* nor *10-K Readability* exhibits sufficient within-firm-year variation, we follow best practices outlined by deHaan (2021) and only use this specification for *Information Staleness*.[16]

Next, we address the concern that firm performance determines both accounting disclosure policy and attention from fake news authors. Prior literature documents the relation

[16] We estimate the variation in *Management Forecast Frequency*, *10-K Readability*, and *Information Staleness* after removing the effect of firm-year fixed effects. We find that 22%, 13%, and 95% of variation remain for *Management Forecast Frequency*, *10-K Readability*, and *Information Staleness*, respectively. Hence, we do not estimate *Management Forecast Frequency* and *10-K Readability* with firm-year fixed effects to avoid erroneous inferences based on limited variation within fixed effects (deHaan, 2021). We note that the coefficient for *Information Staleness* is largely unchanged if we include *Management Forecast Frequency* and *10-K Readability* as controls and do not tabulate these results.

between bad performance and decreased voluntary disclosure or 10-K readability (e.g., Li, 2008; Chen, Matsumoto, and Rajgopal, 2011). To the extent that fake news authors are drawn towards writing about firms with worse performance, our results may be driven by this omitted variable. In addition to using return on assets, short-run past returns, and long-run past returns as controls, we perform our main test partitioning by the sign of the earnings surprise of the last earnings announcement and tabulate them in Columns 4 and 5 of Table 5. We continue to find statistically significant results in both partitions for all 3 accounting information variables, alleviating the concern of firm performance as an omitted variable.

Lastly, we conduct a host of subsample analyses to mitigate the concern that our independent variables of interest capture the quality of the general information environment rather than accounting information in particular. In other words, firms with systematically better general information environments may provide more management forecasts or more readable 10-Ks than firms that do not, creating uncertainty about whether we can attribute our inferences specifically to accounting information. To alleviate this concern, we run our main test within subsamples of firms likely to have similar general information environments, limiting the amount of unobserved variation that exists in our models. We use management forecasts, analyst coverage, institutional ownership, and size as our partitioning variables, as prior literature documents these characteristics as particularly important in determining a firm's general information environment (e.g., Beyer et al., 2010). Table 6 presents the results of these subsample tests. For parsimony, we only report the coefficients for our accounting information variables. We find statistically significant and economically meaningful coefficients within each subsample, with the exception of insignificant coefficients on *Management Forecast Frequency* in the low analyst coverage and small size groups. In summary, our series of tests with alternative specifications yield largely robust results in support of our main inferences, reducing the likelihood of an unobserved variable driving our results.

Overall, our results in this section are consistent with our first hypothesis that accounting information deters the production of financial fake news. The estimated effect sizes are economically meaningful and reasonable, with all three measures of accounting information incremental to each other. Further, the results continue to be largely significant and in the expected direction in a series of additional tests.

4.2.2 Accounting Information and the Quantity of Fake News

To provide additional evidence on the role of accounting information in deterring the production of fake news, we examine our first hypothesis with an alternative dependent variable *# of Fake Articles,* the count of the number of fake Seeking Alpha articles published in a firm-quarter. Since *# of Fake Articles* only takes nonnegative integer values, we use Poisson pseudo-maximum likelihood estimation at the firm-quarter level to analyze how accounting information affects the amount of fake news published about a firm within a quarter:[17]

$$\# \textit{ of Fake Articles}_i = \beta_1 \textit{ Accounting Information}_i + \sum \beta_i \textit{ Controls}_i + \sum \textit{Fixed Effects} + \varepsilon. \quad (2)$$

We modify our independent variables to adjust for the change from estimating Equation (1) on an article level to estimating Equation (2) on a firm-quarter level. *Management Forecast*

[17] We use Poisson pseudo-maximum likelihood estimation as it offers consistent estimators for over-dispersed (i.e., the variance is greater than the mean), highly skewed data distributions (e.g., Gourieroux, Monfort, and Trognon, 1984; Santos Silva and Tenreyro, 2011).

Frequency, *10-K Readability*, and our control variables are measured as of the first article of the quarter. *Information Staleness* is measured as the average *Information Staleness* of all fake and non-fake articles about the firm in the quarter. We continue to use the same control variables and fixed effects as described in Section 4.2.1 with the addition of *Seeking Alpha Articles*, the number of Seeking Alpha articles written about the firm in the quarter.

Table 7 presents the results from estimating Equation (2) using Poisson pseudo-maximum likelihood estimation. Consistent with our first hypothesis as well as results in Tables 4-6, we find negative and significant coefficients for *Management Forecast Frequency* and *10-K Readability* and a positive and significant coefficient for *Information Staleness*. The coefficient estimates in Columns 1-3 continue to be economically meaningful. A one-standard-deviation increase in the number of management forecasts reduces the amount of fake news by 8%. A one-standard-deviation increase in *10-K Readability* is associated with 18% fewer fake articles. Lastly, the *# of Fake Articles* decreases by 11% for a one-standard-deviation decrease in the number of days since the last major accounting information release. In Column 4, all three coefficient estimates remain significant and in the expected direction when estimated in the same regression. Overall, we continue to find evidence in Table 7 consistent with our first hypothesis that accounting information deters the production of fake news.

4.3 The Role of Accounting Information in Reducing the Market Reaction to Fake News

Next, we examine the effect of accounting information on the market reaction to fake news. In accordance with our second hypothesis, we expect that accounting information decreases the ability of fake news to influence investor judgments, resulting in a lower market reaction to these fake articles. To test our hypothesis, we estimate the following model using OLS at the article level:

$$Market\ Reaction_{t,t+2} = \beta_1\ Accounting\ Information_i + \sum \beta_i\ Controls_i + \sum Fixed\ Effects + \varepsilon. \tag{3}$$

Our dependent variable *Market Reaction* is one of two variables used to measure the market response to fake Seeking Alpha articles: *Abnormal Volume,* a measure based on trading activity, and *Idiosyncratic Return Volatility,* a measure based on price movement. *Abnormal Volume* is the sum of scaled trading volume on the publication date of the Seeking Alpha article and the following two trading days, where scaled trading volume is calculated as the daily trading volume scaled by the average trading volume between the 20 and 140 trading days prior. *Idiosyncratic Return Volatility* is the sum of squared abnormal returns on the article publication date and the following two trading days multiplied by 100. We measure a firm's abnormal return as the daily return minus the return on a 5x5x5 size-, B/M-, and momentum-matched portfolio (Daniel, Grinblatt, Titman, and Wermers, 1997). We avoid using a signed measure of price reaction because assigning an expected direction of price movement to Seeking Alpha articles is challenging and noisy (Kogan et al., 2020). We use both trade- and price-based reaction variables to obtain a more holistic view of the market reaction to fake news as well as address concerns that excess trading can occur without impacting prices (e.g., Fama, 1970) or that substantial price movement can occur without any trade (e.g., Milgrom and Stokey, 1982).

We control for market reactions to other events in two ways. First, we exclude articles from these analyses if they are published within two days of an earnings announcement, management forecast, 10-K, 10-Q, or 8-K both prior to and after these events because we cannot disentangle the market reaction to these events from the reaction to the Seeking Alpha articles.[18] Second, we include single-day measurements of our two market reaction variables for the three trading days before article publication to control for other unobserved events that cause market reactions. We continue to use the control variables and fixed effects described in Section 4.2.1.

Table 8 presents the results for our second hypothesis examining whether accounting information affects the market reaction to fake news. Panel A estimates Equation (3) with *Abnormal Volume* as the dependent variable. In Column 1, we examine the association between *Management Forecast Frequency* and *Abnormal Volume* and find a negative but statistically insignificant coefficient. In Column 2, we examine whether a more readable 10-K results in a lower market reaction to fake news. The negative and significant coefficient estimate for *10-K Readability* implies that a one-standard-deviation increase in *10-K Readability* is associated with a 6% decrease in *Abnormal Volume* to the publication of fake news. In Column 3, we examine whether the market reacts less to fake news if the most recently issued accounting information is newer and more informative. Consistent with this hypothesis, we find a positive and significant coefficient for *Information Staleness* implying that a one-standard-deviation decrease in *Information Staleness* is associated with a 3% decrease in *Abnormal Volume.* Column 4 estimates Equation (3) including all three accounting information variables. Our inferences from Column 4 are consistent with those from Columns 1-3.

Table 8 Panel B reports the results from estimating Equation (3) using *Idiosyncratic Return Volatility* as the dependent variable. We find that a one-standard-deviation increase in the

[18] Due to the importance and informativeness of these events, market reactions measured during this window are likely to be overwhelmingly in response to the event and not the Seeking Alpha article.

number of management forecasts is associated with a 16% decrease in return volatility following a fake news article in Column 1. In Column 2, we find that a one-standard-deviation increase in *10-K Readability* is associated with a 23% lower *Idiosyncratic Return Volatility*. In Column 3, we estimate that a one-standard-deviation decrease in *Information Staleness* is associated with 17% decrease in *Idiosyncratic Return Volatility*. Our inferences from Column 4 that includes all our accounting information variables are consistent with those from Columns 1-3. Overall, we find evidence in support of our second hypothesis that accounting information attenuates the market reaction to fake news.

5. Bunching Analyses Examining Fake News Publication Timing Preferences

To further alleviate concerns that we are attributing the impact of the general information environment or another unobserved variable to accounting information, we use a bunching identification strategy to provide causal evidence for accounting information changing publication behavior in fake news authors. In general, bunching is an empirical methodology developed in the economics literature to attribute distortions in behavioral outcomes to a known discontinuous change in incentives due to certain thresholds (Kleven, 2016).[19] Intuitively, the existence of certain thresholds with discontinuities in incentives cause outcomes on one side of the threshold to dominate those on the other side in preferences, inducing behavioral distortions in actions or reporting so that the observed outcome is on the preferred side of the threshold. Thus, the density distribution of outcomes exhibits excess mass (i.e., "bunching") in the region

[19] This methodology has gained popularity in the public economics and finance literatures to study a diverse range of topics, such as taxpayer responses to tax schedule cutoffs and lenders' supply of credit in response to government loan guarantees (e.g., Saez, 2010; Chetty, Friedman, Olsen, and Pistaferri, 2011; Kleven and Waseem, 2013; Bachas, Kim, and Yannelis, 2021).

of preferred outcomes on one side of the threshold and a missing mass of avoided outcomes on the other.[20]

Applying an identification strategy following Sallee (2011) that combines the bunching and difference-in-differences methodologies (henceforth, difference-in-bunching), we use earnings announcements to study the publication timing preferences of fake news authors.[21] Earnings announcements, through the disclosure of meaningful accounting information, induce significant discontinuous increases in the proportion of informed investors (e.g., Beaver, 1968; Atiase and Bamber, 1994). In accordance with our theoretical development in Section 2, the relatively lower proportion of informed investors prior to the earnings announcement allows fake news to be more successful in influencing investor opinions than afterwards, when the proportion of informed investors increases sharply due to the announcement (Schmidt, 2020). We exploit the fact that Seeking Alpha authors are freelancers with both the discretion and incentive to produce fake news when there is a relatively lower proportion of informed investors. Hence, in our bunching analyses, we expect an excess mass in the density distribution of fake articles before earnings announcements and a missing mass afterwards if accounting information deters fake news publication.

[20] The bunching methodology is conceptually related to the distribution discontinuity methods used to study the effect of salient thresholds on earnings management behavior (e.g., Burgstahler and Dichev, 1997) but is different in certain underlying assumptions. In the context of earnings management, earnings distributions exhibit excess mass just above salient performance thresholds and missing mass just below.

[21] Sallee (2011) incorporates a temporal element to bunching to show that consumers accelerated their purchase of the Toyota Prius prior to decreases to hybrid vehicle tax subsidies. Sallee uses a difference-in-bunching approach by utilizing consumer purchases of non-hybrid Toyota sedans as a control behavior (see Figures 2-7 in his paper).

5.1 Research Design and Results

We construct density distributions of both fake and non-fake article publication for bunching analyses. We first match our sample of articles to the earnings announcements of each firm for articles published within 45 days of the announcement date, retaining only matched articles in our bunching sample. We use the number of hours between the time of Seeking Alpha article publication and the earnings announcement to create the *Days to EA* variable. Specifically, we create 90 blocks of 24-hour periods (henceforth, "days") centered on the time of the earnings announcement to the nearest minute. For example, an article published 26 hours prior to an earnings announcement is classified as being two days prior to an earnings announcement (i.e., *Days to EA* = -2).[22] To create distributions of publication behavior for both fake and non-fake articles, we count the number of articles published in event time relative to the earnings announcement date. *Fake Articles*$_t$ is equal to the number of fake articles published on *Days to EA*$_t$ summed across all earnings announcements and scaled by the total number of fake articles in our sample. *Non-Fake Articles*$_t$ follows the same procedure for non-fake articles. Figure 2 Panel A depicts the resulting density distributions for both fake and non-fake articles.

We use the distribution of non-fake articles as the counterfactual in applying the difference-in-bunching methodology. Analogous to the control group in difference-in-differences research designs, the counterfactual behavior approximates what would be observed in our behavior of interest absent the change in incentives at the threshold. To be an appropriate counterfactual for our study, non-fake article publication must exhibit parallel trends in behavior to fake article publication without the direct influence of accounting information. We expect this

[22] There are no articles published at the exact same time as an earnings announcement in our data.

to be the case, as the primary incentive to publish both fake and non-fake articles is compensation linked to readership (e.g., Seeking Alpha payment per view, internet clout, etc.) (Dyer and Kim, 2021). While counterfactuals cannot be formally tested for suitability, Figure 2 Panel A provides compelling visual evidence that the distributions of fake and non-fake articles closely mirror each other outside of the immediate days surrounding the earnings announcement.

Next, we present graphical evidence consistent with earnings announcements disincentivizing fake news production. For ease of interpretation, we compute *Abnormal Mass*$_t$, the difference between *Fake Articles*$_t$ and *Non-Fake Articles*$_t$. In Figure 2 Panel B, we plot *Abnormal Mass* for each event day. We note that *Abnormal Mass* spikes in t-2 and t-1 and drops sharply t+1 through t+8. This pattern indicates that relative fake news publication bunches prior to earnings announcements when the accounting information environment is weaker and exhibits a large missing mass post-announcement when the accounting information environment is stronger. Thus, our graphical evidence is consistent with fake news authors strategically choosing to avoid publishing in stronger accounting information environments.

