Do Non-professional Investors Influence Managers and Financial Analysts?

Wanyun Li

This thesis is presented for the degree of

Doctor of Philosophy

Research School of Accounting

College of Business & Economics

The Australian National University

June 2022

Declaration of Originality

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any university and that to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed:__________________On 19/06/2022

Wanyun Li

Acknowledgements

Throughout the several years that have gone into this thesis, support, enthusiasm, and effort have come from many quarters. I owe so many people that I don’t know where to start. But there are people without whom this thesis simply would not have come to pass, and I’ll focus on them.

First of all, I would like to thank my principal supervisor and Chair of the Supervision Panel, Prof Shuk Ying (Susanna) Ho, who led me into the research world and has guided me through this PhD journey. Thank you for always being available on campus and via email whenever I had questions, especially during the most challenging time in the Covid-19 lockdown. Thank you for spending lots of time meticulously reviewing and revising my drafts. Thank you also for tolerating my many mistakes throughout these years and encouraging me to move forward. I am incredibly blessed to have had you as my supervisor since the master program.

I would also like to thank the members of my Supervision Panel, Prof Mark Wilson and Dr Ka Wai (Stanley) Choi. Prof Mark Wilson taught me to be rigorous in accessing, processing, and analyzing data. Dr Stanley Choi helped me collect data and perform textual analyses. They challenged me to think critically in every aspect of my research. This thesis would not have been possible without them.

My sincere thanks go to all other academics at the Research School of Accounting for their helpful comments and constructive criticisms at different stages of my research. I am also grateful to all my fellow PhD students, particularly Kathy Jiang, for taking this PhD journey with me. Knowing that a companion was there to help when I couldn’t cope with academic rigor gave me so much joy. Also, to Zeo Zhang and Kenuo Li, I appreciate the time we spent working, dining, and shopping together. Thank you for bringing me peace and delight in my time of need.

Moreover, I am indebted to my family and friends. I am appreciative of my mother, Rong Jia, for always being a role model in my life and showing me what it takes to accomplish whatever I want in life; of my father, Pingsheng Li, for being my friend and sharing your amazing sense of humor with me; of my cousin, Qing Wu, for lending an ear when nobody else would and for the bond we share; of Da Li, for sharing the house with me and being there at times when I got stressed out or worried; and of Sugar, dear puppy, for laughter.

Finally, I am thankful for all of it. The highs. The lows. The blessings. The lessons. The setbacks. The comebacks. Everything.

This thesis was edited by Elite Editing, and editorial intervention was restricted to Standards D and E of the *Australian Standards for Editing Practice*.

Information for the Examiners

Dear Examiners,

I would like to thank you for agreeing to examine my thesis. I understand that you are all eminent professors and examining a thesis requires significant investment of your time and intellectual inputs. I have thus prepared the following with the hope of saving you some precious time:

- Digital copy of the thesis (pdf)
- PowerPoint presentation slides
- PowerPoint presentation slides (with audio walkthrough)

Thank you again for examining my thesis.

Best regards,

Wanyun Li

Abstract

With the advent of online social networks, individuals' sentiments and opinions about firms are more accessible and processable. These individuals are often non-professional investors (NPIs). This thesis investigates whether NPIs' sentiments and opinions about particular firms as expressed on two types of social network—namely social media and crowdsourcing platforms—are noise or information; and whether their sentiments and opinions can help management and financial analysts make disclosures about future earnings.

The first objective of this thesis is to investigate whether managers consider NPI sentiment toward their firms in issuing optimistic earnings guidance to promote desirable market reactions. I infer firm-level NPI sentiment from social media discussions concerning 3,212 distinct firms on StockTwits, which published over 17 million tweets from 118,685 users between May 2008 and January 2017. Following Aboody et al. (2018), I adopt overnight stock return to measure NPI reaction and find that NPI reaction to positive guidance is stronger when sentiment is high, and that managers are more likely to issue positive guidance at these times. This association between the likelihood of issuing positive guidance and NPI sentiment is stronger in firms in which NPIs have greater proportionate shareholdings and where managers' equity incentives are highly contingent on short-term stock price increases. These findings are consistent with managers opportunistically manipulating guidance to exploit NPI sentiment, and contribute to research by creating a research avenue on the role of NPI sentiment in management decision making.

The second objective of this thesis is to investigate whether financial analysts exploit NPIs' earnings expectations for a firm to 'walk down' their earnings forecasts with an aim to create beatable forecasts. I infer NPIs' earnings expectations from crowdsourced earnings estimates on Estimize, which contains 879,015 'street earnings' estimates submitted by 70,926 investors

for the period January 2012 to September 2018. I document a positive association between analyst forecast revision and a change in investors' earnings expectations. I also find that the likelihood of analysts issuing forecasts that generate pessimistic errors is high when they revise their forecasts down. These findings are consistent with analysts opportunistically manipulating earnings forecasts to exploit investors' expectations of future earnings. This study extends the literature on how market forces constrain analysts' conflicts of interest by demonstrating that new crowdsourcing technologies disrupt the market for traditional earnings forecast providers. It also informs regulators on how sharing their opinions on these public platforms may potentially expose NPIs to exploitation by more sophisticated market participants who possess the processing power to harness the enormous volume of these scattered pieces of information.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2SLS	Two-stage least squares
AFE	Absolute forecast errors
CRSP	Center for Research in Security Prices
EPS	Earnings per share
GDP	Gross Domestic Product
IT	Information technology
MCSI	Michigan Consumer Sentiment Index
NPI	Non-professional investor
OLS	Ordinary least squares
PSM	Propensity score matching
Reg FD	Regulation Fair Disclosure
SEC	Securities and Exchange Commission
VADER	Valence Aware Dictionary and sEntiment Reasoner
NYU	New York University
PCE	Personal Consumption Expenditure
SIC	Standard Industrial Classification
US	United States

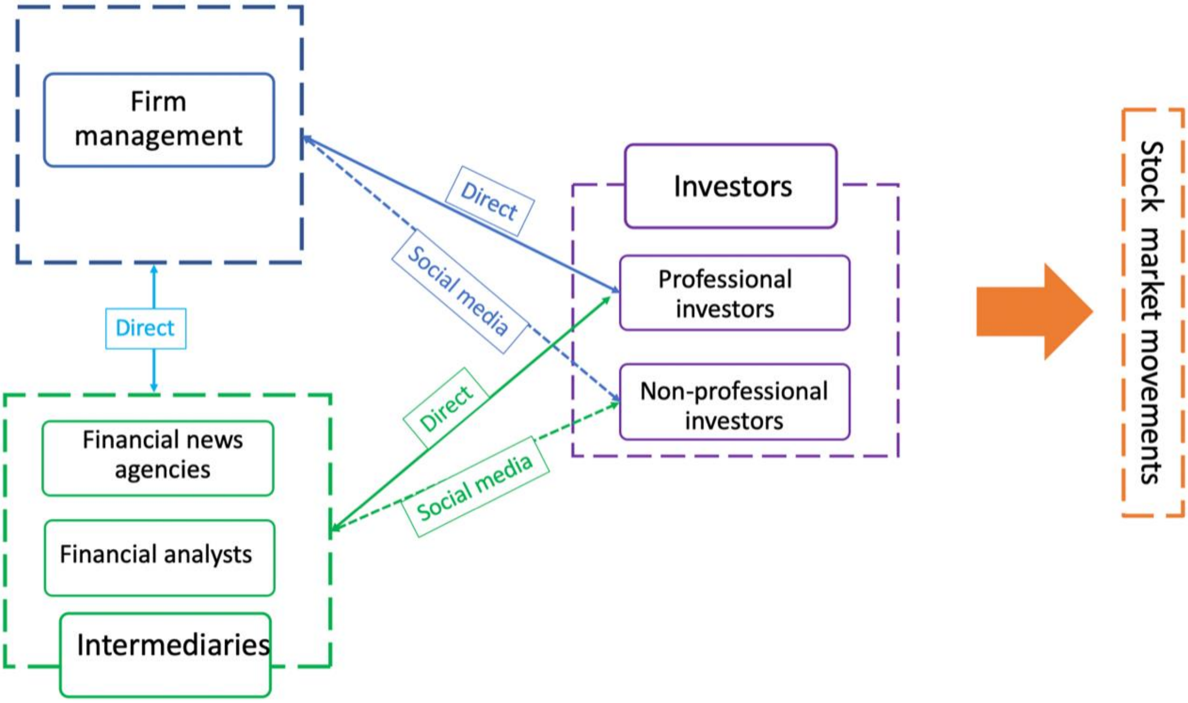
Chapter 1 Introduction

1.1 Research Questions and Motivations

This thesis examines the role of investors' opinions about firms' financial performance in capital market phenomena. The capital market consists of three key players: managers who run firms in a manner that maximizes profit for their investors (Jensen & Meckling, 1976; Obeng et al., 2020); investors who identify firms seeking to expand their businesses and invest capital in them (Jensen & Meckling, 1976; Obeng et al., 2020); and information intermediaries that facilitate information sharing between investors and firm management (Maber et al., 2020). There are two types of investor: professional investors and non-professional investors (NPIs). This thesis focuses on NPIs, who are largely neglected in the accounting literature.

Figure 1 presents a conceptual diagram that illustrates relations among the three players. In this thesis, I focus on the key player—the investor—who procures transactions affecting the forces of demand and supply in the capital market (Brown et al., 2015). Firms provide investors with reports and disclosures to help investors make investment decisions, and investors purchase shares of firms with strong future prospects, which promotes the growth of firms. Financial intermediaries act between investors and firms. Financial analysts (hereafter, analysts) and news agencies obtain information from firms, run analyses, and generate reports for investors. In the traditional view of the capital market, investors are information seekers. Extending this traditional view, I focus on whether and how the information extracted from investors' opinions alters firms' and financial intermediaries' subsequent information to investors. I leverage recent technology—online social networks—to collect these real-time opinions from investors. My thesis aims to investigate whether and how investors' opinions about firms' financial performance that are spread via online social networks provide useful information to the other two players: firm management and information intermediaries.

Figure 1 Main participants in capital markets



Investors can be broadly classified as professional or sophisticated investors, and NPIs. Professional or sophisticated investors are experienced and able to evaluate investment opportunities without needing a prospectus or other regulated disclosure documents (New York Stock Exchange [NYSE], 2016). They are often employed in the financial industry or registered with a financial regulatory body (e.g., the Securities and Exchange Commission [SEC] or stock exchange) (NYSE, 2016). Professional investors can interact with managers and intermediaries via one-to-one meetings, investment conferences, and other shareholder communication modes (Palter et al., 2008). In contrast, NPIs trade with their own money (NYSE, 2016). They are not registered with any financial regulatory body (NYSE, 2016) and have no direct line to management (Cade, 2018). Although NPIs are large in number, they often lack a voice with the firms in which they invest. The recent advent of online social networks facilitates interactions between NPIs and managers, and between NPIs and intermediaries. That is, online social networks enable NPIs to publish and share their opinions about firm prospects (Cade, 2018). Further, online social networks allow management and information intermediaries to access and aggregate NPIs' opinions about their firms as expressed on these social networks (Hales et al., 2018).

I divide my examination into two parts and consider the three parties in Figure 1 in pairs. First, I focus on NPIs and managers. The disclosure of management earnings guidance (hereafter, management guidance) is an important means by which managers communicate anticipated future performance, and is thus useful to NPIs when making investment decisions. For example, in the study of Beyer et al. (2010), management guidance accounted for around 55% of accounting information used by investors on average. Managers may consider NPIs' opinions about their firm's prospects when preparing earnings guidance because they want to correct investors' opinions and minimize investors' misevaluation of their firm (Brown et al., 2012; Seybert & Yang, 2012; Hurwitz, 2018) or because they want to exploit investors' opinions to influence their stock prices in pursuit of private benefits (e.g., increasing the worth of their equity incentives) (Balakrishnan et al., 2014; D'Augusta, 2022). Chapter 1.2 explains my rationale for why NPIs' opinions about a firm's prospects (particularly the sentiment of NPIs regarding the firm) disseminated on online social networks may influence the firm's managers when preparing earnings guidance.

Second, I consider NPIs and analysts. According to Figure 1, there are two types of information intermediary in the capital market: financial news agents and analysts. Financial news agents spend considerable time and effort disseminating news about firm earnings and other relevant economic indicators (Dzieliński, 2017; Tetlock, 2007). Like financial news agents, analysts also collect and disseminate information on firm earnings (Groysberg et al., 2011; Kothari et al., 2016; Maber et al., 2020). More importantly, they interpret firms' accounting disclosures (e.g., 10-K or 10-Q reports, earnings guidance, proxy statements, and press releases), gather additional information, and produce earnings forecasts to assist investors with investment decisions (Kothari et al., 2016; Maber et al., 2020). Analysts may consider NPIs' opinions about firm's prospects when preparing earnings forecasts because they seek to use NPIs' opinions to improve their own forecasts (Jame et al., 2021) or exploit investors' opinions to affect their stock prices in pursuit of private benefits (e.g., better access to management) (Bradley et al., 2020; Lourie, 2019). Chapter 1.3 explains my rationale around why NPIs' opinions about a firm's prospects disseminated on online social networks are useful to analysts when preparing earnings forecasts.

1.2 NPI Sentiment and Management Earnings Guidance

Regarding NPIs and management, I investigate whether and how the likelihood of a manager issuing positive earnings guidance (hereafter 'positive guidance') is associated with the sentiment of NPIs toward the firm. 'Positive' ('negative') guidance occurs where management's forecast level of earnings is higher (lower) than the prevailing consensus analyst forecast for the same financial period (Billings et al., 2015; Han & Tan, 2010). Managers have self-interested incentives to opportunistically influence the level and timing of earnings guidance provided, including the desire to temporarily boost or maintain stock price and increase liquidity (Balakrishnan et al., 2014). Investor sentiment, in particular NPI sentiment, plausibly conditions investors' reactions to earnings guidance.

I follow Baker and Wurgler (2006) in characterizing investor sentiment as reflecting the propensity to speculate. Speculation is particularly common when investors are confronted with relatively incomplete information regarding a security (Baker & Wurgler, 2006, 2007). When making investment decisions, investors—especially NPIs—often let their sentiment alter their interpretation of corporate information whenever consistencies lead to psychological conflicts in beliefs (Brown et al., 2012; Simpson, 2013). That is, in periods of high (low) NPI sentiment, NPIs perceive positive guidance as relatively credible (incredible) because it aligns (contrasts) with their existing beliefs about a firm's future performance, and this leads to stronger (weaker) NPI reactions.

I explore whether managers consider positive NPI sentiment to positively bias earnings guidance hoping for a stronger NPI reaction . Specifically, I test the association between NPI sentiment and the likelihood that guidance bundled with earnings announcements is positive, after controlling market- and total firm-level sentiment. I infer firm-level NPI sentiment from individual discussions on StockTwits (stocktwits.com), which publishes over 17 million tweets posted by 118,685 users concerning 3,212 distinct firms between May 2008 and January 2017. StockTwits is a social media platform whose content focuses on financial discussions between NPIs; it was used in earlier research such as Deng et al. (2018) and Renault (2017) when examining intraday stock returns. I first show that NPI reaction to positive guidance is conditioned by NPI sentiment. NPI reaction to positive guidance is measured from stock returns during the overnight period following issuance of positive guidance. I use the overnight return because NPIs tend to place orders outside of regular working hours (Aboody et al., 2018; Berkman et al., 2012). I regress the overnight return on the interaction between positive guidance and NPI sentiment using an ordinary least squares (OLS) regression, the relevant main effects, and controls. I then show that the likelihood of managers issuing positive guidance bundled with earnings announcements increases with NPI sentiment. I estimate a logistic regression for the likelihood of managers issuing bundled positive guidance in the current quarter on the NPI sentiment with the relevant controls. This association between the likelihood of issuing positive guidance and NPI sentiment is stronger for firms in which NPIs have greater proportionate shareholdings and where managers' equity incentives are highly contingent on short-term stock price increases. The findings are consistent with managers opportunistically manipulating guidance to exploit the sentiment of NPIs.

This investigation has the potential to contribute to the literature in three ways. First, while much of the literature regards NPI sentiment as 'noise' and dedicates little attention to its effects on managers' decision making (e.g.,Tetlock., 2007; Kurov, 2008), the study suggests that NPI sentiment is informative, rather than pure 'noise' and may help identify the factors affecting a manager's decision to issue positive guidance. Second, prior to the advent of social media, observing and robustly aggregating individuals' opinions about a particular firm was problematic for researchers. Social media has become a popular communication tool among investors in recent years. NPIs keep up with the latest news and trends in the finance world. They submit tweets about firm performance and foster discussion. Using social media data to infer NPI sentiment opens up future research to collect NPIs' opinions about individual firms in a natural environment. Third, in April 2019, the SEC issued an investor bulletin to express its concern regarding the ethical use of social sentiment-investing tools by firms; in particular, how these tools can be used to manipulate a stock's price.[1] Responding to the SEC's concern, this study shows that managers can exploit social media discussions to infer NPI sentiment and trigger NPIs to react to positive guidance in the way that managers hope for.

1.3 Crowdsourced Earnings Forecasts: Implications for Sell-Side Analysts' Earnings Forecasts Strategy

Regarding NPIs and analysts, I investigate whether analysts' walk-downs to beatable earnings forecasts are associated with investors' expectations of these firms' future earnings. A walk-down to beatable analyst earnings forecasts (hereafter 'analyst forecasts') occurs where analysts first issue optimistic earnings forecasts and then—at the official earnings announcement—'walk down' their estimates to a level that firms can beat (Jame et al., 2016, 2021). Analysts have self-interested incentives to opportunistically influence their forecasts to

[1] For further information, see *Investor bulletin: Social sentiment investing tools—think twice before trading based on social media*. https://www.sec.gov/oiea/investor-alerts-and-bulletins/ib_sentimentinvesting

help management beat them, including access to management inside information (Bradley et al., 2020; Feng & McVay, 2010; Ke & Yu, 2006) and future employment opportunities (Horton et al., 2017; Lourie, 2019). One factor that plausibly conditions analysts' walk-down to beatable earnings forecasts is investors' earnings expectations.

The analyst profession thrives because of its reputation for unbiased assessments of a firm's potential (Cote, 2000; Kadous et al., 2009; Meng, 2015). When investors consider easy-to-beat analyst forecasts as biased, the costs of worsening reputation may offset the intended benefits (Jame et al., 2016, 2021; Meng, 2015). When assessing bias in analyst forecasts, investors often anchor on their own earnings expectations, which causes a contrast effect (Schafhäutle & Veenman, 2021). That is, in periods of higher (lower) investor earnings expectations, investors perceive analysts' walk-downs as relatively credible (incredible) because they contrast (align) with their own existing beliefs about a firm's future performance, and this leads investors to give a low (high) rate for the analyst.

I investigate whether analysts exploit investors' earnings expectations toward a firm to walk down their forecasts, and whether this walk-down is likely to induce pessimistic forecast errors (i.e., the forecasted earnings are less than the actual earnings). Specifically, I test the association between changes in investors' earnings expectations and subsequent revisions to analyst forecast consensuses after controlling for firm characteristics likely to affect analyst forecast decisions. I infer investors' earnings expectations from crowdsourced earnings estimates on Estimize, which contains 879,015 'street earnings' estimates submitted by 70,926 participants in the period January 2012 to September 2018. Estimize is a crowdsourcing platform whose content focuses on financial estimates and that has been used in accounting research such as that by Jame et al. (2016) and Schafhäutle & Veenman (2021) to proxy for investors' earnings expectations. I first show that the level of revision in analyst forecast consensuses increases with the level of change in investors' earnings expectations. The revision of analyst forecast consensus is measured as the difference between an analyst forecast consensus issued [–30, –1] days before an actual earnings announcement for firm *j* in quarter *t*, and an analyst forecast consensus issued [–60, –31] days before the actual earnings announcement, deflated by the stock price at the end of the previous quarter. The change in investors' earnings expectations is measured as the difference between an Estimize estimates consensus issued in the [–60, –31] days before an actual earnings announcement for firm *j* in quarter *t* and an Estimize estimates consensus issued in the [–90, –61] days before the actual earnings announcement, deflated by the stock price at the end of the previous quarter. I regress the revision of analyst forecast consensus on the change of investors' earnings expectations using an OLS regression and the relevant controls. I then show that the likelihood of analysts' forecast errors being pessimistic is greater when analysts' prior forecast revisions follow downward revisions in investor earnings expectations. I estimate a logistic regression for the likelihood of analysts issuing forecasts that generate pessimistic errors for the analysts' downward revision (i.e., a binary that equals 1 when the revision to the analyst forecast consensus is less than 0; and 0 otherwise) following changes in investors' earnings expectations, with the relevant controls. The findings are consistent with analysts opportunistically manipulating earnings forecasts to exploit investors' expectations of future earnings.

This investigation contributes to the literature in three ways. First, recent research provides evidence that crowdsourced earnings forecasts can be useful for investors in predicting firms' future earnings (Adebambo & Bliss, 2015; Jame et al., 2016) and pricing earnings news (Schafhäutle & Veenman, 2021). Adding to this research, I find a positive association between changes in crowdsourced earnings estimates and analyst forecast revisions, suggesting that crowdsourced earnings forecasts also provide analysts with valuable information that affects their forecast decisions. Second, this study contributes to our understanding of the market forces that constrain analysts' conflicts of interest. My findings confirm the role of reputational considerations in disciplining analysts, and show that analysts are likely to opportunistically issue pessimistic forecasts following a downgrade in investors' earnings expectations to mitigate the risk of losing reputation. Third, on 30 August 2010, the SEC published an investor publication that describes analysts' conflicts of interest and encourages investors to uncover these conflicts. The arrival of Estimize provides a means for investors to assess the extent to which analysts opportunistically bias their forecasts. They can select which analyst forecast to rely on when faced with multiple forecasts from different analysts.

1.4 Structure of the Thesis

The remainder of this thesis is structured as follows. Chapter 2 provides a literature review on the rise of NPI investing and the development of online social networks. It also reviews the literature on the role of management guidance and analyst forecast in capital markets. Chapter 3 presents the first study, which investigates whether and how the likelihood of a manager issuing positive guidance is associated with the sentiment of NPIs toward the firm. The second study, presented in Chapter 4, examines how analysts' walk-downs to beatable earnings forecasts are associated with investors' expectations of a firm's future earnings. Chapter 5 discusses the findings of this thesis and provides suggestions for future research. Chapter 6 concludes the thesis.