We formalize our inferences by conducting statistical analyses testing for the existence of fake article publication bunching around earnings announcements. Following prior literature, we first identify a narrow affected region by visually inspecting the distribution of fake articles in Figure 2 Panel A. Visual inspection suggests that abnormal publication behavior starts two days prior to earnings announcements and lasts until around eight days post announcement. Thus, we set the affected region equal to t-2, t+8.[23,24] We then construct four different variables of interest

---

[23] Our results are robust to alternative windows: (1) t-2, t+7, (2) t-3, t+8, and (3) t-3, t+7.

[24] We interchangeably use the terms "affected region," "earnings announcement window," and "announcement window" in our paper.

as follows. (1) *Pre EA Abnormal Mass*$_{t-2,t-1}$ is the sum of *Abnormal Mass*$_t$ for days t-2 and t-1. (2) *Post EA Abnormal Mass*$_{t+1,t+8}$ is the sum of *Abnormal Mass*$_t$ for days between t+1 and t+8. *(3) Differential Abnormal Mass*$_{t-2,t+8}$ is the difference between *Pre EA Abnormal Mass*$_{t-2,t-1}$ and *Post EA Abnormal Mass*$_{t+1,t+8}$. (4) *Total Abnormal Mass*$_{t-2,t+8}$ is the sum of *Abnormal Mass*$_t$ for days between t-2 and t+8. We follow the bootstrap procedure by Chetty, Friedman, Olsen and Pistaferri (2011) to compute standard errors for statistical inferences. Specifically, we create a bootstrap distribution by randomly sampling *Abnormal Mass*$_t$ from the observed distribution in Figure 2 Panel B for each of the 90 days of the distribution. We then calculate our four variable estimates using the bootstrap distribution. We repeat this procedure 1,000 times and define the standard error as the standard deviation of the estimates from this procedure.[25]

Table 9 presents the numerical density estimates from our difference-in-bunching methodology for our full sample of articles. All four estimates are statistically significant at least at the 5% level. We first examine fake news publication behavior relative to the earnings announcement using Columns 1-3. The estimates for *Pre EA Abnormal Mass*$_{t-2,t-1}$ and *Post EA Abnormal Mass*$_{t+1,t+8}$ indicate abnormal fake article densities of 5% and -11%, respectively. These results show fake news articles bunching prior to earnings announcement and exhibiting a missing mass post announcement, consistent with fake news authors avoiding publication during periods of relatively strong accounting information. In addition, the estimate for *Differential Abnormal Mass*$_{t-2,t+8}$ denotes a 16% gap in the abnormal density of fake articles pre- and post-announcement, indicating the economically significant difference in preferences of publishing in

[25] As an additional safeguard against an invalid parallel trends assumption, our standard errors represent differences in fake and non-fake article publication behavior outside the earnings announcement window. To the extent that these differences exhibit excess variance (i.e., a potential sign that the parallel trends assumption is violated), the standard error will be large and result in statistically insignificant estimates.

these two periods. Thus, our bunching analyses find that fake news authors avoid publishing after earnings announcements when the accounting information environment is relatively stronger, electing to publish fake articles prior to earnings announcements instead.

Next, we examine the overall net impact of earnings announcements on fake news publication during the announcement window. Our estimate for *Total Abnormal Mass*$_{t-2,t+8}$ indicates that the density of fake articles is 6% lower than the density of non-fake articles during the same time period. Although the raw counts of fake and non-fake articles both increase around earnings announcements likely due to attention effects (as seen in Figure 2 Panel A), we find that there is a relative deficit in fake news inside the earnings announcement window compared to outside the window, providing some preliminary evidence that accounting information decreases net publication of fake news in equilibrium. We have two conjectures for why we observe this behavior. One, while some fake news authors are able to shift their fake news publication from post-announcement to pre-announcement, others opt to not publish at all, creating short-term slippage in the total amount of fake news. Two, the spike in information search and acquisition around earnings announcements increases the proportion of informed investors, disincentivizing fake news authors from writing during the announcement window in general (e.g., Drake, Roulstone, and Thornock, 2012).

To further validate our bunching results, we conduct the same analyses after partitioning our sample by articles classified by LDA as containing accounting content.[26] For the subsample analyses of articles with accounting content, the density distributions of fake and non-fake articles are graphed in Figure 3 with numerical estimates presented in Table 9. Our inferences

[26] We use the same methodology as our accounting content partitions in Section 4.2.1.

from both the graphical and statistical evidence remain the same as the full sample. In addition, within the subsample of articles without accounting content (Figure 4 and Table 9), there is no meaningful publication pattern nor discontinuity around earnings announcements, providing falsification evidence against correlated omitted variables expected to influence the publication of fake non-accounting articles. This pair of analyses provides validation that our results are driven by articles for which accounting information is particularly relevant. Similar to our prior empirical logit tests partitioning by accounting content, finding our hypothesized pattern in the expected subsample (i.e., articles with accounting content) and not in the subsample where we do not expect a pattern (i.e., articles without accounting content) bolsters our inference that the abnormal bunching behavior is driven by articles for which accounting information is particularly relevant.

In aggregate, our difference-in-bunching analyses show that accounting information deters the publication of fake news by providing evidence that fake news authors strategically avoid publishing fake news articles in robust accounting information environments. We show that the publication of fake news bunches prior to earnings announcements and drops after earnings announcements not only in our full sample of articles but also in a restricted subsample of articles containing accounting content. Furthermore, within the subsample of articles without accounting content, we do not find any evidence of publication behavior changes, providing reassurance that the behavioral patterns we document are not artifacts of the publication of fake non-accounting articles. Overall, the results from our difference-in-bunching analyses allow us to make clear inferences regarding the role of accounting information in deterring fake news that are distinct from the general information environment for two reasons. One, the empirical pattern of fake news publication behavior that we document around earnings announcements (i.e., excess mass prior to earnings announcements and missing mass afterwards) is consistent with accounting information as a salient disincentive to fake news authors when publishing fake articles. Two, we are unable to offer alternative explanations that attribute this pattern to an unobserved variable such as the general information environment, which does not change discontinuously at the time of earnings announcements for reasons independent of the accounting information event itself. [27] Hence, we interpret the results from our bunching analyses as compelling evidence that accounting information in particular deters the publication of fake news.

6. Conclusion

Our paper investigates two research questions regarding the impact of accounting information on financial fake news: Does accounting information deter the production of fake news? Does accounting information affect the market reaction to the fake news that is produced? We examine these questions using three measures of accounting information: (1) management forecast frequency, (2) 10-K readability, and (3) information staleness.

Using a sample of Seeking Alpha articles, we find results consistent with accounting information deterring the production of fake news and attenuating the market reaction to fake news. Specifically, firms with more management forecasts in the past year, more readable 10-Ks, and more recent accounting information releases have a lower probability of being the target of a

[27] While the general information environment does change around earnings announcements, these changes are generally attributable to the earnings announcement itself. Therefore, we view changes to the general information environment during this time as directly accounting-induced.

fake article and have a fewer number of fake articles written about them. In addition, abnormal volume and idiosyncratic return volatility are lower following the publication of a fake article about these firms. As further evidence, bunching analyses demonstrate that fake news authors strategically avoid publishing fake news articles after earnings announcements, when the accounting information environment is relatively more robust. While the absence of random variation limits us from definitively establishing that our results are not driven by unobservable confounds, our compilation of primary analyses in combination with validation and falsification tests, an alternative specification with firm-year fixed effects, and extensive subsample tests provide compelling evidence for our hypotheses. Collectively, our results are consistent with accounting information reducing the production of and the market reaction to fake news, providing evidence of an *ex ante* and *ex post* role accounting information plays in safeguarding firms from financial misinformation.

References

Atiase, R. K., and Bamber, L. S. 1994. Trading volume reactions to annual accounting earnings announcements: The incremental role of predisclosure information asymmetry. *Journal of Accounting and Economics*, *17(*3), 309-329. https://doi.org/10.1016/0165-4101(94)90031-0

Bachas, N., Kim, O. S., and Yannelis, C. 2021. Loan guarantees and credit supply. *Journal of Financial Economics*, *139*(3), 872–894. https://doi.org/10.1016/j.jfineco.2020.08.008

Backus, M., Conlon, C., and Sinkinson, M. 2021. Common Ownership in America: 1980–2017. *American Economic Journal: Microeconomics*, *13*(3), 273-308. https://doi.org/10.1257/mic.20190389

Ball, C., Hoberg, G., and Maksimovic, V. 2015. Disclosure, Business Change, and Earnings Quality. *Working Paper*. https://ssrn.com/abstract=2260371

Ball, R., Jayaraman, S., and Shivakumar, L. 2012. Audited financial reporting and voluntary disclosure as complements: A test of the Confirmation Hypothesis. *Journal of Accounting and Economics*, *53*(1–2), 136–166. https://doi.org/10.1016/j.jacceco.2011.11.005

Beaver, W. H. 1968. The Information Content of Annual Earnings Announcements. *Journal of Accounting Research*, *6*(1968), 67–92. https://doi.org/10.2307/2490070

Benabou, R., and Laroque, G. 1992. Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility. *The Quarterly Journal of Economics*, *107*(3), 921–958. https://doi.org/10.2307/2118369

Berger, P. G., and Hann, R. 2003. The impact of SFAS no. 131 on information and monitoring. *Journal of Accounting Research*, *41*(2), 163–223. https://doi.org/10.1111/1475-679X.00100

Beyer, A., Cohen, D. A., Lys, T. Z., and Walther, B. R. 2010. The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics*, *50*(2–3), 296–343. https://doi.org/10.1016/j.jacceco.2010.10.003

Blankespoor, E., deHaan, E., and Marinovic, I. 2020. Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics, 70*(2-3), 101344. https://doi.org/10.1016/j.jacceco.2020.101344

Bloomfield, R. J. 2002. The "Incomplete revelation hypothesis" and financial reporting. *Accounting Horizons*, *16*(3), 233–243. https://doi.org/10.2308/acch.2002.16.3.233

Bonsall, S. B., Leone, A. J., Miller, B. P., and Rennekamp, K. 2017. A plain English measure of financial reporting readability. *Journal of Accounting and Economics*, *63*(2–3), 329–357. https://doi.org/10.1016/j.jacceco.2017.03.002

Burgstahler, D. and Dichev, I. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics*, *24*(1), 99-126. https://doi.org/10.1016/S0165-4101(97)00017-7

Carvana. 2019. *Prospectus Supplement.* Available from: sec.gov/Archives/edgar/data/1690820/000119312519155493/d746066d424b2.htm.

Chen, S., Matsumoto, D., and Rajgopal, S. 2011. Is silence golden? An empirical analysis of firms that stop giving quarterly earnings guidance. *Journal of Accounting and Economics*, *51*(1–2), 134–150. https://doi.org/10.1016/j.jacceco.2010.10.006

Chetty, R., Friedman, J. N., Olsen, T., Pistaferri, L. 2011. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: evidence from Danish tax records. *Quarterly Journal of Economics, 126*(2):749–804. http://dx.doi.org/10.1093/qje/qjr013

Collins, D. W., Li, O. Z., and Xie, H. 2009. What drives the increased informativeness of earnings announcements over time? *Review of Accounting Studies*, *14*(1), 1–30. https://doi.org/10.1007/s11142-007-9055-y

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of Finance*, *52*(3), 1035– 1058. https://doi.org/10.1111/j.1540-6261.1997.tb02724.x

De Franco, G., Lu, H., and Vasvari, F. P. 2007. Wealth transfer effects of analysts' misleading behavior. *Journal of Accounting Research*, *45*(1), 71–110. https://doi.org/10.1111/j.1475-679X.2007.00228.x

deHaan, E. (2021). Practical Guidance on Using and Interpreting Fixed Effects Models. *Working Paper*. https://doi.org/10.2139/ssrn.3699777

Diamond, D. W. 1985. Optimal Release of Information by Firms. *The Journal of Finance*, *40*(4), 1071-1094.

Drake, M.S., Roulstone, D.T. and Thornock, J.R. 2012. Investor Information Demand: Evidence from Google Searches Around Earnings Announcements. *Journal of Accounting Research*, *50*, 1001-1040. https://doi.org/10.1111/j.1475-679X.2012.00443.x

Dye, R. A. 1985. Disclosure of Nonproprietary Information. *Journal of Accounting Research*, *23*(1), 123-145. https://doi.org/10.2307/2490910

Dyer, T., and E. Kim 2021. Anonymous Equity Research. *Journal of Accounting Research, 59*(2), 575-611. https://doi.org/10.1111/1475-679X.12359

Fairfield, P. M., Sweeney, R. J., and Yohn, T. L. 1996. Accounting Classification and the Predictive Content of Earnings. *The Accounting Review*, *71*(3), 337-355.

Fama, E. F. 1970. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25(2)*, 383–417. https://doi.org/10.2307/2325486

Farmland Partners, Inc. 2018. "Farmland Partners Inc. Responds to Unfounded Allegations". Available from: http://ir.farmlandpartners.com/file/Index?KeyFile=394184917.

Greenberg, Herb. 2007. As Buyout Speculation Thrives, Checks, Balances Are in Order. *Wall Street Journal.* Available from: https://www.wsj.com/articles/SB117832449817392917.

Goodman, T. H., Neamtiu, M., Shroff, N., and White, H. D. 2014. Management forecast quality and capital investment decisions. *The Accounting Review*, *89*(1), 331–365. https://doi.org/10.2308/accr-50575

Gourieroux, C., Monfort, A., and Trognon, A. 1984. Pseudo Maximum Likelihood Methods: Applications to Poisson Models. *Econometrica*, *52*(3), 701-720. https://doi.org/10.2307/1913472

Griffin, P.A. 2003. Got Information? Investor Response to Form 10-K and Form 10-Q EDGAR Filings. *Review of Accounting Studies, 8*, 433–460. https://doi.org/10.1023/A:1027351630866

Guay, W., Samuels, D., and Taylor, D. 2016. Guiding through the Fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics*, *62*(2–3), 234–269. https://doi.org/10.1016/j.jacceco.2016.09.001

Healy, P. M., and Palepu, K. G. 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, *31*(1–3), 405–440. https://doi.org/10.1016/S0165-4101(01)00018-0

Hirst, D. E., Koonce, L., and Venkataraman, S. 2008. Management earnings forecasts: A review and framework. *Accounting Horizons*, *22*(3), 315–338. https://doi.org/10.2308/acch.2008.22.3.315

Hu, D. 2019. Detecting Opportunistic Behavior in Public Short Campaigns. *Working Paper.*

Jennings, R. 1987. Unsystematic security price movements, management earnings forecasts, and revisions in consensus analyst earnings forecasts. *Journal of Accounting Research*, *25*, 90–110. https://doi.org/10.2307/2491260

Jia, W., Redigolo, G., Shu, S., & Zhao, J. 2020. Can social media distort price discovery? Evidence from merger rumors. *Journal of Accounting and Economics*, *70*(1), 101334.