Chapter 2 Literature Review

2.1 NPIs

In this thesis. I focus on NPIs in contrast to professional investors for the following reasons. First, NPIs and professional investors trade for different purposes. NPIs are natural persons trading in their own right who are not registered with the SEC or stock exchange and do not work for an institution that requires them to be so registered or qualified. On the other hand, a professional investor is a company or organization that invests money on behalf of clients or members. Hedge funds, mutual funds, and endowments are examples of professional investors. They are considered savvier than NPIs and are often subject to less regulatory oversight. Second, NPIs and professional investors have different investment horizons. NPIs, in general, invest in shorter-term trending stocks and spread their investment across different stocks or even in different sectors. Their short-term focus makes them sensitive to short-term earnings new, in our case, quarterly earnings guidance (Bushee 2004; Kim et al., 2017). In contrast, many professional investors are committed to investing in the long term. Because of the longer investment horizons, they are less focused on near-term earnings and do not trade actively for short-term profits. Even if these investors have near-term forecasts, it may not be optimal for them to trade on it (Ke and Petroni 2004). Finally, NPIs have lower skills and fewer resources and technology available to assess investments compared with professional investors. As such, NPIs often exhibit speculative trading, that is, they invest based on information they hear or believe (Han & Kumar 2013; Pan et al. 2016).

Understanding NPI participation in the stock market is important for several reasons. First, NPIs play an increasingly important role in capital markets. In 2020, these investors accounted for approximately 20% of trading activity in the United States (US).[2] The rise of NPI investment is the result of many factors: online trading platforms such as Robinhood and FreeTrade have made trading on the stock market more accessible for NPIs; online social network services such as StockTwits and Seeking Alpha provide NPIs with easy access to information about the stock market; and the global trend toward working from home during the Covid-19 pandemic gave NPIs more free time and motivated them to find alternative sources of income. Second, owing to NPIs' lack of knowledge, they are less capable than are professional investors (or institutional investors)[3] of accessing and identifying useful information, and may rely more on signals sent by managers and analysts to make investment decisions (Cahill et al., 2017; Schmeling, 2009). Research shows that NPIs have difficulty when processing information about the stock market (e.g., fair value information on liabilities and SOX 404 auditor reports) (Boritz et al., 2020; Lachmann et al., 2011) and may rely more on signals sent by managers and analysts to make investment decisions (Cahill et al., 2017; Schmeling, 2009). Third, rational investor decision making should be driven by economic fundamentals, free of the influence of behavioral biases and emotion (Daniel & Hirshleifer, 2015; Hirshleifer, 2001). However, research shows that NPIs are likely to be influenced by individual factors, such as stereotypes (Brave & Nass, 2007; Frijda, 1994), emotions (Reavis, 2012; Shefrin, 2002), and traits (Munezero et al., 2014) when making investment decisions. This leads to their irrational market behavior inconsistent with the efficient-market hypothesis, and consequently, unpredictability of the stock market as a whole (Daniel & Hirshleifer, 2015; Hirshleifer, 2001). The argument aligns with empirical studies in behavioural finance, which suggest that NPI's s irrational behaviour fuels overpricing in financial markets and is

[2] Please see *Professional investors should not ignore the retail wave*. https://www.ft.com/content/ddc4630c-c27c-47e6-b13e-1e036d16b0f9 and *Individual investor boom reshapes U.S. stock market*. https://www.wsj.com/articles/individual-investor-boom-reshapes-u-s-stock-market-11598866200

[3] An institutional investor is a company or organization that invests money on behalf of clients or members. Hedge funds, mutual funds, and endowments are examples of institutional investors. Institutional investors are considered savvier than the average investor and are often subject to less regulatory oversight.

particularly salient when NPIs are greedy (Janssen et al., 2019; Reavis, 2012). Greed provokes NPIs to concentrate on potential gains and create unrealistic optimism. Their unrealistic optimism overrides objective assessments of the intrinsic value of the firm's fundamentals in their investment decision, resulting in stock overpricing (Barberis et al., 2018; Reavis, 2012).

A classic example of NPI participation that causes unpredictability of the stock market is the GameStop trading frenzy.[4] GameStop, a struggling video game retailer that operates over 5,000 stores in the US, suddenly saw NPIs buying up its stock in droves in January 2021, while many professional investors were shorting the company's shares. Professional investors estimated that GameStop's business would decline in the Covid-19 pandemic, and that its share price was bound to fall. They sold GameStop's shares, intending to buy them back later at lower prices. However, some NPIs took a different view and spread the message to buy GameStop shares on online social networks, which led to GameStop's share price rising by almost 1,900% in less than one month, as many NPIs invested in them (mostly via online trading platforms). Those professional investors who shorted GameStop stock, including hedge funds such as Melvin Capital Management, lost almost US$6 billion. As of 27 August 2021, GameStop's share price was US$204.95, below the US$347.51 peak in January 2021 but still much higher than the US$17.25 mark at the end of 2020. The GameStop trading frenzy shows how a group of NPIs banding together can dramatically push up a stock in the short term. The GameStop example has evident the power of NPIs in nowadays capital market, however, it is hard to conclude GameStop case is an example of a strategic win of NPIs. To elaborate, NPIs rely on posting on Reddit when deciding to invest in GameStop, but this does not help them identify traders with superior skills. Indeed, large increases in Robinhood users are often accompanied

[4] Please see *GameStop? Reddit? Explaining what's happening in the stock market.* https://www.nbcnews.com/business/business-news/gamestop-reddit-explainer-what-s-happening-stock-market-n1255922 and *How retail investors are changing the stock market.* https://www.fairobserver.com/economics/finance/heya-shah-retail-investors-stock-market-social-media-gamestop-robinhood-news-25541/.

by large price spikes and are followed by reliably negative returns (Barber et al.,2021). Also, NPIs on Reddit are driven by their sentiment, typically excited about firms that do end up exceeding expectations; their enthusiasm was excessive and resulted in negative post-announcement returns (Hu et al., 2021). In addition, it should be noted that there are some fraudsters in the GameStop case. GameStop trader Keith Gill, a chartered financial analyst with multiple broker licenses, impersonates amateur "Robinhood" and induces other NPIs to buy GameStop stock[5]".

2.2 Online Social Networks

2.2.1 Social Media Platform

Social media platforms such as Twitter and Seeking Alpha enable individuals who have no direct line to management to ask questions publicly and interact in ways that motivate managers to take action (Cade, 2018; Hales et al., 2018). On social media platforms, NPIs can keep abreast of breaking news and industry developments. If they can purchase or offload stocks before the rest of the market becomes aware of a significant change, they put themselves at a considerable advantage when it comes to making a profit. NPIs can also use the platform to connect with others in any industry to obtain further insight into their market. According to a survey by Consumer News and Business Channel (CNBC, 2021), about 35 percent of NPIs said that they use social media to look into a possible investment, compared with 25% cited conversations with family and friends, and 24% said financial guidance or investment websites.

Recent research suggests that messages posted on social media provide information relevant for predicting investor perception for a firm and their sentiment in the stock market. For example, Cade (2018) examines cases where Twitter users criticize firms' discretionary accrual

---

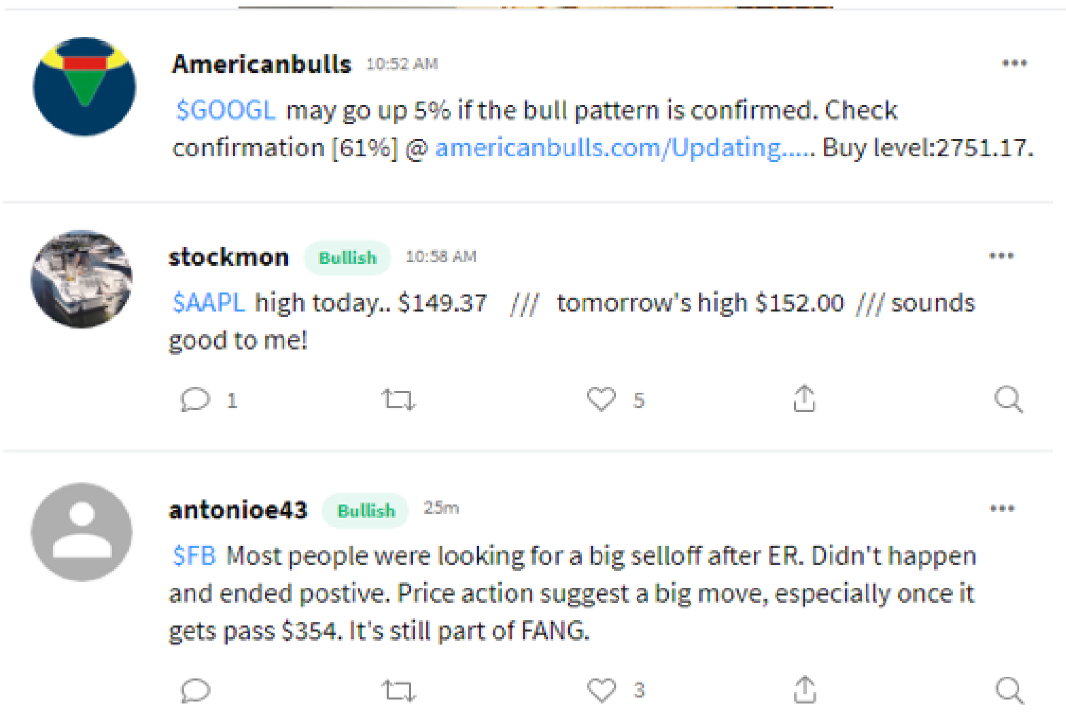
[5] For further information, please see "GameStop trader Deep F—ing Value, a licensed broker, sued for "fake persona" https://www.fintechfutures.com/2021/02/gamestop-trader-deep-f-ing-value-a-licensed-broker-sued-for-fake-persona/

adjustment and NPIs respond to the criticism. She finds that NPIs evaluate firms' future financial performance less favorably in response to criticism, and this response is more pronounced for criticisms with more retweets. Extending Cade (2018), subsequent studies associate sentiment derived from social media to stock market movements. These studies include that of Behrendt and Schmidt (2018), who discover some significant co-movements of intraday return volatility and Twitter sentiment for all constituents of the Dow Jones Industrial Average; Campbell et al. (2019), who find that sentiment derived from financial articles and commentaries on Seeking Alpha predicts future stock returns; and Dunham and Garcia (2020), who find that high (low) Twitter sentiment leads to a decrease (an increase) in the average firm's share liquidity.

As social media platforms make individual public opinions about firms more easily accessible and aggregated than before, some researchers examine whether these platforms provide information relevant to managers in disclosing information to investors. In particular, Chen et al. (2014) investigate whether sentiment derived from financial articles and commentaries on Seeking Alpha can predict earnings surprise (i.e., the difference between the reported earnings per share [EPS] and the average of analysts' EPS forecasts issued/updated within 30 days prior to the earnings announcement). Bartov et al. (2018) and Tang (2018) link Twitter sentiment prior to a firm's earnings announcement to its upcoming quarterly earnings and sales growth. While these studies employ social media platforms such as Twitter and Seeking Alpha to measure investor sentiment, their results need to be interpreted with caution. According to Oliveira et al. (2013) and Hales et al. (2018), Twitter and Seeking Alpha users include the general public who may not have any interest or knowledge in investment; hence, their posts are not specific to the stock market and contain noise without any real investor sentiment behind them. Rather than focusing on general posts on Twitter and Seeking Alpha, Hales et al. (2018) examine a social media platform, Glassdoor.com, where employees voluntarily share their insider opinions about their firms' prospects. The results reveal that employee sentiment is related to future growth in key income statement information, transitory reporting items (e.g., restructuring charges), earnings surprises, and management guidance news.

More recent studies extract data from StockTwits, a social media platform designed explicitly for investors and traders to overcome noise when seeking investor sentiment, which is present in more generalist social media such as Twitter and Seeking Alpha (Oliveira et al., 2013). StockTwits is an investor-oriented platform with more than one million users who share information related to the market and individual stocks (Deng et al., 2018). Similarly to Twitter, because of the character limit on StockTwits posts (i.e., limited to 140 characters), tweets are relevant and succinct, and relatively devoid of the typical noise included in many other social media platforms (e.g., Seeking Alpha) (Renault, 2017). StockTwits users share posts about an individual stock or index by adding a '$' symbol before the ticker symbol; for example, $AAPL (Apple Inc.), $FB (Facebook, Inc.), and $GOOGL (Google LLC) (Oliveira et al., 2013). The $TICKER tag (hashtag) allows StockTwits to filter streams of information about a particular stock or index (Oliveira et al., 2013). Figure 2 shows a real-time screenshot from the StockTwits platform with a tweet for $AAPL, $FB, and $GOOGL. In addition, StockTwits began offering a service - StockTwits IR suite in 2011 that allows managers to monitor can monitor in real-time any messages about their firm and also the level of discussion on StockTwits about their stock.

Figure 2 A screenshot from the StockTwits platform



Investor sentiment obtained from StockTwits has been used to examine stock market behavior. Renault (2017) and Deng et al. (2018) use StockTwits tweets to test the relationship between investor sentiment and intraday stock returns, providing evidence that investor sentiment helps predict intraday stock index returns. Audrino et al. (2020) use StockTwits tweets to analyze the relationship between investor sentiment and stock market volatility, and find that investor sentiment has significant predictive power for realized volatility.

2.2.2 Crowdsourcing Platform

Crowdsourcing uses an internet task market, usually called a crowdsourcing platform, to engage a crowd or group in collaborative brainstorming (Hamadi et al., 2020). This collaborative brainstorming model enables a large and open group of people to work together on a single task and achieve cumulative results that conventional means could never manage. Crowdsourcing platforms connect people to various tasks; for example, participating in experiments (Farrell et al., 2017), data coding (Stol et al., 2017), and project funding (Coakley & Lazos, 2021).

Crowdsourcing platforms such as Amazon Mechanical Turk (MTurk) have been introduced to accounting and finance research in recent years, and these platforms have been mainly used for conducting surveys or experiments. An early study by Brandon et al. (2014) compares the online participant recruitment services provided by MTurk to traditional participant recruitment processes in accounting research and concludes that MTurk provides researchers with relatively inexpensive access to participants and flexibility to incentivize participants appropriately. Subsequent studies investigate the quality of participants recruited using crowdsourcing platforms. These studies include that of Farrell et al. (2017), who compare the decision making of MTurk participants with that of student research assistants in an accounting setting, and find that MTurk participants have honesty preferences and exert effort similar to student research assistants. After demonstrating the validity of crowdsourcing accounting research, researchers use crowdsourcing platforms to collect data through surveys and conducting experiments. For example, Tadesse and Murthy (2018) recruit NPIs on MTurk to examine their perceptions of the partial remediation of information technology (IT) control weaknesses and non-IT control weaknesses. Stuart et al. (2021) employ NPIs on MTurk to evaluate whether they judge firms differently based on managers' stated purpose for undertaking corporate social responsibility activities in the presence versus absence of a firm-specific negative event.

Recent studies take advantage of Estimize data, a crowdsourcing platform for financial estimates, to study investor expectations about a firm's prospects. Launched in 2011, Estimize is designed to collect forward-looking financial estimates from independent, buy-side and sell-side analysts, and portfolio managers, along with those from NPIs and academics. Figure 3 presents a screenshot from the Estimize interface. Currently, Estimize has over 100,000 contributors, resulting in coverage of over 2,200 stocks and 80 economic indicators each quarter.[6] Estimize provides its daily proprietary 'consensus' earnings estimates for the next earnings announcement for each stock in its database. Estimize earnings consensus has been used as a proxy for investors' earnings expectations to examine how investors react to earnings that miss their earnings expectations (Veenman & Verwijmeren, 2018) and how their earnings expectations influence analyst forecast bias (Jame et al., 2021; Schafhäutle & Veenman, 2021).

Figure 3 A screenshot from the Estimize interface

GOOGL Alphabet Inc.

EPS	FQ4 '20	FQ1 '21	FQ2 '21	FQ3 '21	FQ4 '21	FQ1 '22	FQ2 '22	FQ3 '22
You	-	-	-	-				
Wall St.	15.77	15.76	19.24	23.73	26.91	24.58	26.42	28.68
Estimize	16.82	17.82	22.34	25.65	26.81	27.13	28.08	-
Actual	22.30	26.29	27.26	27.99	-	-	-	-
YoY Growth	45%	166%	169%	71%	-	-	-	-

AAPL Apple Inc.

EPS	FQ4 '20	FQ1 '21	FQ2 '21	FQ3 '21	FQ4 '21	FQ1 '22	FQ2 '22	FQ3 '22
You	-	-	-	-				
Wall St.	0.71	1.42	0.99	1.01	1.24	1.85	1.29	1.16
Estimize	0.76	1.47	1.07	1.16	1.35	1.95	1.38	1.26
Actual	0.73	1.68	1.40	1.30	-	-	-	-
YoY Growth	-4%	35%	120%	102%	-	-	-	-

FB Facebook

EPS	FQ4 '20	FQ1 '21	FQ2 '21	FQ3 '21	FQ4 '21	FQ1 '22	FQ2 '22	FQ3 '22
You	-	-	-	-				
Wall St.	3.19	2.35	3.04	3.19	3.87	3.10	3.39	3.38
Estimize	3.31	2.63	3.35	3.39	4.11	3.81	3.88	3.63
Actual	3.88	3.30	3.61	3.22	-	-	-	-
YoY Growth	52%	93%	101%	19%	-	-	-	-

There are several reasons why contributors have the incentive to release their earnings estimates on Estimize. For instance, all contributors can benefit from a comprehensive earnings expectation dataset and seek to build one by sharing information with their peers (Da & Huang,

[6] Please see https://www.estimize.com/faq#what

2020). Among its contributors, analysts may use Estimize to build a record of their accuracy and foresight for the upcoming earnings of the firms they cover (Jame et al., 2016). This record can distinguish their ability to predict future earnings from that of their peers and help them attract clients. Some portfolio managers may contribute their estimates because they want to ensure that their information is reflected more quickly in prices (Crawford et al., 2018), whereas others may contribute in an attempt to manipulate prices. NPIs may participate because they want to develop their forecasting skills. Estimize allows them to compare their estimates with those of peers and receive feedback regarding their estimation accuracy (Jame et al., 2016).

Estimize takes steps to incentivize the accuracy of estimates and protect against bad estimates that may be contributed for both innocent and nefarious reasons. First, Estimize assigns users virtual tokens for each estimate they make. The point system rewards users who offer more accurate estimates than the Wall Street consensus, while penalizing them for less accurate estimates. Contributors who gain the most points are recognized on the website, featured in podcasts, and awarded prizes.[7] Second, when a new user joins Estimize, their first five estimates are reviewed manually for reliability, and their subsequent estimates continue to be algorithmically reviewed for reliability.[8] Estimates whose reliability is believed to be low are flagged throughout the platform and are not included in the Estimize consensus. Further, in 2015, together with Da and Huang (2020), Estimize experimented with investigating the question—by allowing contributors to view each other's estimates before contributing their own (or simply viewing the data)—of whether contributors were herding together more than they otherwise would, and how this was impacting the accuracy and representativeness of the dataset as a whole. They randomly assigned contributors to a blind group in which contributors

[7] Please see https://www.estimize.com/faq#scores
[8] Please see https://www.estimize.com/faq#consensus

were asked to provide their own estimate before seeing any others and a control group in which contributors were allowed to view others before making their own. They find that the Estimize earnings consensus among the blind group on average is more accurate than that for the control group, implying that the Estimize earnings consensus becomes more accurate with a more significant number of independent opinions. On this basis, Estimize switched to a blind platform in November 2015.

2.3 The Role of Management Earnings Guidance in Capital Markets

Earnings guidance is the discretionary disclosures by management that provide its own prediction of future earnings. Early studies on earnings guidance investigate whether earnings guidance conveys information about firm value to investors. For example, Foster (1973) uses the trading volume reaction and the adjustment of stock prices to earnings guidance to infer that investors consider earnings guidance to have information content. Subsequent studies (Nichols & Tsay, 1979; Patell, 1976; Penman, 1980) extend those results by examining whether the information contained in earnings guidance is associated with the direction of the guidance news. These studies distinguish between positive and negative earnings guidance by comparing the guidance estimates to a mechanical model of investors' earnings expectations based on prior actual earnings. They find evidence of excess stock returns to positive guidance. Waymire (1984) uses analyst forecasts to measure guidance news, rather than a model of investors' earnings. He compares the guidance estimates to analyst forecasted earnings and finds significant abnormal stock returns following both positive and negative earnings guidance in their respective news direction. Since the enactment of Regulation Fair Disclosure (Reg FD) in 2000,[9] managers increasingly 'bundle' guidance with earnings announcements (Anilowski

[9] Reg FD is a rule passed by the SEC that is intended to prohibit private communications between public firms, analysts, and certain public shareholders. Public firms are allowed to conduct earnings and forecasts call to inform analysts and shareholders that are matched with simultaneously issued press release of the statements made during those calls (Kross et al., 2011).

et al., 2007; Billings et al., 2015; Rogers & Van Buskirk, 2013). In particular, the proportion of earnings announcements with bundled guidance rose to 45% by 2004, and more than 55% around 2011–13 and has remained at just over 50% in recent years (Billings et al., 2015). More impressively, the proportion of guidance bundled with earnings has increased steadily since the early 2000s and is now over 80% (Beaver et al., 2020). Zhang (2012) focuses on post-earnings-announcement drift when earnings guidance about the next quarter's earnings are bundled with the current quarter's earnings announcements, to investigate whether bundled earnings provide information to investors that resolve the uncertainty related to future earnings. She finds that bundled earnings guidance mitigates post-earnings-announcement drift when the guidance has higher accuracy.

Earnings guidance is informative for investors in forming their earnings expectation, however, there are increasing concerns about the potential agency problems associated with guidance, due to its voluntary and non-audited nature (Core, 2001; Healy & Palepu, 2001). To elaborate, the practice of guidance has become "misguided" due to an excessive focus on short-term "number games" rather than long-term business conditions. This myopic focus potentially induces managers to sugarcoat poor business conditions with misleading guidance, which leads to further opportunistic behaviors (e.g., earnings management to meet or beat the guidance) (Graham et al., 2005; Kothari et al., 2009). In response to these concerns, researchers studied management incentives and attempts to discern managers' goals using forecast characteristics. One early study is that of Ajinkya and Gift (1984), who argue that managers have the incentive to issue earnings guidance to move market expectations toward their beliefs about future earnings. They use analyst forecasts as surrogates for market earnings expectations, and find results consistent with their argument. Their results also reveal that managers issue both positive and negative earnings guidance to adjust market earnings expectations, rather than selectively disclosing only when they have good news regarding earnings. Contrary to the findings of Ajinkya and Gift (1984), Kothari et al. (2009) find the magnitude of the negative stock price reaction to negative guidance to be greater than the magnitude of the positive stock price reaction to positive guidance, implying that, on average, managers delay the release of bad news regarding earnings while immediately disclosing good news to investors.