Kleven, H. J. 2016. Bunching. *Annual Review of Economics, 8*, 435-64. https://doi.org/10.1146/annurev-economics-080315-015234

Kleven, H., and Waseem, M. 2013. Using notches to uncover optimization frictions and structural elasticities: theory and evidence from Pakistan. Quarterly Journal of Economics, 128, 669–723.

Kogan, S., T. Moskowitz, and M. Niessner. 2020. Fake news in Financial Markets. *Working Paper.* https://doi.org/10.2139/ssrn.3237763

Landsman, W. R., and Maydew, E. L. 2002. Has the Information Content of Quarterly Earnings Announcements Declined in the Past Three Decades? *Journal of Accounting Research*, *40*(3), 797–808. https://doi.org/10.1111/1475-679X.00071

Lansford, B., Lev, B., and Tucker, J. W. 2013. Causes and Consequences of Disaggregating Earnings Guidance. *Journal of Business Finance & Accounting*, *40*(1-2), 26-54. https://doi.org/10.1111/jbfa.12002

Lawrence, A. 2013. Individual investors and financial disclosure. *Journal of Accounting and Economics*, *56*(1), 130–147. https://doi.org/10.1016/j.jacceco.2013.05.001

Lazer, D., M. Baum, J. Benkler, A. Berinsky, K. Greenhill, F. Menczer, M. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, S. Sloman, C. Sunstein, E. Thorson, D. Watts, and J. Zittrain. 2018. The science of fake news. *Science 359* (6380): 1094–1096. https://doi.org/10.1126/science.aao2998

Leuz, C., Meyer, S., Muhn, M., Soltes, E., and Hackethal, A. 2017. Who Falls Prey to the Wolf of Wall Street? Investor Participation in Market Manipulation. In *National Bureau of Economic Research Working Paper Series*. https://doi.org/10.3386/w24083

Li, F. 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics, 45*(2–3), 221–247. https://doi.org/10.1016/j.jacceco.2008.02.003

Li, X. and Yang, H. 2016. Mandatory Financial Reporting and Voluntary Disclosure: The Effect of Mandatory IFRS Adoption on Management Forecasts. *The Accounting Review, 91*(3), 933-953. https://doi.org/10.2308/accr-51296

Miller, B. P. 2010. The effects of reporting complexity on small and large investor trading. *The Accounting Review 85(6)*, 2107–2143. https://doi.org/10.2308/accr.00000001

Milgrom, P. and N. Stokey. 1982. Information, trade and common knowledge. *Journal of Economic Theory 26(1),* 17–27. https://doi.org/10.1016/0022-0531(82)90046-1

O'Brien, P. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics 10,* 53–83. https://doi.org/10.1016/0165-4101(88)90023-7

Plymouth Industrial REIT. 2019. *Q3 2019 Earnings Conference Call.* Available from: sec.gov/Archives/edgar/data/1515816/000117152019000338/ex99-1.htm

Pennebaker, J., R. Booth, R. Boyd, and M. Francis. 2015. *Linguistic inquiry and word count: LIWC 2015 operator's manual*.

Pennebaker, J. W. 2011. *The Secret Life of Pronouns: What Our Words Say About Us.* Chapter 6. New York: Bloomsbury Press.

Regen BioPharma. 2019. *Regen BioPharma, Inc. Responds to Fake News Release.* Available from: sec.gov/Archives/edgar/data/1589150/000160706219000195/ex99_1.htm

Saez, E. 2010. Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy, 2*(3),180–212. https://doi.org/10.1257/pol.2.3.180

Sallee, J. M. 2011. The Surprising Incidence of Tax Credits for the Toyota Prius. *American Economic Journal: Economic Policy, 3*(2), 189-219. https://doi.org/10.1257/pol.3.2.189

Santos Silva, J. M. C., and Tenreyro, S. 2011. Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. *Economics Letters, 112*(2), 220–222. https://doi.org/10.1016/j.econlet.2011.05.008

Schmidt, D. 2020. Stock market rumors and credibility. *Review of Financial Studies 33(8),* 3804-3853. https://doi.org/10.1093/rfs/hhz120

Securities and Exchange Commission. 2014. SEC vs. Lidingo Holdings LLC et al. Available from: https://www.sec.gov/litigation/complaints/2017/comp23802-lidingo.pdf.

Securities and Exchange Commission. 2017. *SEC: Payments for Bullish Articles on Stocks Must Be Disclosed to Investors.* Available from: https://www.sec.gov/news/press-release/2017-79.

Securities Class Action Clearinghouse. 2014. Cho et al. vs. Galena Biopharma, Inc. et al. Available from: https://securities.stanford.edu/filings-documents/1051/GBI00_01/20141031_r01c_14CV00367.pdf

Seeking Alpha. 2020. *Who Reads Seeking Alpha.* Available from: https://seekingalpha.com/page/who_reads_sa.

Stepner, Michael. 2014. Binned Scatterplots: introducing -binscatter- and exploring its applications. *2014 Stata Conference 4*, Stata Users Group.

Stice, E. 1991. The Market Reaction to 10-K and 10-Q Filings and to Subsequent The Wall Street Journal Earnings Announcements. *The Accounting Review,* 66(1), 42-55. www.jstor.org/stable/247704

Stocken, P. C. 2000. Credibility of Voluntary Disclosure. *The RAND Journal of Economics*, *31*(2), 359–374. https://doi.org/10.2307/2601045

Twedt, B. 2016. Spreading the Word: Price discovery and newswire dissemination of management earnings guidance. *The Accounting Review*, *91*(1), 317–346. https://doi.org/10.2308/accr-51129

Van Bommel, J. 2003. Rumors. *The Journal of Finance*, *58*(4), 1499–1520. https://doi.org/10.1111/1540-6261.00575

Verrecchia, R. E. 2001. Essays on disclosure. *Journal of Accounting and Economics*, *32*(1–3), 97–180. https://doi.org/10.1016/S0165-4101(01)00025-8

Weiner, B. P. M., Weber, R. D., and Hsu, K. 2017. The growing menace of ‘short and distort’ campaigns. *Westlaw Journal Securities Litigation & Regulation*.

Yang, H. 2012. Capital market consequences of managers’ voluntary disclosure styles. *Journal of Accounting and Economics*, *53*, 167–184. https://doi.org/10.1016/j.jacceco.2011.08.003

You, H., and Zhang, X. 2009. Financial reporting complexity and investor underreaction to 10-K information. *Review of Accounting Studies*, *14*, 559–586. https://doi.org/10.1007/s11142-008-9083-2

Appendix A: Variable Definitions

Variable	*Definition*
Dependent Variables:	
Fake Article$_t$	An indicator variable equal to one when the Seeking Alpha article is classified as fake and zero when non-fake using the methodology in Kogan et al. (2020). Source: Seeking Alpha
# of Fake Articles$_q$	The number of fake articles written about a firm in the quarter. Source: Seeking Alpha
Abnormal Volume$_{t,t+2}$	The sum of the scaled trading volume on the day of publication and the following two trading days. Scaled trading volume is defined as trading volume scaled by the average trading volume between 20 and 140 trading days prior. Source: CRSP
Idiosyncratic Return Volatility$_{t,t+2}$ (%)	The sum of the squared abnormal returns on the day of publication and the following two trading days multiplied by 100. Abnormal return is defined as a firm's daily return minus the daily return on a 5x5x5 size-, B/M-, and momentum-matched portfolio. Source: CRSP
Accounting Information Variables:	
Management Forecast Frequency$_{t-365,t}$	The natural logarithm of one plus the number of management forecasts in the past year. Source: IBES
10-K Readability$_{y-1}$	The Bog Index from Bonsall et al. (2017) multiplied by -1. This variable is available for 10-Ks filed on or prior to March 31st, 2018. Source: Sam Bonsall Data Library (https://sites.psu.edu/sambonsall/data/)
Information Staleness$_t$	The natural logarithm of one plus the minimum number of days since the most recent earnings announcement, management forecast, 10-K filing, or 10-Q filing. Sources: Compustat, IBES, WRDS SEC Analytics

(Continued)

Appendix A: Variable Definitions (Continued)

Variable	*Definition*
Control Variables:	
Adj. ROA_{q-1}	Return on assets (i.e., earnings before extraordinary items divided by total assets) less the average return on assets for firms within the same two-digit standard industrial classification code, year, and quarter. Source: Compustat
Analyst $Coverage_{q-1}$	The natural logarithm of one plus the number of analysts who provided an EPS forecast between the prior quarter's earnings announcement and two days before the forecasted earnings announcement. Source: IBES
Business $Segments_{y-1}$	The number of segments with non-zero revenue in the Compustat Segments file as of the prior fiscal year-end. Source: Compustat
Institutional $Ownership_{q-1}$	The sum of shares owned by institutional investors scaled by the number of shares outstanding. This value is set equal to zero if no institutional ownership is reported and set equal to one if reported institutional ownership exceeds shares outstanding. Source: Backus et al. (2021) via Michael Sinkinson Data Library (https://sites.google.com/view/msinkinson/research/common-ownership-data)
M/B_{q-1}	Market value of equity scaled by book equity. Source: Compustat
Media $Coverage_{t-180,t}$	The natural logarithm of one plus the number of news articles about the firm within the past 180 days. Source: RavenPack Analytics Dow Jones Edition
$Returns_{m-12,m-1}$	The firm's returns over the 12-month period ending the month prior to the article publication date. Source: CRSP
$Returns_{t-10,t-1}$	The firm's returns over the 10-trading day period ending the day prior to the article publication date. Source: CRSP
$Size_{q-1}$	The natural logarithm of market value of equity. Source: Compustat

(Continued)

Appendix A: Variable Definitions (Continued)

Variable	*Definition*
Bunching Variables:	
Days to EA_t	The signed number of 24-hour blocks between the time of Seeking Alpha article publication and the earnings announcement rounded away from zero to the next integer. For example, an article published 26 hours prior to (after) an earnings announcement is classified as being two days prior to (after) an earnings announcement.
Fake $Articles_t$	The number of fake articles published on *Days to* EA_t summed across all earnings announcements and scaled by the total number of fake articles in the sample.
Non-Fake $Articles_t$	The number of non-fake articles published on *Days to* EA_t summed across all earnings announcements and scaled by the total number of non-fake articles in the sample.
Abnormal $Mass_t$	The difference between *Fake* $Articles_t$ and *Non-Fake* $Articles_t$.
Pre EA Abnormal $Mass_{t-2,t-1}$	The sum of *Abnormal* $Mass_t$ for days t-2 and t-1.
Post EA Abnormal $Mass_{t+1,t+8}$	The sum of *Abnormal* $Mass_t$ for days between t+1 and t+8.
Differential Abnormal $Mass_{t-2,t+8}$	The difference between *Pre EA Abnormal* $Mass_{t-2,t-1}$ and *Post EA Abnormal* $Mass_{t+1,t+8}$.
Total Abnormal $Mass_{t-2,t+8}$	The sum of *Abnormal* $Mass_t$ for days between t-2 and t+8.

This table presents the definitions for the primary variables used in our analyses. For the dependent variables, accounting information variables, and control variables, the *y*, *q*, *m*, and *t* subscripts represent year, quarter, month, and day, respectively, and represent when the variable is measured relative to article publication on day *t*. Unless otherwise noted, our dependent variables and accounting information variables are measured as of the article publication date. Analyst coverage is measured as of the most recent earnings announcement occurring on or before article publication. Accounting data and market values are measured as of the fiscal quarter-end in which the earnings announcement for the quarter occurs on or before article publication. For the bunching variables, *t* represents the event date relative to the earnings announcement occurring at $t = 0$.

Appendix B
Exhibit A: Fake News Article on Galena Biopharma

Seeking Alpha α
Read. Decide. Invest.

Home | Portfolio | Market Currents | Investing Ideas | Dividends & Income | ETFs | Macro View | ALERTS

Galena Biopharma: Best And Worst Case Scenario

Aug 14 2013, 04:09 | about: GALE BOOKMARK / READ LATER

Brian Nichols

3 on Quick Picks & Lists

Articles (441)
Instablog (4)
StockTalks (113)
Comments (1,985)
Profile

6,488 Followers
501 Following

Follow

Send Message

Disclosure: I am long GALE. **(More...)**

Last Thursday, Needham & Company initiated coverage on shares of **Galena Biopharma** (GALE) with a "Buy" rating and a $3.50 price target. According to streetinsider.com, Galena has "Buy" or "Outperform" ratings by all of the analysts who cover the stock. Yet, certain bloggers and retail investors have been quite cynical about the company's future. Thus, let's look at both the best and worst case scenario with Galena Biopharma to determine if the risk is worth the reward.

Abstral: Now Available

Abstral is a rapidly-dissolving sublingual tablet for the management of breakthrough pain. The drug is a best-in-class fentanyl product, with plasma concentrations of fentanyl seen within 10 min. This rapid absorption is what separates Abstral from other breakthrough pain and fentanyl drugs and is why many believe it will be a successful product.

According to Galena's recent quarterly report, Abstral is now available at nationwide pharmacies. Galena will now market the drug, and hopes to create a profit by next year.

Best Case

In Europe, Abstral produced sales of $54 million in 2012. In Q4 2012, Abstral sales grew 42% year-over-year, thus showing that it could be a success in the U.S.

In the U.S., the market for fentanyl products is $400 million annually, and Galena believes that it can control 10%-15% of the market within five years. This means that Abstral could generate annual sales of $60 million.

If Abstral is viewed as reliable and efficient, it is highly likely that Abstral could command an even larger share of the fentanyl market, or expand the market in size. At first glance, it appears as though Galena is being conservative with their guidance, as Abstral continues to grow and maintains a 30% market share in Europe.

If Galena can control a 30% market share in the U.S., sales could rise to $120 million. If we use a four times sales ratio then we arrive at a market capitalization of $480 million, or $5.75, and that's only accounting for Abstral.

Worst Case

The worst case is that Abstral is a dud and never reaches $60 million in sales. Currently, Galena is trading with a market cap of $150 million. On March 18, when Galena announced the acquisition of Abstral, it traded with a market cap of $166 million.

Due to Galena being cheaper today than it was in March, we can conclude that none of Galena's valuation is tied to Abstral's success or failure. This means that any upside will be viewed as a bonus, also suggesting that a failed Abstral campaign should not alter the stock.

NeuVax: Blockbuster Potential

NeuVax is a Phase 3 vaccine that is being tested to prevent breast cancer recurrence in the 50%-75% of patients who are not eligible for Herceptin. The vaccine targets those who have low to intermediate levels of HER2, while Herceptin targets those who have high levels of HER2.

In a Phase 2, 187 patient study, NeuVax reduced the risk of recurrence by 78% in patients who were node-positive. Compared to the control arm of the study, only 5.6% of patients recurred after 60 months compared to 25.9% of those who were not vaccinated with NeuVax. In the company's ongoing Phase 3 study, the company hopes to prove that NeuVax can keep patients from redeveloping the disease.

Best Case

As of now, we have no way of knowing how much revenue NeuVax can produce if proven successful in its Phase 3 study. NeuVax is being tested alone, with Herceptin, and as a booster. If all three studies are successful, then we know that NeuVax will be a blockbuster product.

In comparison, Herceptin generates $7 billion annually by targeting just one-fourth to one-third of breast cancer patients. If NeuVax is successful, it will target at least one-half of patients. Hence, it is not unreasonable to estimate $2 billion in peak sales, which is most likely conservative.

At $2 billion in sales, adding a four times sales multiple, Galena could be worth $8 billion long-term. After approval, and awaiting an FDA decision, if Galena trades at 0.5 times peak sales then it would support a $1 billion valuation. In other words, Galena has upside of 700% short-term, and over 5,000% long-term if NeuVax is proven successful.

Worst Case

If NeuVax fails, then Galena would fall sharply. Right now, all of the stock's valuation is tied into the potential of NeuVax, not quite accounting for any success or lack thereof in marketing Abstral.

If unsuccessful, a 50% loss should be expected, or a market cap of $75 million. The unknown piece of the puzzle will be sales of Abstral. If Abstral is successful and generates sales between $60 and $120 million, then Galena's market capitalization will likely carry a 100% to 300% increase from its current price.

This means that if Abstral is successful, Galena should trade higher as data progresses. Then, if NeuVax fails, Abstral's success should still carry a stock price that is greater than its current price.

Overall Outlook

In bringing this discussion back down to earth, let's pretend that you invest $7,500 in Galena, or purchased 4,000 shares. Below I have included a table to show how each of the discussed scenarios could play out in stock performance.

Situation	Stock Price	Return (rounded)
Abstral fails + NeuVax fails	$0.90 or less.	($3,600)
Abstral $60m + NeuVax fails	$2.88	$11,500
Abstral $120m + NeuVax fails	$5.75	$43,000
Abstral $60m + NeuVax @ 0.5x peak sales	$15	$60,000
Abstral $60m +NeuVax @ 4x peak sales	$100	$400,000

As charted, there aren't too many scenarios where an investment returns a loss, which may be the driving force behind positive sentiment from analysts. Granted, this is speculative and theoretical in using standard price times sales ratios, but given each situation it's reasonable that the noted stock price should follow.

Like I explained, the wild card is Abstral. We don't know how it will perform. But NeuVax looks to be a very compelling product, that when used on the appropriate patient population, produced significant results. After assessing the company, including its risk and reward, it is difficult to determine why some could be overly bearish, as the downside is extraordinarily minimal. The decision of whether to invest is then determined on your own assessment of risk: Is the risk worth the reward

Exhibit B: Article Disputing Bullish Sentiment on Galena Biopharma

We reproduce only the most relevant sections of this article for brevity. Full article is found here: https://web.archive.org/web/20140301202559/https://seekingalpha.com/article/1984371-galena-biopharma-numerous-red-flags-suggest-a-significant-overvaluation

Seeking Alpha α
Read. Decide. Invest.

Home | Portfolio | Breaking News | Investing Ideas | Dividends & Income | ETFs | Macro View | ALERTS | PRO

Matt Gravitt

Long/short equity, value, special situations, momentum

Profile | Send Message | + **Follow** (125)

Galena Biopharma: Numerous Red Flags Suggest A Significant Overvaluation

TOP IDEA Jan. 31, 2014 10:15 AM ET | About: GALE

Disclosure: I am short GALE. **(More...)**

Disclosure

I am short Galena Biopharma (GALE) in via the purchase of put options at varying expiration dates/strike prices.

Summary

> *"There is no training - classroom or otherwise - that can prepare for trading the last third of a move, whether it's the end of a bull market or the end of a bear market. There's typically no logic to it; irrationality reigns supreme..."* - Paul Tudor Jones

While pondering of a way to sum up Galena Biopharma (GALE), I remembered that quote from Paul Tudor Jones and realized how applicable it is to GALE. The ~200% rise GALE has seen over the past two months is nothing short of incredible. After trading range bound for over three months, shares of GALE saw a massive break-out in mid-November and momentum traders who got in on the move during the early stages were handsomely rewarded with triple digit returns by early January.

Now that the parabolic move has finally exhausted itself and momentum traders (along with company insiders) are exited out of their long positions, reality and logic can set in.

On the surface one *might* think that Galena is a good speculative biotech stock with a nice risk/reward profile. However, the purpose of this article is to shed some light on the numerous risks the GALE poses and present the case for why the risk/reward profile favors the short side.

Before we start moving on to the more interesting topics, there are a couple of key points that should be taken away here:

- Interim results the NueVax phase II trial data were presented in 2006 by Dr. George Peoples. (That is an important name to remember as well)
- It took 4 years of the market essentially writing off NeuVax as a legitimate therapy before they found a willing buyer.
- The combination of a spin-off, mergers, and name changes make it extremely difficult to track down accurate data/SEC filings.

Galena Lacks Vital Exclusivity Rights to NeuVax

Even PRESENT Phase III trial does prove to meet its agreed endpoint, and the FDA does not require an additional Phase III study (which is possible based on GALE's assessment of the situation stated in their latest 10K), Neuvax faces some difficult challenges as it relates to intellectual property.

Galena's Pipeline and "Partnerships" add very little value

The Press Releases concerning partnerships have been misleading.

In December 2012, GALE announced "signature of commercialization partnership with Teva in Israel". (Link to PR Here)

However, GALE's sec filings provide some clarity into the specifics of the "partnership" with Teva. *"Effective December 3, 2012, we entered into a license and supply agreement with ABIC Marketing Limited."*

This press release raises a red flag for a couple of reasons:

1. While the full financials terms haven't been disclosed, it appears GALE did not receive any upfront payment in the deal. It seems GALE will be entitled to royalties based on future sales of Neuvax in Israel. My question is, if Neuvax was the potential blockbuster drug like GALE management claims, then why would they agree to a deal with such miniscule financial benefit?
2. The timing of the press release, coupled with the terminology and verbiage used seems to be misguided.

Abstral

In fact, Galena warns about the potential for Generic competition in their 10Q filed in August 2013:

> ***"We may not be able to obtain and enforce patent rights or other intellectual property rights that cover Abstral and that are of sufficient breadth to prevent third parties from competing against us.***
>
> *Our success with respect to Abstral will depend in part on our ability to obtain and maintain patent protection in the United States, to preserve our trade secrets, and to prevent third parties from infringing upon our proprietary rights. Fentanyl, the sole active pharmaceutical ingredient, or "API," in Abstral, has been approved for many years and therefore our ability to obtain any patent protection is limited. Composition of matter*

> *patents are a particularly effective form of intellectual property protection for pharmaceutical products, as they apply without regard to any method of use. However, we will not be able to obtain composition of matter patents or methods of use patents that cover the APIs in Abstral. As a result, competitors who obtain the requisite regulatory approval can offer products with the same active ingredients as Abstral so long as the competitors do not infringe any formulation patents that we may have or may obtain or license, if any"*

It is also concerning that GALE limited their contractually obligated marketing responsibilities to a two-year span. Logic would dictate that if a company expected a drug to succeed and achieve growth in sales, they would have no problem committing to marketing it throughout its lifespan.

GALE states the following in their latest 10Q:

> *"Under our agreement with Orexo, we assumed responsibility for the U.S. commercialization of Abstral and for all regulatory and reporting matters in the U.S.* ***We also agreed to establish and maintain through 2015 a specified minimum commercial field force to market, sell and distribute Abstral and to use commercially reasonable efforts to reach the specified sales milestones.*** *Orexo is entitled to reacquire the U.S. rights to Abstral from us for no consideration if we breach our obligations to establish and maintain the requisite sales force throughout the marketing period."*

Paying Companies for Stock Promotion and Significant Insider Selling are Major Red Flags

Stock Promotion

When I first started investing in/trading biotech stocks, I was fortunate enough to have several trading mentors impart valuable insight and words of wisdom that has helped contribute to my success. These "words of wisdom" included a warning about stocks that are constantly "pumped" over the internet.

As outlined in a March 2012 article by SeekingAlpha contributor Michael Morhamus, GALE's moves higher can be partially attributed to heavy promotion of the stock via the internet. After reading the article, I thought it brought up some interesting points and additional research into the matter was warranted. Needless to say, I came across some noteworthy pieces of information as it relates to Galena's stock being "promoted" via various outlets.

However, further investigation revealed that Galena was paying for these promotions. This, for me, is definitely a red flag.

According to a disclaimer found on the tip.us website, MissionIR received compensation from "GALE for 240 days of advertising, branding, marketing, investor relations and social media services provided by MissionIR and affiliate DreamTeamGroup Business Brands."

This potentially explains a part of the massive increase in the company's SG&A Expenses (obviously a large part of the increase was due to the Abstral launch).

Selling, General and Administrative Expense

Selling, general and administrative expense includes compensation-related costs for our employees dedicated to sales and marketing, general and administrative activities, legal fees, audit and tax fees, consultants and professional services, and general corporate expenses. Selling, general and administrative expense for the three and nine month periods ended September 30, 2013 and 2012 , was as follows (dollars in thousands):

	Three Months Ended September 30,			Nine Months Ended September 30,		
	2013	2012	% Change	2013	2012	% Change
Selling, general and administrative expense	$ 4,129	$ 1,359	204%	$ 8,369	$ 5,068	65%

Selling, general and administrative expense increased $2.8 million for the three months ended September 30, 2013 , compared with the three months ended September 30, 2012 . The increase was primarily due to a $2.7 million increase in personnel related costs, associated with the establishment of our Abstral commercial force and marketing team, and professional and outside services, and a $0.2 million increase in non-cash employee stock based compensation expense.

Selling, general and administrative expense increased $3.3 million for the nine months ended September 30, 2013 , compared with the nine months ended September 30, 2012 . The increase was primarily due to a $3.3 million increase related to personnel related costs, associated with the establishment of our Abstral commercial force and marketing team, and professional and outside services, and a $ 0.4 million increase in non-cash employee stock based compensation expense, which was partially offset by a decrease of $ 0.4 million in non-cash non-employee stock based compensation expense.

Insider Selling

As shown in the illustration below, there has been significant insider selling in GALE since the beginning of 2014. The insider-selling activity was also called out in a recent article by SeekingAlpha contributor Markus Aarnio.

Transaction Date	Name	Position	Type	Shares	Range	Market Value	Total Holdings
01/22/2014	KRIEGSMAN, STEVEN A.	Director	Option Execute	250,000	$0.72 - $1.18	180.0K	255,000
01/22/2014	KRIEGSMAN, STEVEN A.	Director	Sell	250,000	$6.13	1.5M	5,000
01/17/2014	KRIEGSMAN, STEVEN A.	Director	Option Execute	200,000	$0.85	170.0K	205,000
01/17/2014	KRIEGSMAN, STEVEN A.	Director	Sell	200,000	$7.00	1.4M	5,000
01/17/2014	HILLSBERG, SANFORD	Director	Sell	200,000	$6.93	1.4M	110,447
01/17/2014	NISI, RUDOLPH	Director	Sell	200,000	$6.90	1.4M	3,500
01/15/2014	NISI, RUDOLPH	Director	Option Execute	200,000	$0.85 - $1.18	170.0K	203,500
01/14/2014	HILLSBERG, SANFORD	Director	Disposition (Non Open Market)	24,426	$6.96	170.0K	310,447
01/14/2014	HILLSBERG, SANFORD	Director	Option Execute	200,000	$0.85	170.0K	334,873

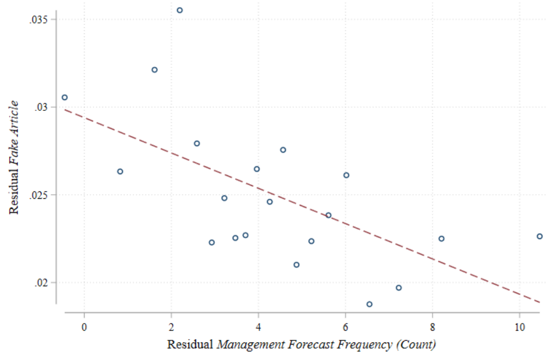
So if members of Senior Management (including the CEO) are liquidating significant portions of their holdings at current levels, what does that suggest about their views of future prices?

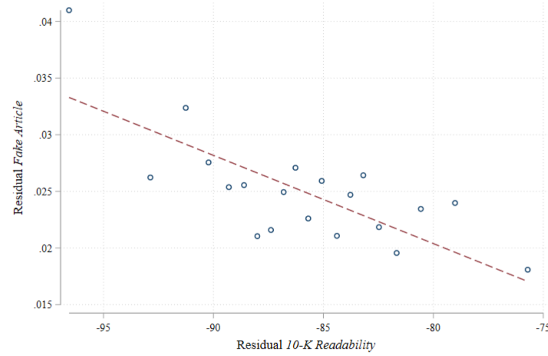
Per the latest 10Q, GALE has approximately 35M in shares reserved for future issuance. And if history is any indication of future actions, I am a strong believer that those warrants and options will be exercised well before they expire. Additionally, the large number of warrants add another negative dimension due to the potential for an increase in short interest. It is common to see an increase in selling pressure on stocks with a large number of warrants, let me explain. Once the stock price reaches a level that satisfies a warrant holder's expectations and desired return, they will short/sell the stock at those levels and essentially lock in a fixed return while using the warrants as a cover.