Other studies explore the rationale behind managers' selective disclosure of positive and negative guidance. These studies include Frankel et al. (1995), who examines the effect of litigation risks and market price on managers' disclosure of positive and negative guidance prior to equity financing. They find that the additional litigation risk shortly before an equity financing decreases (increases) managers' incentive to provide positive (negative) guidance. Cheng and Lo (2006) examine whether managers take advantage of voluntary disclosure to profit from insider trading. They show that managers tend to disclose negative guidance to decrease the purchase share before their share purchase. Moreover, Kross et al. (2011) explore whether consistency in meeting or beating analysts' earnings forecasts affects managers' guidance decisions. They document that when their firms have a consistent record of meeting or beating analysts' earnings forecasts, managers are more likely to provide negative guidance to guide analysts' earnings expectations down, to maintain their consistency.

Another stream of research focused on whether investor responses to earnings guidance vary with factors that reflect the manager's credibility. An early example is Jennings (1987), who uses the magnitude of analyst forecast revisions subsequent to a manager's earnings guidance to measure the credibility of the manager, finding that stock returns around the release of earnings guidance are more pronounced for the more credible managers. This association is more significant in cases of positive guidance relative to negative earnings. Consistent with Jennings (1987), subsequent studies (Ng et al., 2013; Rogers & Stocken, 2005) show that stock returns around the release of earnings guidance are larger for more credible managers, with credibility proxies by managers' prior guidance accuracy, litigation risk, and the level of competition their firms face, which affects how they bias their guidance. These studies also find that investors often interpret negative earnings guidance as more credible than positive guidance, even though they do not find negative earnings guidance unbiased or even less biased than positive guidance.

2.4 The Role of Analyst Earnings Forecasts in Capital Markets

An earnings forecast is an analyst's prediction for a firm's future earnings. Researchers began examining analyst forecasts to investigate their usefulness as a surrogate for time series forecasts in studies of the efficiency of capital markets. For example, Givoly and Lakonishok (1979) find a significant stock price reaction to disclosures of analyst forecast revisions, implying analyst forecasts have information content that might proxy for investor expectation. Fried and Givoly (1982) show that analyst forecast errors are more closely related to stock price movement than time series forecasts, implying that analyst forecasts provide a better surrogate for investor expectations than do forecasts generated by time series models. Later studies such as those of Brown et al. (1987) and Eddy and Seifert (1992) explore the source of analyst forecast superiority as a proxy for investor expectation, and find that analyst forecast superiority is positively related to the dimensionality of the information set (proxied by firm size), and negatively related to variance in the information observations (proxied by past variability of earnings). Building on studies that compare analyst forecasts with time series forecasts, researchers examine factors correlated with analyst forecast accuracy. These studies show that analyst forecast accuracy increases with individual analysts' ability (Stickel, 1990) and experience (Clement, 1999), and decreases with portfolio complexity (Clement, 1999) and forecast horizon (Brown & Kim, 1991; O'Brien, 1988).

Since then, interest in analyst forecasts has shifted to explore whether they inform investors in forming their earnings expectations. Some studies described above also examine investor reactions to analyst forecasts. For example, Givoly and Lakonishok (1980) and Stickel (1991) show that stock prices initially underreact to forecast revisions, resulting in short-term return drift. The initial underreaction is often viewed as resulting from information uncertainty that inhibits the efficiency with which prices reflect available information, and investors' information processing biases regarding specific attributes of analysts' forecasts. Specifically, Zhang (2006) investigates how information uncertainty (proxied by firm size, age, analyst coverage, dispersion in analysts' forecasts, return volatility, and cash flow volatility) affects short-term return drift. They find that short-term return drift decreases with information uncertainty, suggesting that lower information uncertainty enables investors to react more strongly to analysts' forecast revisions. Gleason and Lee (2003) investigate how analyst reputation affects short-term return drift. They find that short-term return drift is lower for analyst forecast revision disclosed by star analysts than that by less well-known analysts from smaller brokerage houses, suggesting a higher analyst reputation enables investors to react more strongly to analyst forecast revisions.

Another stream of studies regards analysts as an interesting economic agent in their own right, much like literature focusing on managers. Analysts typically perform as intermediaries in regard to information in the capital market; however, their role is characterized by a vast number of principal agency relations and conflicts of interest. That is, investors, especially NPIs, rely on analysts' forecasts to make informed investment decisions. Accordingly, investors expect that analysts provide independent, unbiased, and accurate earnings forecasts to the best of their knowledge. However, managers of firms followed by analysts are not always interested in the objectivity and accuracy of analyst forecasts; rather, the outcome of analyst forecasts as reflected in stock price movement. Moreover, analysts often act in their own interest and strive to maximize their own profits rather than those of investors or their employers (i.e., an investment bank or a broker). This conflict of interest is more pronounced when an analyst's employer is an investment bank or a broker that also offers corporate finance services (Feng & McVay, 2010; Karamanou, 2011). These affiliated analysts have a strong incentive to curry favor with management to generate additional business for the investment banking division. Studies identify empirical evidence for this 'conflict-of-interest-hypothesis.' Affiliated analysts wishing to please management tend to publish optimistic long-term earnings forecasts to secure investment banking mandates (Karamanou, 2011), but are pessimistic in their short-term earnings forecasts to allow management to overcome earnings expectations when earnings are realized (Feng & McVay, 2010).

2.5 Chapter Summary

This chapter describes NPIs and their role in capital markets; provides an overview of internet platforms in accounting studies; and reviews the literature on management guidance and analyst forecasts. This thesis defines NPIs as natural persons trading in their own right, not registered with the SEC or stock exchange, and who do not work for an institution that requires them to be so registered or qualified (NYSE, 2016). In the first study of my thesis, I obtain NPI sentiment from a social media platform—StockTwits—and examine its impact on managers' guidance decisions. In the second study, I obtain investors' earnings expectations (professional investors and NPIs) from a crowdsourcing platform—Estimize—and examine their impact on analysts' forecast decisions. The first study, "Non-professional investor sentiment and management earnings guidance", is presented in the next chapter.

Chapter 3 NPI Sentiment and Management Earnings Guidance

3.1 Background and Research Problem

Earnings guidance is a public statement of a manager's forecast of the level of future earnings (Rogers & Van Buskirk, 2013) and provides an important signal to the market about firm value (Milian, 2018). However, managers have well-known incentives to opportunistically influence earnings guidance in pursuit of private benefits. For example, issuing positive ('good news') earnings guidance[10] may temporarily boost stock price (D'Augusta, 2022), increase stock liquidity (Balakrishnan et al., 2014) and reduce stock price volatility (Billings et al., 2015). These market effects, although temporary, may benefit both the firm (especially during capital raisings) and managers personally (by increasing the worth of their equity incentives). However, because these motives are apparent to many investors, the intended market reactions may not eventuate—unless investors become overconfident (a form of psychological bias) and care less about the precision and objectivity of public information signals (Barberis et al., 1998; Daniel et al., 1998). I examine one factor that plausibly influences the extent to which investors will anticipate and impound managers' private incentives when interpreting guidance: the sentiment of NPIs, which has become far more visible to managers since the advent of social media platforms specialising in discussions regarding listed securities (Cookson et al., 2021; Deng et al., 2018). I argue and show that managers are more likely to issue positive guidance when NPIs' sentiment towards the firms is high and that this is consistent with the pursuit of a stronger market reaction by these investors.

Prior studies that examine the association between investor sentiment and managers' guidance decisions (Hurwitz, 2018; Seybert & Yang, 2012) suggest that such an association may exist

[10] 'Positive' ('negative') guidance occurs where a management's forecast level of earnings is higher (lower) than the prevailing consensus analyst forecast for the same financial period (Billings et al., 2015; Twedt, 2016).

simply because managerial sentiment may be correlated with that of investors, or alternately may reflect informative or opportunistic reasons. The 'informative view' suggests that managers issue earnings guidance to correct or avoid under-valuation of their firm associated with investor sentiment (Seybert & Yang, 2012). The 'opportunistic view' posits that managers recognise the prevailing investor sentiment towards their firm and form their guidance decisions in pursuit of private benefits (Hurwitz, 2018).

While the opportunistic view has found little empirical support in the very limited prior literature, these earlier studies conceptualise and measure sentiment across all potential investors in a stock (e.g., Brown et al., 2012; Hurwitz, 2018; Seybert & Yang, 2012). However, it seems reasonable to assume considerable cross-sectional variation in both sentiment and willingness to act on the beliefs entailed. In particular, NPIs are less likely to have knowledge relating to fundamental analysis (Aboody et al., 2018; Weißofner & Wessels, 2020), and accordingly, their sentiment tends to reflect factors such as their memories of a firm's historical performance, their psychological acceptance of risk (Hirshleifer, 2001) and their perceptions of a firm's social reputation (Cordeiro & Tewari, 2015) and CEO social ties (Kaplan et al., 2015). Because it is practically easier to trade on positive sentiment than negative sentiment (i.e. taking a short position is typically more difficult and costly than purchasing stock), the impact of sentiment on NPI's willingness to engage in speculative trades should be stronger in cases of positive sentiment, and there is evidence consistent with this contention (Baker & Wurgler, 2006, 2007; Xiong et al., 2020). These factors make NPIs' investment decisions more vulnerable to managers' opportunistic guidance decisions during times of high sentiment. Consequently, the analysis focuses on NPIs, and the extent to which their investment decision-making is associated with opportunistic guidance behaviour.

I focus on the issuance of positive guidance because the opportunistic abuse of this decision is more likely to align with managers' private incentives than is negative ('bad news') or neutral guidance. While stock prices typically react in the same direction as the guidance provided (Han & Tan, 2010; Twedt, 2016), the magnitude of the reaction is asymmetric (Han & Tan, 2010; Twedt, 2016). Ng et al. (2013) show that market reaction to positive guidance is only one-third of that following negative guidance. The literature attributes this asymmetry to investor awareness of managers' self-interested motivations (e.g., incentives to please stockholders arising from equity compensation plans and insider trading), which most commonly encourage 'good news' disclosures, and thus positive guidance is perceived to be less credible (Asay et al., 2018; D'Augusta, 2022). In contrast, the issuance of negative guidance typically conflicts with managers' personal incentives (Asay et al., 2018; D'Augusta, 2022) and thus suffers less of a credibility threat.

Given the generally lower perceived credibility of positive guidance, the timing of its issuance may be particularly important for managers with incentives to make such disclosures. When NPI sentiment is high, NPIs may give greater credence to news that signals improving financial prospects (Baker & Wurgler, 2006, p.1648). In these circumstances, NPIs have confidence in the firm's future performance that cannot be explained by fundamentals, making them less sensitive to the opposing information (Barberis et al., 1998; Daniel et al., 1998). Because NPIs are less likely than others to possess the requisite knowledge and expertise to robustly appraise the investability of a firm's stock, their sentiment may color their interpretations of corporate information (Brown et al., 2012; Darrough et al., 2020; Simpson, 2013). Thus, these investors are likely to perceive positive guidance to be credible when investor sentiment is high. Taken together, when positive guidance is issued during periods of high NPI sentiment, it is likely to trigger a strong market response from these investors. To the extent that managers can benefit from increases in stock price (e.g. through the value of option-based compensation) or just a broadening of the firm's investor base, there exists scope for the opportunistic issuance of positive guidance in pursuit of managerial interests.