In the long term, GALE's stock price will be determined by most of the underlying factors I discuss in this article. And for reasons that I discuss in this article, I remain quite bearish in the long term.

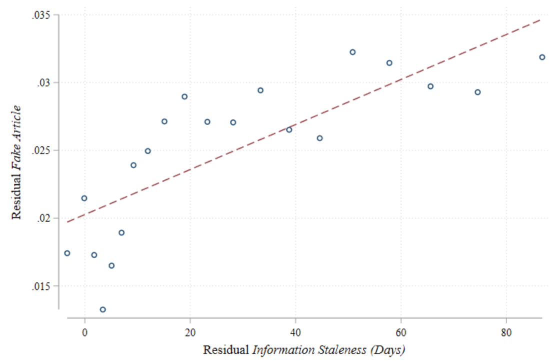
Figure 1: The Fake News Deterrence Role of Accounting Information

Panel A: Management Forecast Frequency (Count) *Panel B: 10-K Readability*





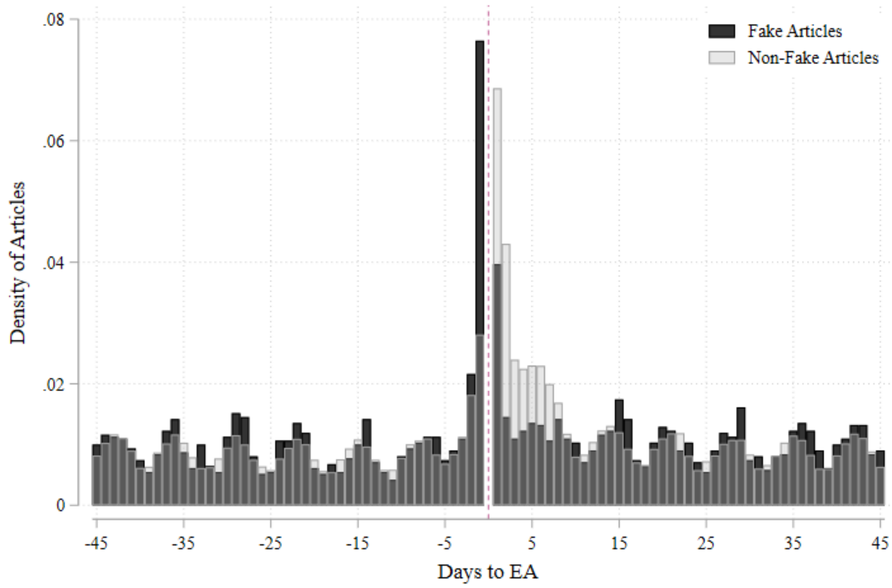
Panel C: Information Staleness (Days)



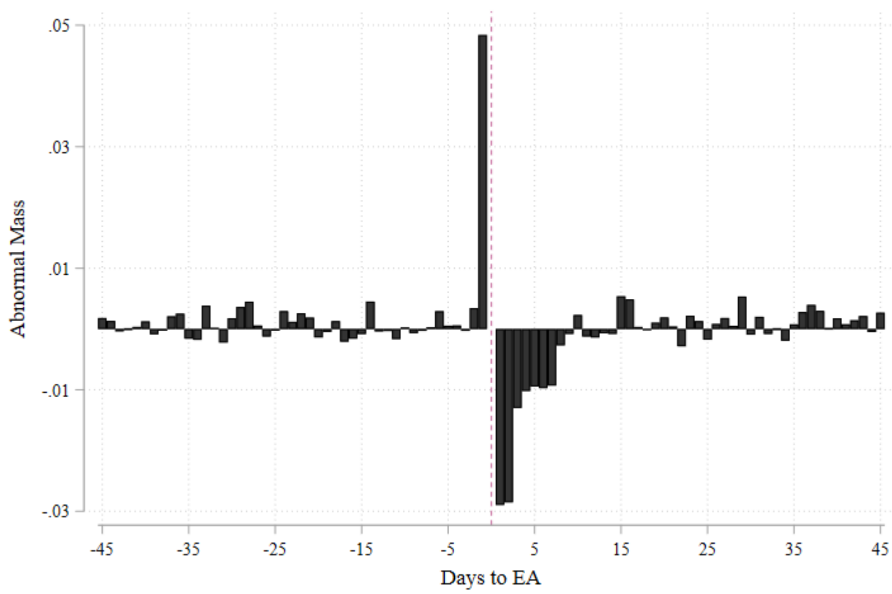
This figure plots the conditional probability of a fake news article versus our three measures of accounting information. Panel A is a binned scatterplot of the probability of a fake article (i.e., *Fake Article*) versus the number of management forecasts in the past year (i.e., *Management Forecast Frequency (Count)*). Panel B is a binned scatterplot of *Fake Article* versus the Bog Index from Bonsall et al. (2017) multiplied by -1 (i.e., *10-K Readability*). Panel C is a binned scatterplot of *Fake Article* versus the minimum number of days since the most recent earnings announcement, management forecast, 10-K filing, or 10-Q filing (i.e., *Information Staleness (Days)*). To construct these binned scatterplots, we first residualize both *Fake Article* and the respective accounting information variables (collectively referred to as *Accounting Information*) with respect to the control variables described in Table 3 as well as industry (two-digit SIC) and year fixed effects using partitioned regressions following the Frisch-Waugh-Lovell theorem. We then rank and divide the observations into 20 equal-size groups (ventiles) based on residual *Accounting Information* and plot the means of residual *Fake Article* within each bin against the mean value of residual *Accounting Information* within each bin. Finally, we add back the unconditional mean of *Fake Article* and *Accounting Information* in the estimation sample to facilitate interpretation of the scale. We use the binscatter Stata program for this procedure (Stepner, 2014). The dashed line shows the best linear fit estimated on the underlying sample of articles using an OLS regression. All three slope estimates are significantly different than zero at conventional significance levels (i.e., p-value < .05).

Figure 2: Bunching Analysis of Full Sample of Articles

Panel A: Distribution of Fake and Non-Fake Articles Around Earnings Announcements



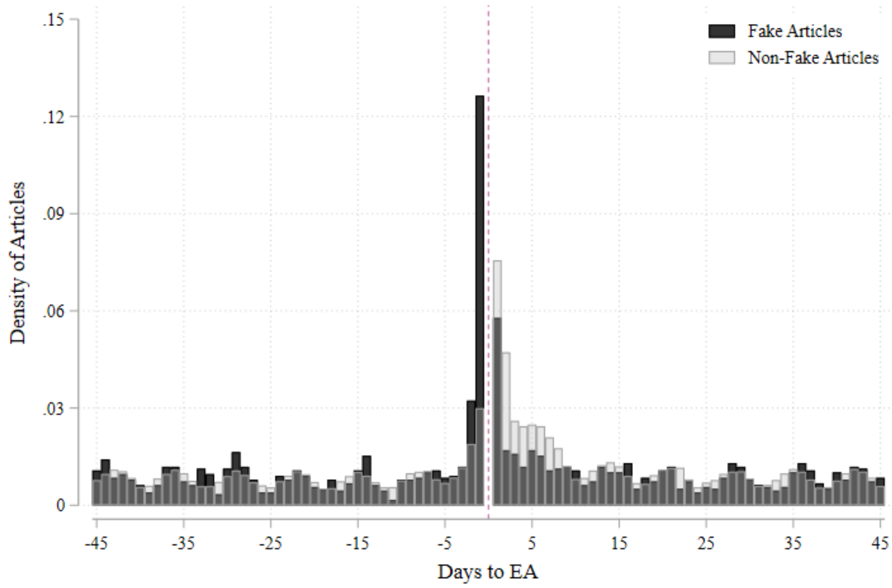
Panel B: Distribution of Abnormal Mass Around Earnings Announcements



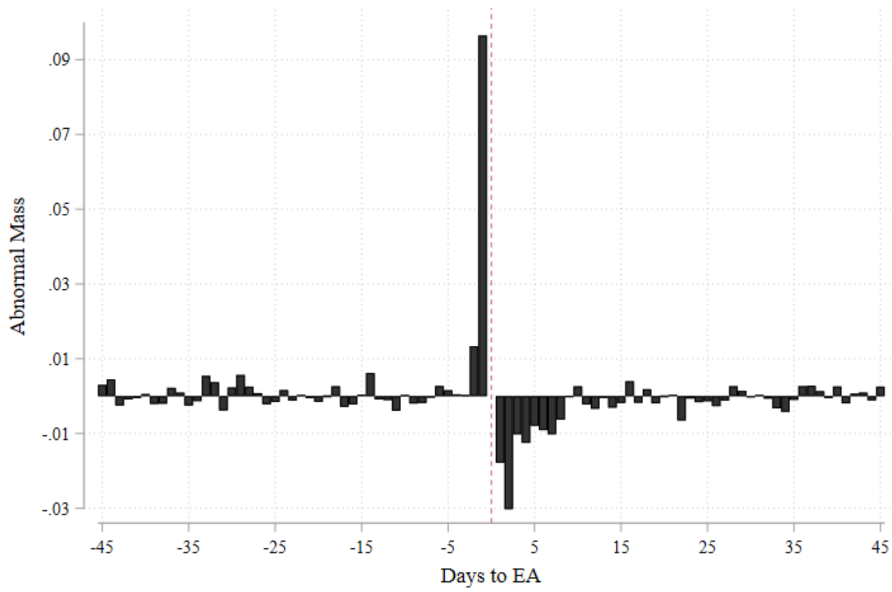
This figure presents graphical evidence from bunching analyses using the full sample of articles. Panel A plots the density distributions for both fake and non-fake articles around earnings announcements. Panel B plots *Abnormal Mass*, directly showing the difference between the fake and non-fake distributions depicted in Panel A.

Figure 3: Bunching Analysis of Articles with Accounting Content

Panel A: Distribution of Fake and Non-Fake Articles Around Earnings Announcement



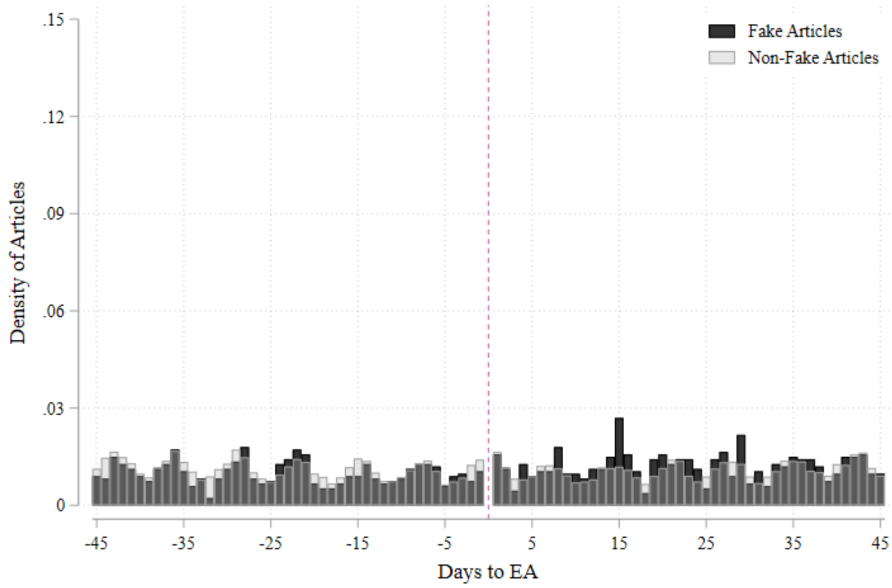
Panel B: Distribution of Abnormal Mass Around Earnings Announcement



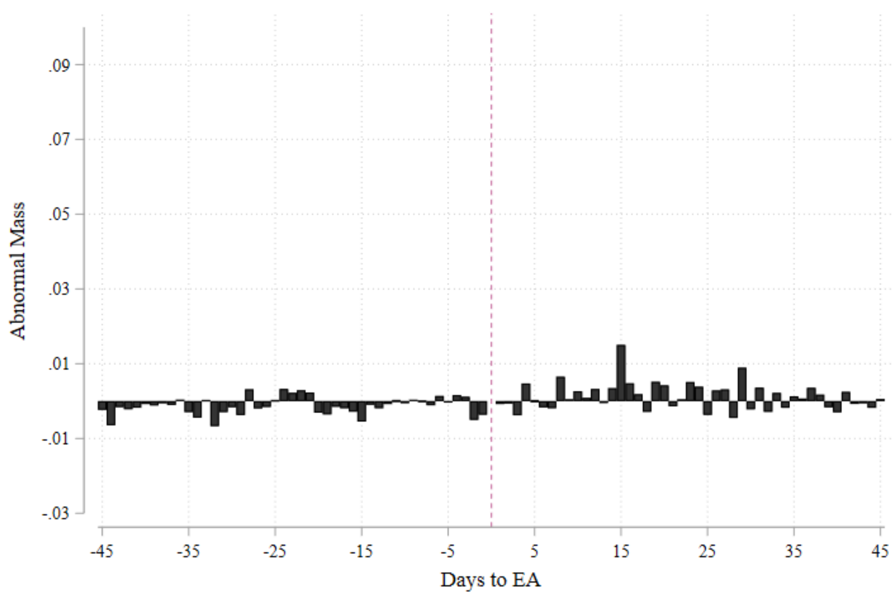
This figure presents graphical evidence from bunching analyses using only articles with accounting content. Panel A plots the density distributions for both fake and non-fake articles around earnings announcements. Panel B plots *Abnormal Mass*, directly showing the difference between the fake and non-fake distributions depicted in Panel A.

Figure 4: Bunching Analysis of Articles without Accounting Content

Panel A: Distribution of Fake and Non-Fake Articles Around Earnings Announcement



Panel B: Distribution of Abnormal Mass Around Earnings Announcement



This figure presents graphical evidence from bunching analyses using only articles without accounting content. Panel A plots the density distributions for both fake and non-fake articles around earnings announcements. Panel B plots *Abnormal Mass*, directly showing the difference between the fake and non-fake distributions depicted in Panel A.