I infer NPI sentiment towards a firm from individuals' discussions about the firm on the StockTwits. StockTwits focuses on financial discussions between NPIs[11] and has been used in prior research to examine intraday stock returns (Deng et al., 2018; Renault, 2017) and stock market volatility (Audrino et al., 2020). I extract over 17 million tweets posted by 118,685 users about 3,212 distinct firms. Higher NPI sentiment implies that NPIs collectively hold a more optimistic attitude towards a firm. While the empirical measure is based on activity on a single social media platform, I expect the tone of conversation on that platform to be indicative of that on social media platforms more generally.

The main tests focus on guidance issued contemporaneously with current period earnings announcements (i.e., 'bundled guidance') because this is by far the most common trigger for the issuance of guidance and is typically part of a regular disclosure strategy.[12] This enables us to focus on the direction, rather than incidence, of guidance for future earnings provided. I first show that NPI reaction to positive guidance is conditioned by NPI sentiment. To this end, I follow the literature and use stock returns across the overnight period immediately following the issuance of guidance to measure NPI response (Aboody et al., 2018; Weißofner & Wessels, 2020) and regress this overnight return on an indicator of positive guidance, NPI sentiment and their interaction. The resulting coefficient for this interaction term is positive and significant, suggesting that NPI reaction to positive guidance increases in line with NPI sentiment. I next test the association between NPI sentiment and the likelihood that guidance issued is positive.

[11] I obtained the full dataset from StockTwits since its launch. Each StockTwits user is required to register an account. In registration, they self-report their investment expertise. In our dataset, 4.08 percent of StockTwits users claimed themselves as professional investors. Therefore, most discussions on StockTwits come from NPIs.
[12] During the sample period, 63.1 percent of guidance is bundled with current period earnings announcements and 36.9 percent of guidance is standalone (i.e., guidance issued separately from earnings announcements).

In all these tests, I control for potential confounding influences from broader sentiment measures through the inclusion in my models of measures of firm-level news sentiment (Cade, 2018; Ng et al., 2016), which typically reflects the opinions of professional analysts, business journalists (Giannini et al., 2019) and institutional order flow (Hendershott et al., 2015), and market-wide sentiment (proxied by the Michigan consumer sentiment index [MCSI]). After controlling for these factors and numerous other firm and market attributes, I find consistent evidence of a significant positive association between NPI sentiment and the likelihood that guidance issued is positive. Taken together, the findings are consistent with managers opportunistically issuing positive guidance following higher NPI sentiment periods in pursuit of a stronger NPI reaction.

I conduct a series of tests to address potential endogeneity threats. First, I exploit an exogenous shock to the public awareness of fraudulent stock rumours triggered by the SEC's fraud charges announced in April 2014 to examine whether greater managerial scepticism regarding the tone of social media discussions may moderate the effects demonstrated in the main tests.[13] While I find a positive and significant effect of NPI sentiment on the likelihood of issuing positive guidance in both pre- and post-alert periods, this effect is significantly weaker after the alert. The results are also robust to the use of a two-stage least squared (2SLS) approach to alleviate concerns over reverse causation, and an unobservable selection analysis to mitigate omitted variables concerns. The results are also robust to numerous alternative modelling choices: (i) the use of a Heckman-type correction for selection bias regarding the decision to issue any guidance (Cheng et al., 2013; Fang & Peress, 2009), (ii) the inclusion of additional controls for macroeconomic conditions as per Hribar et al. (2017), (iii) the use of a residual measure of NPI sentiment that parses out effects of investor attention and (iv) the use of alternative sentiment

[13] For further information, see "*SEC v. JCS Enterprises, Inc. et al.*, Civil Action No. 14-civ-80468 (S.D. Fla.) (April 7, 2014)" https://www.sec.gov/litigation/litreleases/2014/lr22969.htm and "*In the Matter of Keiko Kawamura*" https://www.sec.gov/litigation/admin/2014/33-9574.pdf.

lexicons and alternative time-windows to measure NPI sentiment. I also show that the impact of NPI sentiment is stronger in firms with lower institutional ownership concentration (an inverse proxy of the size of the NPI base).

I perform a series of additional tests to shed more light on competing explanations for the main findings. First, consistent with the opportunistic, rather than informative, perspective on guidance behaviour, I find that managers holding a greater value of exercisable stock options and those with more frequent insider trading activities are abnormally likely to issue positive guidance during high NPI sentiment. I address the possibility that biases in earnings guidance reflect managerial sentiment in two ways. First, the main tests include controls for broader sentiment measures, comprising market-level and firm-level total investor sentiment, which appear far more likely to influence managers' sentiment than would NPI sentiment. To obtain stronger identification, the additional tests show that when firm-level NPI sentiment is in contrast to market-level sentiment, the effect of NPI sentiment on guidance behaviour remains significant. Collectively, these tests provide further support for the contention that managerial opportunism drives the main results.

This study contributes to the literature in three ways. First, I focus on an important group of investors—NPIs—who have been often neglected in the earnings guidance literature but whose wealth is potentially the most vulnerable opportunistic behaviour. I present evidence on the conditioning role of NPI sentiment on NPI reaction to positive guidance and that managers are likely to opportunistically issue positive guidance following periods of high NPI sentiment. Interestingly, while much of the literature regards NPI sentiment as 'noise' and dedicates little attention to its effects on managers' decision-making, the findings suggest that NPI sentiment contains information to managers and may affect their guidance strategy.

Second, this study is among the first to use social media data to measure NPI sentiment at the firm level. Prior to the advent of social media platforms, observing and robustly aggregating individuals' opinions about a particular firm was problematic for researchers. Aboody et al. (2018) acknowledge that proxies for market-level sentiment 'although varying over time, are invariant in the cross-section. This makes market-level indices not well suited to investigate firm events. Although it would be preferable to use a firm-specific measure of investor sentiment, such a measure has not generally been available' (p. 486). Social media platforms have become a popular communication tool among investors in recent years. Examples include StockTwits, Yahoo! Finance and Seeking Alpha. NPIs keep up with the latest news and trends in the finance world. They submit tweets about firm performance and foster discussions. Using social media data to infer NPI sentiment opens up future research to collect NPIs' opinions of individual firms in a natural environment.

Third, in a Report of Investigation under Section 21(a) published on 2 April 2013, the SEC permitted social media platforms to be a legitimate platform for public companies to disseminate material information without running afoul of Regulation FD (SEC, 2013). In April 2019, the SEC issued an investor bulletin to further express their concerns regarding the ethical use of social sentiment investing tools by firms, especially how they can manipulate stock prices.[14] The findings suggest that managers may have exploited social media discussions to infer NPI sentiment for amplified NPI reactions to positive guidance to benefit themselves. It is possible that other resourceful market participants, such as IPO underwriters and analysts associated with them, are also harnessing sentiment analysis on social media data to influence investor decision-making. Future research in these areas is encouraged. Taken together, social

[14] For further information, see "Investor Bulletin: Social Sentiment Investing Tools - Think Twice Before Trading Based on Social Media" https://www.sec.gov/oiea/investor-alerts-and-bulletins/ib_sentimentinvesting.

media data bring new challenges to securities regulators where existing regulations may not be sufficient. This calls for a new consideration of best practices in this regard.

The chapter is organised as follows: Chapter 3.2 reviews the related literature, and Chapter 3.3 develops the hypothesis. Chapter 3.4 discusses the sample selection and research design. Chapter 3.5 presents the results of hypothesis testing. Chapter 3.6 presents the additional analyses, followed by sensitivity analyses in Chapter 3.7. The last section concludes this study.

3.2 Literature Review

3.2.1 Asymmetric Investor Reaction to Guidance

Guidance is forward-looking and unaudited, and thus carries potential risks if incorrect. For this reason, safe harbour provisions were instituted to protect managers from litigation should their expectations of future earnings growth fail to eventuate (Houston et al., 2019; Yang, 2012). In 1995, the United States Congress enacted the Private Securities Litigation Reform Act (PSLRA), which helps protect managers from securities fraud lawsuits stemming from unfulfilled expectations. The PSLRA also exempts managers from obligations to revise their guidance after it is issued, even if market events render their projections unlikely. Managers further protect themselves from lawsuits by adding disclaimer statements to highlight that their projections are by no means guaranteed. Prior studies reveal that firms often fail to meet their own managers' predictions of future earnings growth—from 2002 to 2005, only about 21 percent of quarterly guidance was met (Houston et al., 2010), and from 2000 to 2007, only about 55 percent of annual guidance was met (Hribar & Yang, 2016). When failing to meet market expectations, managers may provide a variety of justifications, such as the occurrence of unanticipated economic events (Hirst et al., 2008), unexpected incidences in internal control systems (Feng et al., 2009) and their overconfidence about the future (Hribar & Yang, 2016).

While the issuance of guidance (i.e. guidance that suggests future earnings above current market expectations) is shown to reduce stock price volatility (Billings et al., 2015; Rogers et al., 2009), managers' ability to influence investors by issuing positive guidance is weakened because investors consider positive guidance less credible and do not rely on it to make their buying decisions. For example, Balakrishnan et al. (2014) find that from the period 1999 through 2009, quarterly stock liquidity increased significantly after the issuance of negative guidance but by a smaller amount after the issuance of positive guidance. Ng et al. (2013) find that from 1995 to 2008, the average stock returns in the three days after issuing positive guidance were +3.79 percent but were –10.90 percent following negative guidance. This asymmetric reaction to positive and negative guidance suggests that investors perceive negative guidance as more credible. To amplify investor reaction to positive guidance, managers may refer to investor sentiment when issuing positive guidance because investor sentiment is strongly associated with the trading activity and direction of prices within a particular market (Audrino et al., 2020; Deng et al., 2018).

3.2.2 NPI Sentiment and Management Guidance Decision

Prior studies (e.g. Brown et al., 2012; Seybert & Yang, 2012; Hurwitz, 2018) propose three views that explain why investor sentiment, which refers to individuals' beliefs about a focal firm that are not consistent with available fundamental information (Baker & Wurgler, 2006; Livnat & Petrovits, 2009; Simpson, 2013), is associated with managers' guidance decision: opportunistic, informative and managerial sentiment. Opportunistic and informative views posit that managers are able to recognize the prevailing investor sentiment for their firms and disclose guidance to either opportunistically exploit investor sentiment for their self-interest or to correct investor sentiment to reduce investors' sentiment-driven misevaluation of firm (Brown et al., 2012; Seybert & Yang, 2012; Hurwitz, 2018). While the managerial sentiment view posits that managers may be susceptible to investor sentiment and disclose guidance corresponding to their beliefs driven by such sentiment (e.g., Brown et al., 2012; Hurwitz, 2018). Although some studies provide support for the association between investor sentiment and managers' guidance decision (e.g., Bergman & Roychowdhury, 2008; Seybert & Yang, 2012; Hurwitz, 2018), it is difficult to disentangle whether the association is because of opportunistic, informative or managerial sentiment view. Also, it is noted that these studies do not attempt to distinguish sentiment of professional investors and sentiment of NPIs.