Table 1: Sample Selection

Sample Selection Criteria	*# of Articles*	*# of Firm-quarters*
Seeking Alpha articles (January 1st, 2006 – December 31st, 2018)	221,103	
Exclude: Articles without more than 100 words	(2,789)	
Exclude: Articles that cannot be classified as fake or non-fake	(86,205)	
Exclude: Articles missing 10-K Readability	(4,440)	
Exclude: Missing firm-level controls	(2,194)	
Article sample	125,475	37,864

This table lists the sample selection criteria for Seeking Alpha articles. The starting point for our sample is a file, provided by Seeking Alpha, of all published Seeking Alpha articles from January 1st, 2006 – December 31st, 2018 that match to a CRSP historical stock ticker with a CRSP share code of 10 or 11. To exclude conference call transcripts and other news releases we require that the article is not written by a Seeking Alpha editor or other staff member. We have an initial sample of 221,103 articles. We retain articles with more than 100 words and those that we can classify as either fake or non-fake using the methodology in Kogan et al. (2020), excluding 2,789 and 86,205 articles, respectively. The Bog Index from Bonsal et al. (2017) is available for 10-Ks filed on or prior to March 31st, 2018, and requiring this variable eliminates 4,440 articles. Requiring the control variables used in our primary analyses eliminates an additional 2,194 articles. Our final sample comprises of 125,475 articles and 37,864 firm-quarters. The exact number of observations in regression analyses will differ slightly because we drop observations for which the fixed effects perfectly predict the dependent variables from estimation samples as needed across different models.

Table 2: Latent Dirichlet Allocation Textual Analysis

Topic #	*Topic Label*	*# of Articles*	*Fake %*
Topic 1	Fiscal Policy	428	7%
Topic 2	Green Technology	2,739	4%
Topic 3	Energy	5,685	1%
Topic 4	Passive Management	234	6%
Topic 5	Accounting	22,356	1%
Topic 6	Retail Industry	10,307	1%
Topic 7	Streaming Services	978	3%
Topic 8	Real Estate	526	1%
Topic 9	Macroeconomy	2,911	1%
Topic 10	Entertainment Industry	993	2%
Topic 11	Graphical Evidence	4,506	0%
Topic 12	Precious Metals	785	2%
Topic 13	Mobile Device Technology	1,725	4%
Topic 14	Unclassified / General	14,841	1%
Topic 15	Healthcare	1,784	3%
Topic 16	Risk Modeling	4,503	1%
Topic 17	General Business	7,236	3%
Topic 18	Legal	1,372	17%
Topic 19	Portfolio Management	1,122	4%
Topic 20	Dividend Investing	4,192	0%
Topic 21	Bonds	409	3%
Topic 22	Capital Raises	3,013	19%
Topic 23	Social Media	3,695	4%
Topic 24	Technology Industry	4,023	4%
Topic 25	Accounting Forecasts	17,255	3%
Topic 26	Global Markets	486	1%
Topic 27	Pharmaceutical Industry	2,994	10%
Topic 28	Financial Services Industry	2,130	5%
Topic 29	Foreign Currency Exchange	91	9%
Topic 30	E-Commerce	2,145	3%

This table presents descriptive statistics for topics identified using Latent Dirichlet Allocation (LDA). Topic # is the original topic number designated by LDA. Topic Label is a descriptive name for the topic based on researcher examination of the most prominent words for the topic. # of Articles is the number of articles which have the highest probability of containing content in that topic. Fake % is the percentage of fake articles within all the articles assigned to that topic. In later analyses, we use Topic 5 (i.e., Accounting) and Topic 25 (i.e., Accounting Forecasts) to define whether articles contain accounting content. We have highlighted these topics in the table.

Table 3: Descriptive Statistics for Primary Regression Variables

Variable	*N*	*Mean*	*Std. Dev.*	*P1*	*P25*	*Median*	*P75*	*P99*
Dependent Variables:								
Fake Article$_{t}$	125,475	0.025	0.156					
# of Fake Articles$_{q}$	37,864	0.088	0.328	0	0	0	0	2
Abnormal Volume$_{t,t+2}$	1,380	3.353	2.626	0.669	2.057	2.692	3.694	19.493
Idiosyncratic Return Volatility$_{t,t+2}$ (%)	1,380	0.208	0.504	0.001	0.019	0.054	0.162	3.576
Accounting Information Variables:								
Management Forecast Frequency$_{t-365,t}$	125,475	1.447	0.767	0.000	1.099	1.609	1.946	2.639
10-K Readability$_{y-1}$	125,475	-85.941	6.216	-102	-90	-86	-81	-72
Information Staleness$_{t}$	125,475	2.744	1.346	0.000	1.792	3.091	3.892	4.564
Control Variables:								
Adj. ROA$_{q-1}$	125,475	0.020	0.046	-0.158	-0.000	0.014	0.038	0.170
Analyst Coverage$_{q-1}$	125,475	2.762	0.797	0.000	2.485	2.944	3.296	3.932
Business Segments$_{y-1}$	125,475	1.731	1.785	0.000	1.000	1.000	3.000	8.000
Institutional Ownership$_{q-1}$	125,475	0.680	0.217	0.000	0.582	0.700	0.832	1.000
M/B$_{q-1}$	125,475	4.834	8.128	-24.339	1.534	3.019	5.610	46.692
Media Coverage$_{t-180,t}$	125,475	3.777	1.270	0.000	3.045	3.871	4.682	6.198
Returns$_{m-12,m-1}$	125,475	0.161	0.496	-0.762	-0.111	0.102	0.338	2.506
Returns$_{t-10,t-1}$	125,475	0.004	0.090	-0.288	-0.037	0.004	0.044	0.329
Size$_{q-1}$	125,475	9.575	2.296	3.931	7.903	9.853	11.510	13.348

This table presents descriptive statistics for variables used in the regression analyses. The *y*, *q*, *m*, and *t* subscripts represent year, quarter, month, and day, respectively, and indicate when the variable is measured relative to article publication on day *t*. Our dependent variables are *Fake Article*, *# of Fake Articles*, *Abnormal Volume*, and *Idiosyncratic Return Volatility*. Our primary independent variables are three distinct measures of accounting information: (1) *Management Forecast Frequency*, (2) *10-K Readability*, and (3) *Information Staleness*. Dependent variables are tabulated at the same level as the analysis in which they appear. Independent variables are tabulated at the article level. The definitions for all these variables can be found in Appendix A. Except for variables with natural lower or upper bounds, we winsorize all variables at the 1st and 99th percentiles.

Table 4: The Fake News Deterrence Role of Accounting Information

Fake Article as Dependent Variable	*(1)*	*(2)*	*(3)*	*(4)*
Accounting Information Variables:				
Management Forecast Frequency	-0.279***			-0.220***
	(-4.21)			(-3.01)
10-K Readability		-0.042***		-0.041***
		(3.82)		(4.43)
Information Staleness			0.271***	0.253***
			(10.59)	(10.32)
Control Variables:				
Adj. ROA	-2.900***	-2.480**	-3.106***	-1.686
	(-2.79)	(-2.27)	(-2.96)	(-1.61)
Analyst Coverage	-0.126	-0.217**	-0.185*	-0.125
	(-1.23)	(-1.98)	(-1.79)	(-1.22)
Business Segments	0.049	0.016	0.039	0.033
	(1.12)	(0.37)	(0.95)	(0.81)
Institutional Ownership	0.256	-0.007	0.182	0.177
	(0.99)	(-0.03)	(0.75)	(0.72)
M/B	-0.005	-0.004	-0.005	-0.005
	(-0.86)	(-0.58)	(-0.77)	(-0.81)
Media Coverage	0.266***	0.285***	0.283***	0.258***
	(4.00)	(4.43)	(4.41)	(4.15)
$Returns_{m-12,m-1}$	-0.248***	-0.243***	-0.240***	-0.252***
	(-2.81)	(-2.81)	(-2.80)	(-2.98)
$Returns_{t-10,t-1}$	0.578**	0.590**	0.605**	0.563**
	(2.08)	(2.13)	(2.16)	(2.05)
Size	-0.070	-0.072	-0.076*	-0.069*
	(-1.61)	(-1.63)	(-1.77)	(-1.66)
Industry & Year Fixed Effects	Included	Included	Included	Included
Mean of *Fake Article* (%)	2.50	2.50	2.50	2.50
Economic Magnitude (%)	-8.6	-10.4	14.6	-
Pseudo R^2	0.116	0.116	0.119	0.122
N	124,602	124,602	124,602	124,602
Model	Logit	Logit	Logit	Logit

(Continued)

Table 4 (Continued)

This table reports analyses on the effect of accounting information on the probability of fake news. The dependent variable is *Fake Article*, which is an indicator variable equal to one when the article is classified as fake and equal to zero for non-fake articles using the methodology in Kogan et al. (2020). Our primary independent variables of interest are three distinct measures of accounting information: (1) *Management Forecast Frequency* is the natural logarithm of one plus the number of management forecasts in the last year. (2) *10-K Readability* is the Bog Index from Bonsall et al. (2017) multiplied by -1. (3) *Information Staleness* is the natural logarithm of one plus the minimum number of days since the most recent earnings announcement, management forecast, 10-K filing, or 10-Q filing. See Appendix A for details on the remaining variables. The table reports marginal effect estimates from a logit regression and (in parentheses) *z*-statistics based on robust standard errors clustered by firm. Marginal effect estimates are calculated at the means of the regressors and multiplied by 100 so that they can be interpreted directly as percentage changes. We include industry (two-digit SIC) and year fixed effects in the regressions as indicated but do not report the coefficients. Observations for which the fixed effects perfectly predict *Fake Article* are dropped from the estimation sample. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% p-levels (two-tailed), respectively. The reported economic magnitude is calculated by multiplying the estimated coefficient by the standard deviation of the *Accounting Information* variable and then scaled by the mean of the dependent variable.

Table 5: Robustness Tests of the Fake News Deterrence Role of Accounting Information

Fake Article as Dependent Variable	*(1)* *Accounting Content Articles*	*(2)* *No Accounting Content Articles*	*(3)* *Firm-year Fixed Effects*	*(4)* *Negative Earnings* $Surprise_{q-1}$	*(5)* *Positive Earnings* $Surprise_{q-1}$
Accounting Information Variables:					
Management Forecast Frequency	-0.116**	-0.286		-0.263***	-0.230***
	(-2.08)	(-0.83)		(-2.60)	(-2.93)
10-K Readability	-0.033***	-0.013		-0.071***	-0.035***
	(4.99)	(0.25)		(4.75)	(3.63)
Information Staleness	0.179***	-0.148	0.827***	0.201***	0.250***
	(7.66)	(-0.89)	(7.48)	(4.16)	(9.07)
Control Variables:					
Adj. ROA	-1.270*	6.196	-0.396	-0.971	-2.093*
	(-1.67)	(1.07)	(-0.06)	(-0.52)	(-1.83)
Analyst Coverage	0.007	-1.394***	-0.335	-0.199	-0.125
	(0.09)	(-2.58)	(-0.29)	(-0.96)	(-1.20)
Business Segments	0.029	0.121	-0.304	0.049	0.034
	(1.22)	(0.59)	(-0.55)	(0.92)	(0.76)
Institutional Ownership	0.254	2.358*	3.858**	0.094	0.326
	(1.50)	(1.66)	(2.37)	(0.23)	(1.31)
M/B	-0.005	-0.015	-0.003	-0.002	-0.006
	(-0.91)	(-0.52)	(-0.13)	(-0.17)	(-0.80)
Media Coverage	0.111***	0.928***	1.047*	0.221**	0.258***
	(2.71)	(2.69)	(1.90)	(2.25)	(3.64)
$Returns_{m-12,m-1}$	-0.049	-1.747***	-0.430	-0.271	-0.251***
	(-0.78)	(-3.94)	(-0.99)	(-1.58)	(-2.72)
$Returns_{t-10,t-1}$	0.650***	-0.028	2.648	1.972***	-0.274
	(2.62)	(-0.02)	(1.47)	(3.76)	(-0.80)
Size	-0.045	-0.500**	0.070	-0.061	-0.059
	(-1.55)	(-2.27)	(0.07)	(-0.79)	(-1.32)
Industry & Year Fixed Effects	Included	Included	-	Included	Included
Firm-Year Fixed Effects	-	-	Included	-	-
Pseudo R^2 (Adj. R^2 – Col. 3)	0.121	0.099	0.138	0.139	0.124
N	108,614	15,602	44,560	30,038	84,523
Model	Logit	Logit	OLS	Logit	Logit

(Continued)

Table 5 (Continued)

This table reports robustness analyses corresponding to specifications in Table 4. The dependent variable is *Fake Article*. We present results for three analyses. (1) In Columns 1 and 2, we estimate our model partitioning by whether the article contains accounting content (i.e., has positive probability associated with Topic 5 or 25 based on LDA analysis). (2) In Column 3, we utilize a firm-year fixed effects specification. The sample in this column includes only firm-years for which there was at least one fake and one non-fake article written about the firm. (3) In Columns 4 and 5, we estimate our model partitioning by negative or positive earnings surprise. Earnings surprise is defined as actual earnings less the median analyst forecast. See Appendix A for other variable definitions. Column 3 reports OLS coefficient estimates and (in parentheses) *t*-statistics based on robust standard errors clustered by firm. For all other specifications, the table reports marginal effect estimates from a logit regression and (in parentheses) *z*-statistics based on robust standard errors clustered by firm. Marginal effect estimates are calculated at the means of the regressors and multiplied by 100 so that they can be interpreted directly as percentage changes. We include fixed effects in the regressions as indicated but do not report the coefficients. Observations for which the fixed effects perfectly predict *Fake Article* are dropped from the estimation sample. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% p-levels (two-tailed), respectively.