In the study, I focus on NPIs and their sentiment because NPIs are most likely to be affected by sentiment among all market participants due to their lack of professional advice or fundamental analysis (Barber et al., 2008; Beckman et al., 2012; Aboody et al., 2018). The sentiment of NPIs (NPI sentiment) tends to be irrational and based on individual factors, such as stereotypes (Brave & Nass, 2007; Frijda, 1994), emotions (Reavis, 2012; Shefrin, 2002) and traits (Munezero et al., 2014). However, either informative or managerial sentiment view is less applicable to explain the association between NPI sentiment and managers' guidance decision because managers neither have much incentive to correct NPI sentiment nor are susceptible to NPI sentiment. I follow the opportunistic view and argue that managers opportunistically decide the direction of guidance to respond to NPI sentiment because of an asymmetric stock price response to positive and negative guidance.

3.3 Hypothesis Development

I follow Baker and Wurgler (2006) and characterize investor sentiment as reflecting the propensity to speculate. Speculation is particularly common when investors are confronted with relatively incomplete information regarding a security (Baker & Wurgler, 2006, 2007). A speculator takes on risk, especially with respect to the anticipation of a future price rise, hoping to make profits that are large enough to offset the risk. Baker and Wurgler (2006)'s characterization may describe investor sentiment at multiple levels of aggregation: the market-level, the firm-level and that affecting particular classes of (potential) investors in each firm. Market sentiment refers to the attitude of investors towards the financial market as a whole (Hurwitz, 2018; Simpson, 2013). Positive (bullish) market sentiment suggests a belief that prices will systematically rise in the future, and thus that the current prices of assets are inefficient. During a bubble period, the anticipation of unlimited future growth induces investors to engage in speculative buying (Baker & Wurgler, 2006, 2007). The lack of fundamental information makes the value of some firms (e.g., small firms, young firms and highly volatile firms) more subjective and thus more vulnerable to shifts in the propensity to speculate than others, even if arbitrage forces are the same across stocks (Baker & Wurgler, 2006, 2007; Berkman et al., 2012). The effect of the sentiment towards individual firms on investors' propensity to speculate is similar to that of the market. The firm-specific sentiment is often inferred from news media articles concerning the focal firm (Bhardwaj & Imam, 2019; Giannini et al., 2019) and captures the average degree of optimism or pessimism of investors and potential investors regarding a particular firm's future value. However, the propensity to speculate in the stock of a given firm is likely to differ among classes of investors. In particular, it appears plausible that NPIs may be more susceptible to sentiment (e.g., Renault, 2017; Xiong et al., 2020).

Since they often have less complete information, NPIs are relatively uncertain about a firm's future performance and are therefore prone to defend a valuation that is too high or too low, as suits his or her sentiment (Baker & Wurgler, 2006; Xiong et al., 2020). When an NPI has high sentiment, he or she favourably evaluates the firm and therefore is willing to bear greater risks (Baker & Wurgler, 2006, 2007). NPIs form their sentiment through both cognitive and affective processing routes. Prior research into the cognitive route suggests that NPIs are more likely to form an opinion based on their memories of a firm's historical performance (Nofsinger & Varma, 2013), their psychological acceptance of risk (Hirshleifer, 2001) and their perceptions of a firm's social reputation (Cordeiro & Tewari, 2015) or CEO social ties (Kaplan et al., 2015) than is the case for professional investors. Decisions made based on memories, psychology and perceptions encompass cognitive biases resulting from information processing and memory errors (Hilbert, 2012; Santos & Rosati, 2015). Prior research concerning the affective route argues that investors' judgments regarding a firm's prospects are often influenced by their moods, which in turn may be caused by features of the natural environment or stock market. For example, there is a significant correlation between sunshine (which boosts positive moods) and stock returns (Hirshleifer & Shumway, 2003). A strong correlation also exists between a market frenzy and stock price (Donoughue & Chau, 2021). When the extent of investor attention on a particular firm increases, other NPIs follow to trade the stock. This also attracts new NPIs, who are often younger investors with no investment experience.[15] These new investors are abnormally likely to make decisions based on their perceptions of what other investors are doing rather than their own analysis (Reavis, 2012; Tauni et al., 2020). When the market atmosphere is hot, one emotion – greed – is dominant. This overrides NPIs' cognition and leads to overconfidence biases (Griffith et al., 2020). Overall, NPIs often construct their sentiment with little substantive basis (Griffith et al., 2020; Simpson, 2013). This aligns with empirical findings by Baker and Wurgler (2006) and Xiong et al. (2020) that NPI sentiment exerts larger effects on securities whose valuations are highly subjective and difficult to arbitrage. As a result, when NPIs have high sentiment, their subjectivity guides them to

[15] Over a short period between January 1, 2021 to January 29, 2021, despite no real change in the underlying business, GameStop's share price has surged 1,915 per cent (from $US17.25 to $US347.51). The surge in demand for GameStop stock has been driven by an enthusiastic bunch of NPIs on Reddit, most of them between the age of 18 and 34 (Donoughue & Chau, 2021). In January 2021, Reddit had its highest number of monthly downloads to date. The 6.6 million global downloads represented a 2x increase over the previous year. https://backlinko.com/reddit-users

comfortably bypass opposing opinions and to conduct a financial transaction with a substantial risk of losing value but still hold an expectation of a significant gain.

Taken together, NPI sentiment is conceptually distinct from market sentiment and firm sentiment. Because they are typically at a disadvantage in terms of information, experience and expertise, NPIs' cognitive and affective biases are more likely to dominate their opinions about the firm's attractiveness for investment. When a firm announces new information about its expected future profitability, these biases may guide NPIs to process the new information in a way that suits their sentiment.

The issuance of positive earnings guidance is an example of new information regarding a firm's prospects. I argue that when NPI sentiment is high, positive guidance aligns with NPIs' optimism regarding the firm's future performance, and therefore, this disclosure is less likely to encounter conflicting beliefs that may lead NPIs to discount the news implied (Simpson, 2013). When NPI sentiment is *extremely* high, this implies excessive optimism (i.e., greed occurs). This optimism allows NPIs to accept earnings estimates with less scepticism and reduces their tendency to verify positive guidance with the intrinsic value of the firm's fundamentals (Brown et al., 2012; Simpson, 2013). High sentiment may also make investors overconfident, leading them to be less sensitive to the precision of public information signals and overweigh their private beliefs (Barberis et al., 1998; Daniel et al., 1998). For these reasons, I expect NPIs to react more strongly to positive guidance when NPI sentiment is high than when NPI sentiment is low.

The argument aligns with empirical studies in behavioural finance, which suggest that NPI's speculative behaviour fuels overpricing in financial markets and is particularly salient when NPIs are greedy (Janssen et al., 2019; Reavis, 2012). Greed provokes NPIs to concentrate on potential gains and create unrealistic optimism. Their unrealistic optimism overrides objective assessments of the intrinsic value of the firm's fundamentals in their investment decision, resulting in stock overpricing (Barberis et al., 2018; Reavis, 2012). Theranos is known as the biggest fraud scandal since Madoff in late 2008. In 2014, NPIs held an extremely positive attitude towards the novel blood-testing technology developed by Theranos and did not verify or doubt its unrealistic earnings estimate ($100 million), which was 1,000 times greater than its actual revenue.[16, 17]

Overall, when NPI sentiment is high, NPIs hold a positive opinion about a firm's future value, which induces them to more readily believe good news disclosed by the firm, including that contained in earnings guidance. It is also plausible that managers may benefit from increases in NPI demand for stocks through channels such as the value of equity-based compensation and trading opportunities. Thus, I predict that managers are more inclined to issue positive guidance following periods of high NPI sentiment than low NPI sentiment for a stronger reaction from NPIs. I hypothesise the following:

> HYPOTHESIS. *The likelihood of issued guidance being positive increases with firm-level NPI sentiment*.

---

[16] For further information, see "Theranos President Exaggerated the Company's Revenue by 1,000 Times to Investors, Says SEC" https://www.cnbc.com/2018/03/14/theranos-president-exaggerated-revenue-by-1000x-says-sec.html.

[17] It is noted that none of the major biotech venture capital firms, who took time to understand the business and technology of Theranos, invested in Theranos as they had doubt about the validity of the technology. Please see supporting information, "The Theranos Scandal Is Just The Beginning" https://www.fastcompany.com/3059230/the-theranos-scandal-is-just-the-beginning.

3.4 Sample, Data and Research Design

3.4.1 Sample and Data

The analysis focuses on guidance 'bundled' with an earnings announcement for two reasons.[18] First, following the implementation of Regulation FD in 2000, bundled guidance has become far more common than 'standalone' guidance (Rogers & Van Buskirk, 2013).[19] Immediately prior to Regulation FD, only 25 percent of guidance was bundled with earnings announcements. By 2010, this proportion had increased to over 80 percent (Billings et al., 2015). Second, and more importantly, standalone guidance is typically event-driven (Milian, 2018) and thus issued contemporaneously with other ad hoc disclosures relating to major events such as corporate scandals (Leuz & Schrand, 2009), mergers and acquisitions (Kimbrough & Louis, 2011), and senior management turnover (Brochet et al., 2011). The reporting incentives surrounding these irregular events are diverse and potentially quite different to those surrounding planned regular disclosures, and as such the main tests focus on bundled guidance. However, I discuss the results of analysis using standalone guidance in the additional tests.

Table 1 details the sample selection criteria. The study period extends from the launch of the StockTwits service in May 2008 through to January 2017. The initial sample comprises the intersection of realised quarterly 'street earnings' on the I/B/E/S Actuals file and one period ahead quarterly earnings guidance data from the I/B/E/S Guidance database (46,259 observations). After requiring that each firm-quarter has sufficient data in the I/B/E/S Detail

[18] Following prior research, guidance issued within the five-day window commencing at the earnings announcement date are classified as 'bundled' (Billings et al., 2015; Rogers et al., 2009).

[19] Regulation FD is the rule passed by the SEC that is intended to prohibit private communications between public firms, analysts and certain public shareholders. Public firms are allowed to conduct earnings and forecasts calls to inform analysts and shareholders that are matched with simultaneously issued press release of the statements made during those calls (https://www.sec.gov/fast-answers/answers-regfdhtm.html).

History file to compute analyst consensus forecasts three days prior to each earnings announcement, the potential sample is reduced to 33,197.