Table 6: Subsample Tests of the Fake News Deterrence Role of Accounting Information

Fake Article as Dependent Variable	Coefficient Estimates for: (1) *Management Forecast Frequency*	(2) *10-K Readability*	(3) *Information Staleness*	(4) *# of Observations*
Management Forecast Provision				
None		-0.071***	0.472***	20,536
		(2.65)	(5.43)	
One or more		-0.033***	0.223***	103,036
		(3.37)	(8.64)	
Analyst Coverage				
Low	-0.005	-0.061***	0.235***	59,512
	(-0.05)	(5.23)	(6.12)	
High	-0.291***	-0.029**	0.248***	63.437
	(-3.01)	(2.19)	(7.55)	
Institutional Ownership %				
Low	-0.298***	-0.030**	0.281***	61,655
	(-2.85)	(2.25)	(8.44)	
High	-0.161*	-0.049***	0.232***	61,887
	(-1.67)	(4.61)	(6.57)	
Size				
Small	0.012	-0.055***	0.197***	61,869
	(0.13)	(4.14)	(5.12)	
Large	-0.344***	-0.037***	0.254***	62,449
	(-4.50)	(3.88)	(9.24)	

This table reports subsample analyses corresponding to the specification presented in Table 4 Column 4. The coefficients for the accounting information variables are reported in columns 1-3 as indicated for each subsample analysis. The dependent variable is *Fake Article*. All subsample analyses include the control variables and fixed effects specified in Table 4 Column 4, but we do not report the coefficients for brevity. The management forecast provision subsamples are partitioned by whether the firm provides at least one management forecast in the past year. Additionally, we exclude *Management Forecast Frequency* as an independent variable from these subsamples to avoid collinearity issues. The analyst coverage, institutional ownership, and size subsamples are created by partitioning at the median for each of these characteristics, respectively. See Appendix A for other variable definitions. The table reports marginal effect estimates from a logit regression and (in parentheses) *z*-statistics based on robust standard errors clustered by firm. Marginal effect estimates are calculated at the means of the regressors and multiplied by 100 so that they can be interpreted directly as percentage changes. Observations for which the fixed effects perfectly predict *Fake Article* are dropped from the estimation sample. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% p-levels (two-tailed), respectively.

Table 7: The Fake News Deterrence Role of Accounting Information

# of Fake Articles as Dependent Variable	*(1)*	*(2)*	*(3)*	*(4)*
Accounting Information Variables:				
Management Forecast Frequency	-0.111***			-0.095**
	(-2.60)			(-2.49)
10-K Readability		-0.026***		-0.026***
		(6.05)		(5.81)
Information Staleness			0.091***	0.082***
			(4.94)	(4.52)
Control Variables:				
Adj. ROA	-2.043***	-1.725***	-2.159***	-1.470***
	(-4.52)	(-3.83)	(-4.73)	(-3.31)
Analyst Coverage	0.033	-0.013	0.008	0.028
	(0.63)	(-0.23)	(0.15)	(0.55)
Business Segments	0.014	0.001	0.009	0.006
	(0.74)	(0.09)	(0.50)	(0.33)
Institutional Ownership	-0.132	-0.214**	-0.167	-0.184*
	(-1.21)	(-2.02)	(-1.55)	(-1.72)
M/B	-0.001	-0.000	-0.000	-0.000
	(-0.21)	(-0.02)	(-0.13)	(-0.10)
Media Coverage	0.124***	0.118***	0.126***	0.119***
	(3.63)	(3.50)	(3.70)	(3.53)
$Returns_{m-12,m-1}$	-0.141***	-0.136***	-0.138***	-0.145***
	(-2.94)	(-2.91)	(-2.90)	(-3.07)
$Returns_{t-10,t-1}$	0.129	0.151	0.141	0.119
	(0.69)	(0.82)	(0.75)	(0.65)
Size	0.101***	0.099***	0.100***	0.104***
	(3.88)	(3.74)	(3.76)	(4.02)
Additional Variables:				
Seeking Alpha Articles	0.044***	0.046***	0.044***	0.044***
	(25.36)	(26.97)	(26.22)	(25.17)
Industry & Year Fixed Effects	Included	Included	Included	Included
Economic Magnitude (%)	-8.9	-18.5	11.0	-
Pseudo R^2	0.207	0.209	0.207	0.210
N	37,690	37,690	37,690	37,690
Model	Poisson	Poisson	Poisson	Poisson

(Continued)

Table 7 (Continued)

The table reports analyses on the effect of accounting information on the quantity of fake news. The observations in this analysis are aggregated to the firm-quarter level. The dependent variable is *# of Fake Articles*, which is a count of fake articles written about the firm in a quarter. In this analysis, *Management Forecast Frequency*, *10-K Readability*, and the control variables are measured as of the first article of the quarter. *Information Staleness* is measured as the average Information Staleness of all fake and non-fake articles about the firm in the quarter. In addition to the control variables described in Table 3, we also include *Seeking Alpha Articles,* which is the number of Seeking Alpha articles written about the firm in a quarter. See Appendix A for other variable definitions. The table reports Poisson pseudo-maximum likelihood regression coefficient estimates and (in parentheses) z-statistics based on robust standard errors clustered by firm. We include industry (two-digit SIC) and year fixed effects in the regressions as indicated, but do not report the coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% p-levels (two-tailed), respectively. The reported economic magnitude is calculated by exponentiating the reported Poisson regression coefficient multiplied by the standard deviation of the *Accounting Information* variable (untabulated), and then subtracting off one.

Table 8: The Impact of Accounting Information on the Market Reaction to Fake News

Panel A: Trade-based Market Reaction

Abnormal Volume$_{t,t+2}$ as Dependent Variable	*(1)*	*(2)*	*(3)*	*(4)*
Accounting Information Variables:				
Management Forecast Frequency	-0.126			-0.109
	(-1.60)			(-1.54)
10-K Readability		-0.026**		-0.025**
		(2.53)		(2.49)
Information Staleness			0.157**	0.144**
			(2.27)	(2.05)
Control Variables:				
Adj. ROA	-1.672	-1.466	-1.892*	-1.323
	(-1.58)	(-1.37)	(-1.76)	(-1.24)
Analyst Coverage	-0.281**	-0.311**	-0.315**	-0.275**
	(-2.02)	(-2.20)	(-2.22)	(-1.97)
Business Segments	-0.009	-0.029	-0.014	-0.024
	(-0.31)	(-0.93)	(-0.49)	(-0.76)
Institutional Ownership	0.496	0.326	0.455	0.399
	(1.26)	(0.86)	(1.16)	(1.05)
M/B	0.002	0.002	0.003	0.002
	(0.57)	(0.58)	(0.76)	(0.62)
Media Coverage	-0.086	-0.078	-0.085	-0.082
	(-1.12)	(-1.02)	(-1.09)	(-1.08)
Returns$_{m-12,m-1}$	-0.245**	-0.256**	-0.237**	-0.270**
	(-2.00)	(-2.09)	(-1.97)	(-2.23)
Returns$_{t-10,t-1}$	0.067	0.017	0.099	0.070
	(0.08)	(0.02)	(0.12)	(0.09)
Size	0.101*	0.104*	0.105*	0.102*
	(1.80)	(1.84)	(1.86)	(1.81)
Lagged Abnormal Volume Variables	Included	Included	Included	Included
Industry & Year Fixed Effects	Included	Included	Included	Included
Economic Magnitude (%)	-3.2	-6.2	3.5	-
Adjusted R^2	0.504	0.506	0.505	0.508
N	1,371	1,371	1,371	1,371
Model	OLS	OLS	OLS	OLS

(Continued)

Table 8: The Impact of Accounting Information on the Market Reaction to Fake News

Panel B: Price-based Market Reaction

Idiosyncratic Return Volatility$_{t,t+2}$ as Dependent Variable	*(1)*	*(2)*	*(3)*	*(4)*
Accounting Information Variables:				
Management Forecast Frequency	-0.040*			-0.035*
	(-1.79)			(-1.72)
10-K Readability		-0.006***		-0.006***
		(2.60)		(2.60)
Information Staleness			0.048***	0.044***
			(3.03)	(2.82)
Control Variables:				
Adj. ROA	-0.929***	-0.901***	-0.998***	-0.858**
	(-2.74)	(-2.64)	(-2.90)	(-2.54)
Analyst Coverage	-0.025	-0.036	-0.037	-0.025
	(-0.57)	(-0.81)	(-0.84)	(-0.56)
Business Segments	0.006	0.001	0.004	0.002
	(0.71)	(0.10)	(0.55)	(0.29)
Institutional Ownership	0.029	-0.015	0.018	0.008
	(0.33)	(-0.17)	(0.20)	(0.09)
M/B	-0.001	-0.001	-0.001	-0.001
	(-1.18)	(-1.13)	(-1.07)	(-1.18)
Media Coverage	0.035	0.038*	0.036	0.037*
	(1.58)	(1.73)	(1.59)	(1.69)
Returns$_{m-12,m-1}$	-0.014	-0.016	-0.011	-0.020
	(-0.39)	(-0.43)	(-0.30)	(-0.54)
Returns$_{t-10,t-1}$	-0.250	-0.261	-0.242	-0.243
	(-1.18)	(-1.24)	(-1.14)	(-1.16)
Size	-0.029*	-0.028*	-0.027*	-0.029*
	(-1.75)	(-1.71)	(-1.67)	(-1.75)
Lagged Idiosyncratic Return Volatility Variables	Included	Included	Included	Included
Industry & Year Fixed Effects	Included	Included	Included	Included
Economic Magnitude (%)	-16.5	-23.0	17.4	-
Adjusted R^2	0.246	0.248	0.249	0.254
N	1,370	1,370	1,370	1,370
Model	OLS	OLS	OLS	OLS

(Continued)

Table 8 (Continued)

The table reports analyses on the effect of accounting information on the market's trading reaction (Panel A) and price reaction (Panel B) to fake news. Articles published within two days of an earnings announcement, management forecast, 10-K, 10-Q, or 8-K are excluded from the analysis because we cannot attribute the market reaction to the Seeking Alpha article. Similarly, we exclude days when both a fake and non-fake article are published. In Panel A, our dependent variable is *Abnormal Volume*, which is the sum of the scaled volume on the day of publication and the following two trading days. Scaled volume is defined as volume scaled by the average volume between 20 and 140 trading days prior. The dependent variable in Panel B is *Idiosyncratic Return Volatility*, which is the sum of the squared abnormal returns on the day of publication and the following two trading days. Abnormal return is defined as a firm's daily return minus the daily return on a 5x5x5 size-, B/M-, and momentum-matched portfolio. In addition to the *Accounting Information* and *Control Variables* described in Table 3, we include lagged one-day measures of our dependent variables to control for serial correlation and unobserved confounding events, but do not report the coefficients. Panel A includes *Abnormal Volume*$_{t-1}$, *Abnormal Volume*$_{t-2}$, and *Abnormal Volume*$_{t-3}$, which are the scaled trading volumes for the three trading days prior to article publication. In Panel B, we include *Idiosyncratic Return Volatility*$_{t-1}$, *Idiosyncratic Return Volatility*$_{t-2}$, and *Idiosyncratic Return Volatility*$_{t-3}$, which are the squared abnormal returns for the three trading days prior to article publication. See Appendix A for other variable definitions. The table reports OLS coefficient estimates and (in parentheses) *t*-statistics based on robust standard errors clustered by firm. We include industry (two-digit SIC) and year fixed effects in the regressions as indicated, but do not report the coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% p-levels (two-tailed), respectively. The reported economic magnitude is calculated by multiplying the estimated coefficient by the standard deviation of the *Accounting Information* variable (untabulated) and then scaled by the mean of the dependent variable.

Table 9: Bunching Analyses Examining Fake News Publication Timing Preferences

Window of Interest	*(1)* Pre EA Abnormal $Mass_{t-2,t-1}$	*(2)* Post EA Abnormal $Mass_{t+1,t+8}$	*(3)* Differential Abnormal $Mass_{t-2,t+8}$	*(4)* Total Abnormal $Mass_{t-2,t+8}$
Full Sample of Articles	0.052**	-0.112***	0.164***	-0.060***
	(5.07)	(-5.36)	(6.95)	(-2.61)
Sample of Articles with Accounting Content	0.110***	-0.104***	0.214***	0.006
	(7.26)	(-3.26)	(6.00)	(0.16)
Sample of Articles without Accounting Content	-0.009*	0.003	-0.011	-0.006
	(-1.83)	(0.27)	(-1.08)	(-0.60)

This table reports the results from bunching analyses examining the publication timing preferences of fake news authors in an event window around earnings announcements. *Pre EA Abnormal* $Mass_{t-2,t-1}$ is the sum of *Abnormal* $Mass_t$ for days t-1 and t-2. *Post EA Abnormal* $Mass_{t+1,t+8}$ is the sum of *Abnormal* $Mass_t$ for days between t+1 and t+8. *Differential Abnormal* $Mass_{t-2,t+8}$ is the difference between *Pre EA Abnormal* $Mass_{t-2,t-1}$ and *Post EA Abnormal* $Mass_{t+1,t+8}$. *Total Abnormal* $Mass_{t-2,t+8}$ is the sum of *Abnormal* $Mass_t$ for each day between t-2 and t+8. The full sample estimates use the entire sample of fake and non-fake articles from earlier analyses published within 45 days of an earnings announcement. The accounting content subsamples are created using the same LDA methodology as the accounting content partitions described in Section 4.2.1 (Table 5). The table reports effect estimates and (in parentheses) *t*-statistics based on standard errors calculated using a bootstrap procedure following Chetty et al. (2011). ***, **, and * indicate statistical significance at the 1%, 5%, and 10% p-levels (two-tailed), respectively.

Internet Appendix for
"The Role of Accounting Information in an Era of Fake News"

Betty Liu
Tippie College of Business
University of Iowa
yunchen-liu@uiowa.edu

Austin Moss
Tippie College of Business
University of Iowa
austin-moss@uiowa.edu

This internet appendix contains additional discussion and analyses referenced in the main paper.

IA1. Latent Dirichlet Allocation (LDA) implementation details

IA2. LDA Top 10 most prominent words per topic

IA3. Sensitivity analyses

IA1: Latent Dirichlet Allocation (LDA) implementation details

Latent Dirichlet Allocation (LDA) is a natural language processing technique used to identify latent topics in a collection of documents and assign these documents to the most relevant detected topics. More specifically, LDA uses unsupervised machine learning to compute statistics about the likelihood of certain words appearing concurrently in a passage of text and imputes as topics words that tend to occur together. By training the LDA algorithm on large repositories of text, LDA can also be used to assign documents (or even portions of documents, if multiple topics exist within a document) both inside and outside the training sample to the identified topics. In the context of this paper, we use LDA on Seeking Alpha articles to shed light on the type of topics covered by these articles and to generate statistics on how fake news articles span the identified topics.

We detail our LDA methodology below, using common benchmarks and thresholds as input to steps within the algorithm. We treat each Seeking Alpha article as an individual document for our analysis.

Preliminary Cleaning

We first separate each document into a list of words. We stem each word into its root form (e.g., "education" becomes "educat") and discard any common stop words (e.g., "the", "about", etc.), words fewer than 3 letters, as well as numbers and symbols. These lists of words are combined into a dictionary of words with an accompanying count of how many documents each word appears in. We filter out extremely common and uncommon words by discarding any words that appear in fewer than 15 documents and in more than 50% of the documents. From these words, we identify the 100,000 most frequently used words and keep only these words in each document's word list for analysis.

Estimating the Number of Topics Present

LDA requires researcher input on how many topics to identify in a corpus. We want to avoid running LDA for too low or too high a topic number, as the resulting model may not appropriately assign words into semantically coherent topics. As a first step, we generate LDA models for a wide range of potential topic numbers from 10 topics up to 100 topics in multiples of 10 (e.g., 10, 20, 30, etc.). To determine the model with the appropriate number of topics, we use the C_v coherence score to evaluate the coherence of the salient words within the topics generated by each of the models (Roder, Both, and Hinneburg, 2015; Syed and Spruit, 2017). A high coherence score implies that the salient words in determining a topic produce well-defined topics across documents and guards against selecting models that produce topics that are mere statistical artifacts. We compute the coherence score for each of our models. In Figure IA1, we plot the coherence scores and find that the coherence score peaks at around 30 topics. Thus, we keep the model utilizing 30 topics as input for our LDA analysis.

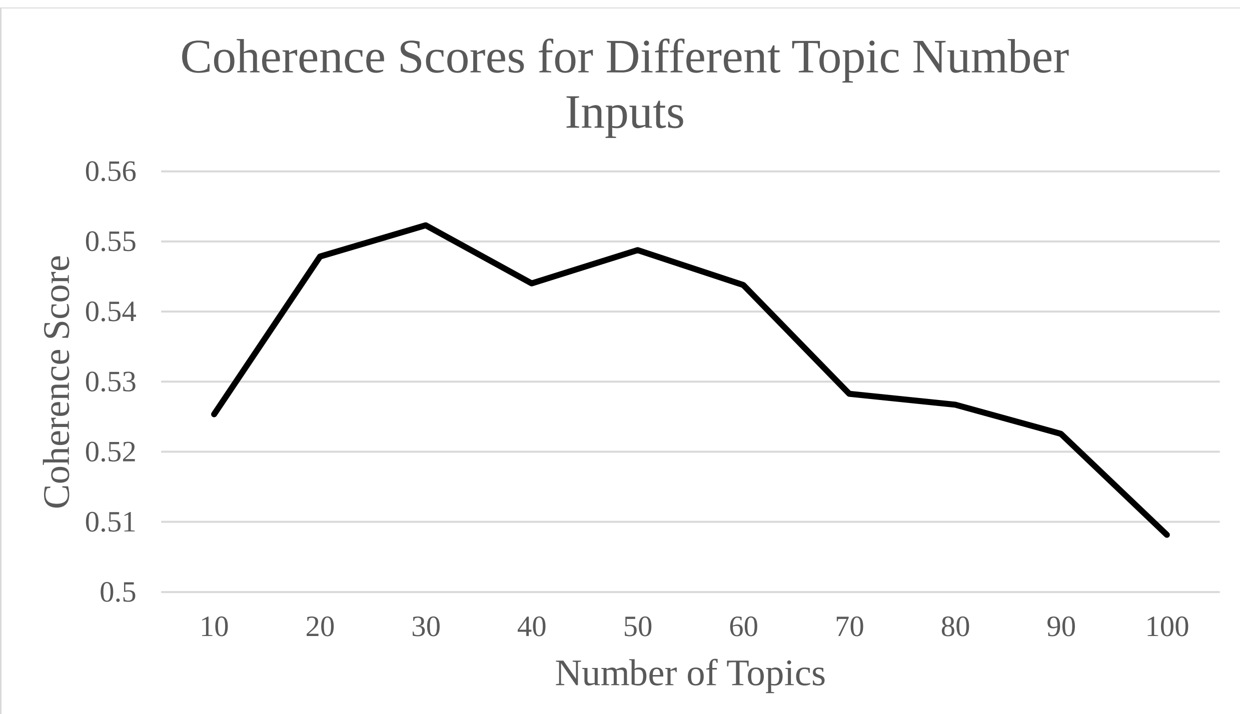
Identifying Representative Topics

While LDA can detect which words belong in latent topics, it is up to the researchers to assign a label or theme to the group of words that comprise a detected topic. Hence, using the LDA model with 30 identified topics, we want to assign each topic a relevant title. To do so, we obtain the top 10 words predictive of a particular topic. Both authors independently went through each set of words and assigned overarching themes to each of the set of the words, compared the potential topic names together, and selected the most suitable topic name for each set of words. We include the sets of words and the topics identified for each of the 30 topics in IA2.

Assigning Topics to Articles

We categorize each article into a specific topic to facilitate descriptive statistics. For each article, the LDA algorithm will generate a list of probabilities of that article containing content in each of the topics. We assign each article to the topic with the highest probability for our descriptive statistics in Table 2. In addition, we compute two numbers for each topic: (1) the number of Seeking Alpha articles that is assigned to that topic and (2) the percentage of fake articles out of the articles classified as that topic.

Figure IA1: Coherence Scores for LDA Implementation



IA2: LDA Top 10 most prominent words per topic

Topic 1 Fiscal Policy

polici	1.33%
economi	1.30%
econom	1.19%
trump	1.01%
govern	0.98%
inflat	0.89%
debt	0.73%
money	0.72%
financi	0.66%
central	0.63%

Topic 2 Green Technology

electr	2.32%
vehicl	2.11%
power	2.02%
industri	1.79%
model	1.58%
energi	1.56%
cost	1.56%
solar	1.45%
manufactur	1.44%
car	1.30%

Topic 3 Energy

product	3.46%
energi	2.32%
natur	1.53%
barrel	1.26%
crude	1.14%
produc	1.14%
million	0.94%
drill	0.85%
demand	0.80%
suppli	0.77%

Topic 4 Passive Management

fund	3.60%
index	3.10%
sector	2.63%
return	2.03%
perform	1.87%
portfolio	1.60%
volatil	1.37%
etf	1.24%
hold	1.21%
analyst	1.19%

Topic 5 Accounting

cash	2.97%
valu	2.25%
flow	2.02%
valuat	1.72%
earn	1.56%
debt	1.50%
margin	1.45%
oper	1.40%
ratio	1.18%
capit	1.08%

Topic 6 Retail Industry

sale	3.62%
store	3.27%
brand	2.26%
retail	2.19%
food	1.31%
product	1.27%
consum	1.10%
custom	0.76%
ford	0.71%
busi	0.70%

Topic 7 Streaming Services

subscrib	5.16%
netflix	5.01%
content	3.57%
stream	2.84%
gilead	2.72%
servic	2.14%
subscript	1.76%
nflx	1.58%
gild	1.42%
warner	1.37%

Topic 8 Real Estate

reit	5.43%
properti	3.10%
real	2.57%
estat	2.42%
home	2.16%
hous	1.68%
leas	1.62%
mortgag	1.40%
trust	1.03%
rent	0.97%

Topic 9 Macroeconomy

declin	1.48%
percent	1.46%
data	1.30%
rise	1.28%
rate	1.18%
remain	1.13%
econom	1.11%
report	1.06%
demand	1.00%
economi	0.97%

Topic 10	
Entertainment Industry	
game	2.40%
disney	2.10%
hotel	1.40%
sport	1.13%
movi	1.07%
entertain	1.03%
travel	0.94%
revenu	0.92%
million	0.84%
film	0.82%

Topic 11	
Graphical Evidence	
week	3.85%
chart	2.09%
level	1.59%
short	1.46%
averag	1.29%
click	1.21%
enlarg	1.11%
move	1.00%
indic	0.98%
higher	0.91%

Topic 12	
Precious Metals	
gold	12.46%
silver	3.15%
metal	2.94%
mine	2.65%
miner	1.91%
product	1.77%
copper	1.54%
project	1.47%
ounc	1.45%
resourc	1.19%

Topic 13	
Mobile Device Technology	
mobil	3.30%
game	2.23%
micron	1.50%
verizon	1.35%
qualcomm	1.21%
wireless	1.17%
network	1.06%
semiconductor	1.04%
tencent	1.02%
billion	0.91%

Topic 14	
Unclassified / General	
go	2.01%
think	2.00%
good	1.18%
thing	1.02%
sell	0.97%
right	0.90%
want	0.86%
know	0.85%
say	0.83%
point	0.73%

Topic 15	
Healthcare	
product	2.30%
boe	2.26%
healthcar	2.24%
sale	1.97%
health	1.90%
medic	1.83%
drug	1.54%
care	1.37%
order	1.29%
airbus	1.21%

Topic 16	
Risk Modeling	
risk	1.28%
articl	0.76%
valu	0.71%
differ	0.71%
model	0.67%
return	0.59%
strategi	0.58%
chang	0.56%
import	0.53%
futur	0.52%

Topic 17	
General Business	
busi	2.42%
servic	2.08%
revenu	1.80%
custom	1.72%
product	1.71%
technolog	1.70%
provid	1.03%
manag	1.02%
industri	1.00%
data	0.97%

Topic 18	
Legal	
report	1.58%
say	1.27%
legal	0.89%
regul	0.85%
claim	0.81%
state	0.79%
street	0.73%
court	0.71%
case	0.71%
issu	0.67%

Topic 19

Portfolio Management

fund	4.79%
manag	3.70%
portfolio	3.18%
asset	2.47%
hedg	1.67%
capit	1.67%
hold	1.65%
valu	1.56%
stake	1.34%
berkshir	1.33%

Topic 20

Dividend Investing

dividend	14.46%
yield	4.94%
incom	3.01%
portfolio	1.97%
payout	1.45%
ratio	1.24%
return	1.19%
distribut	1.18%
pay	1.18%
annual	1.16%

Topic 21

Bonds

bond	4.49%
yield	4.29%
rat	4.26%
rate	3.38%
risk	2.24%
treasuri	1.72%
fund	1.45%
inflat	1.37%
rise	1.27%
asset	1.23%

Topic 22

Capital Raises

million	3.08%
cash	1.26%
deal	1.12%
sharehold	1.06%
manag	1.00%
offer	0.94%
capit	0.93%
sell	0.85%
debt	0.84%
valu	0.83%

Topic 23

Social Media

googl	3.33%
user	2.57%
facebook	2.40%
advertis	1.63%
platform	1.34%
media	1.34%
revenu	1.34%
video	1.01%
content	0.93%
social	0.91%

Topic 24

Technology Industry

appl	9.39%
intel	3.20%
iphon	2.34%
aapl	2.19%
nvidia	1.87%
product	1.61%
devic	1.28%
sale	1.17%
cola	1.01%
chip	0.98%

Topic 25

Accounting Forecasts

quarter	7.39%
million	5.12%
revenu	4.99%
earn	3.99%
billion	2.93%
report	2.37%
sale	1.99%
result	1.94%
estim	1.68%
guidanc	1.38%

Topic 26

Global Markets

china	5.67%
global	2.53%
countri	2.39%
chines	2.16%
world	1.73%
currenc	1.70%
emerg	1.32%
dollar	0.98%
export	0.97%
foreign	0.96%

Topic 27

Pharmaceutical Industry

patient	2.26%
drug	1.77%
trial	1.73%
phase	1.52%
approv	1.32%
studi	1.27%
treatment	1.18%
data	1.05%
develop	0.99%
clinic	0.97%

Topic 28 Financial Services Industry		**Topic 29** Foreign Currency Exchange		**Topic 30** E-Commerce	
bank	13.26%	dollar	2.96%	amazon	6.59%
loan	4.61%	european	1.70%	microsoft	2.78%
financi	3.16%	euro	1.59%	onlin	2.11%
credit	2.85%	week	1.56%	amzn	1.96%
capit	1.38%	bank	1.25%	alibaba	1.83%
asset	1.30%	europ	1.16%	commerc	1.78%
lend	1.30%	currenc	1.10%	retail	1.74%
billion	1.24%	meet	0.91%	payment	1.43%
insur	1.21%	hike	0.76%	busi	1.29%
deposit	1.14%	itali	0.72%	billion	1.16%

IA3: Sensitivity Analyses

Table IA3: Sensitivity Analyses of the Fake News Deterrence Role of Accounting Information

	Coefficient Estimates for:			
	(1)	*(2)*	*(3)*	*(4)*
Fake Article as Dependent Variable	*Management Forecast Frequency*	*10-K Readability*	*Information Staleness*	*# of Observations*
Time Period Subsamples				
Exclude 6 months following 2014 and 2017 scandals	-0.206***	-0.044***	0.239***	108,414
	(-2.78)	(-5.00)	(9.30)	
Exclude pre-2012	-0.280***	-0.031***	0.184***	100,656
	(-3.68)	(3.50)	(7.77)	
Alternative Management Forecast Frequency Windows				
180-day Window	-0.254***	-0.041***	0.249***	124,602
	(-2.62)	(4.33)	(10.04)	
90-day Window	-0.334***	-0.041***	0.241***	124,602
	(-2.82)	(4.26)	(9.61)	
Other Sensitivity Analyses				
Exclude industry-years with fewer than 50 observations	-0.223***	-0.034***	0.246***	116,879
	(-3.01)	(-3.66)	(9.89)	

This table reports sensitivity analyses of our primary specification (Col. 4 in Table 4) examining the role of accounting information in deterring fake news. The dependent variable is *Fake Article*. The coefficients for the accounting information variables are reported in columns 1-3 as indicated for each sensitivity analysis. All analyses include the control variables and fixed effects specified in Table 4 Column 4, but we do not report the coefficients for brevity. See Appendix A for variable definitions. The table reports marginal effect estimates from a logit regression and (in parentheses) *z*-statistics based on robust standard errors clustered by firm. Marginal effect estimates are calculated at the means of the regressors and multiplied by 100 so that they can be interpreted directly as percentage changes. Observations for which the fixed effects perfectly predict *Fake Article* are dropped from the estimation sample. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% p-levels (two-tailed), respectively.