Table 1 Sample development—NPI sentiment and management guidance

Sample selection	Firm-quarter observations	Distinct firms
I/B/E/S: Quarterly earnings announcement	147,712	10,060
Less: observations without earnings guidance in current quarter	(101,453)	(7,309)
Number of observations remaining	46,259	2,751
Less: earnings announcement without matched analyst detail	(13,062)	(516)
Number of observations remaining	**33,197**	**2,235**
Less: insufficient StockTwits data to estimate NPI sentiment	(2,646)	(1,025)
Number of observations remaining	30,551	1,947
Less: insufficient TRNA data to estimate News sentiment	(2,475)	(401)
Number of observations remaining	28,076	1,861
Less: observations with missing control variables		
CRSP – daily stock	(757)	(71)
OptionMetrics – the standardised options	(998)	(186)
Thomson Reuters - stock transactions	(721)	(120)
Number of observations with earnings guidance in the current quarter *(Total useable in main analyses)*	**25,600**	**1,625**

Next, I match these observations to the StockTwits data used to estimate NPI sentiment and Thomson Reuters News Analytics (TRNA) data, which I employ to estimate broader news sentiment. To this end, I collect more than 17 million tweets from StockTwits, submitted by 118,685 users that include hashtag references to the sample firms, and perform sentiment analysis to assign a sentiment score to each tweet. Sentiment analysis is implemented using a computer algorithm to mine textual content so as to infer people's attitudes (positive, neutral and negative) towards a subject matter (Audrino et al., 2020; Renault, 2017). I use Valence Aware Dictionary and sEntiment Reasoner[20] (VADER), which is one of the most common

[20] The VADER lexicon contains a list of sentiment-related words, including sentiment word banks such as LIWC, ANEW and GI; sentiment words used in microblogs; and common Western-style emoticons. It works well in the social media context. It is sensitive to both the words in the text and the intensity of sentiments expressed in the way the words are written. It is empirically proved that VADER outperforms individual human raters at correctly classifying the sentiment of tweets into positive, neutral or negative (Gilbert, 2014).

sentiment analysis tools used in the literature.(Gilbert, 2014), to derive the sentiment of these tweets. Appendix A reports examples of these tweets and the VADER sentiment score awarded to them. Each word in the VADER lexicon is assigned a valence score indicating whether it is positive, neutral or negative. I follow the developers' guide and classify each tweet into positive (score > 0.05), negative (score < −0.05) or neutral (score ≥ −0.05 and score ≤ 0.05) categories according to its corresponding compound sentiment score (Gilbert, 2014). After accommodating these data requirements, and those relating to control variables acquired from CRSP, OptionMetrics and Thomson Reuters Insiders, the sample used in the main tests comprises 25,600 firm-quarter observations representing 1,625 distinct firms.

3.4.2 The Impact of NPI Sentiment on NPI Reaction to Positive Guidance

Prior to testing the hypothesis, I examine whether positive guidance issued during periods of high NPI sentiment exhibits exaggerated NPI reaction. I follow the NPI literature and use the overnight ('close-to-open') stock return to measure NPI reaction (Aboody et al., 2018; Berkman et al., 2012; Weißofner & Wessels, 2020) because NPI trading activity is argued to have a greater effect on this metric than on traditional daily cumulative abnormal returns.[21] Model 1 below is estimated using OLS, with standard errors adjusted for clustering by firm and year-month.

[21] The sample used for NPI reaction (23,060) is slightly lower than that our main sample (25,600) due to missing observations for overnight stock returns.

$$
\begin{aligned}
CTO[0,+1]_{i,t} \quad = \quad \alpha_0 \quad & + \quad \alpha_1 GuidePos_{i,t} + \alpha_2 NPISent_{i,t} \\
& + \quad \alpha_3 GuidePos_{i,t} \times NPISent_{i,t} + \alpha_4 NewsSent_{i,t} \\
& + \quad \alpha_5 GuidePos_{i,t} \times NewsSent_{i,t} + \alpha_6 TweetsCount_{i,t} \\
& + \quad \alpha_7 NewsCount_{i,t} + \alpha_8 GuidePos_{i,t} \times NewsCount_{i,t} \\
& + \quad \alpha_9 MktSent_{i,t} + \alpha_{10} GuidePos_{i,t} \times MktSent_{i,t} \\
& + \quad \alpha_{11} Surprise_{i,t} + \alpha_{12} GuidePos_{i,t} \times Surprise_{i,t} \\
& + \quad \alpha_{13} Loss_{i,t} + \alpha_{14} PriorRet_{i,t} + \alpha_{15} MarketCap_{i,t} \\
& + \quad \alpha_{16} AnalystCov_{i,t} + \alpha_{17} MeetBeat_{i,t} + \alpha_{18} LitiRisk_{i,t} \\
& + \quad \alpha_{19} AnalystDisp_{i,t} + \alpha_{20} RetVol_{i,t} + \alpha_{21} Mtb_{i,t} \\
& + \quad \sum \alpha_i Industry_i + \sum \alpha_j FiscalYearQrt_j + \varepsilon_{i,t} \qquad (1)
\end{aligned}
$$

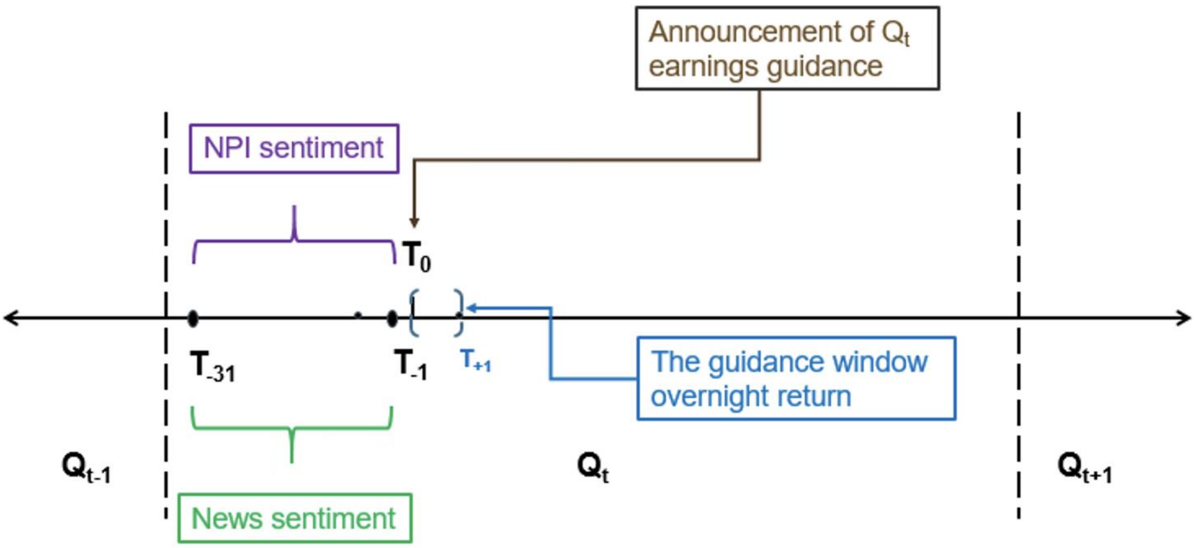
In model 1, the dependent variable is the guidance window overnight return (***CTO[0,+1]***), measured as the natural logarithm of the ratio of the opening price one day after the guidance date to the closing price on the guidance date (Berkman et al., 2012). I identify positive guidance as guidance issued at a level exceeding the consensus (median) analyst forecast for the firm for the same fiscal period observed three days prior and measure this using an indicator variable (***GuidePos***). Guidance issued at the time of quarter t's earning announcement relates to expected earnings per share for the next fiscal quarter (t+1). I define NPI sentiment (***NPISent***) as the difference between the numbers of positive and negative tweets (i.e., net positive tweets) observed in the 30-day window immediately prior to each earnings announcement, consistent with much of the literature (Bhardwaj & Imam, 2019; Bhattacharya et al., 2009).[22] The independent variable of interest is the interaction term between ***GuidePos*** and ***NPISent***, which measures the incremental effect of NPI sentiment on these investors' reactions to positive guidance. I include a number of controls. First, I control for news article sentiment (***NewsSent***), measured as the difference between the numbers of positive and negative news articles in the

[22] Alternatively, the literature proxies NPI sentiment by ***NPISentScaled***, which is the difference between the number of positive tweets and the number of negative tweets in a 30-day window prior to earnings announcement, scaled by the number of total tweets (Audrino et al., 2020; Renault, 2017). I use ***NPISent*** because the number of tweets, which reflects investor attention, plausibly affects managers' likelihood of issuing positive guidance. Nevertheless, using ***NPISentScaled*** in place of ***NPISent*** does not change the substance of our main results.

same time-window, market sentiment (***MktSent***) as measured by the MCSI, and their interaction with ***GuidePos***. I also control for the numbers of tweets (***TweetCount***), news articles (***NewsCount***) and the interaction term between ***GuidePos*** and ***NewsCount***.[23] Further, I follow Seybert and Yang (2012) and Yang (2012) and include variables that may affect the association between guidance and stock return: the quarterly earnings surprise (***Surprise***), measured as the reported actual earnings minus the most recent median analyst estimate, deflated by stock price three days prior to the earnings announcement; loss-making (***Loss***), which equals one if reported earnings are negative, and zero otherwise; the proportion of the four prior quarters for which reported actual earnings met or beat the median analyst estimate (***MeetBeat***); the 90-day stock return ending three days prior to the earnings announcement (***PriorRet***); and the standard deviation of daily stock returns over the 90 days prior to the earnings announcement (***RetVol***). I also include controls for the market value of a firm's equity for each quarter (***MarketCap***), the number of analysts who issued forecasts for the firm in the 90 days prior to the earnings announcement (***AnalystCov***) and the standard deviation of these analyst forecasts (***AnalystDisp***), litigation risk (***LitiRisk***), and market-to-book ratio (***MtB***). All variables used in analyses are defined in detail in Appendix B. Figure 4 shows the timeline and setup in this study.

[23] I do not include the interaction term between ***GuidePos*** and ***TweetCount*** because the correlation between ***NPISent*** and ***TweetCount*** is 0.85, and partially reflects a mechanical relationship between the number of tweets and the difference in the number of positive and negative tweets.

Figure 4 Timeline and setup—NPI sentiment and management guidance



3.4.3 The Impact of NPI Sentiment on the Issuance of Positive Guidance

I then use model 2 to investigate how NPI sentiment affects managers' issuance of positive guidance, using a logistic regression with standard errors adjusted for clustering by firm and year-month.

$$
\begin{aligned}
\Pr(GuidePos_{i,t}) = \alpha_0 &+ \alpha_1 NPISent_{i,t} + \alpha_2 NewsSent_{i,t} + \alpha_3 TweetCount_{i,t} \\
&+ \alpha_4 NewsCount_{i,t} + \alpha_5 MktSent_{i,t} + \alpha_6 IdioVol_{i,t} \\
&+ \alpha_7 MultiGuide_{i,t} + \alpha_8 GuidePrior_{i,t} \\
&+ \alpha_9 InsiderTrade_{i,t} + \alpha_{10} InsiderTradePost_{i,t} \\
&+ \alpha_{11} VIX_{i,t} + \alpha_{12} Surprise_{i,t} + \alpha_{13} Loss_{i,t} \\
&+ \alpha_{14} PriorRet_{i,t} + \alpha_{15} MarketCap_{i,t} + \alpha_{16} AnalystCov_{i,t} \\
&+ \alpha_{17} MeetBeat_{i,t} + \alpha_{18} LitiRisk_{i,t} + \alpha_{19} AnalystDisp_{i,t} \\
&+ \sum \alpha_i Industry_i + \sum \alpha_j FiscalYearQrt_j + \varepsilon_{i,t} \qquad (2)
\end{aligned}
$$

The dependent variable is the indicator for the issuance of positive guidance, ***GuidePos***, measured as described above. The independent variable of interest is ***NPISent***, the coefficient for which measures the impact of NPI sentiment on the likelihood that any guidance issued will be more optimistic than the current analyst consensus forecast. I include controls for other sentiment measures and the volume of news via ***NewsSent***, ***MktSent***, ***TweetCount*** and ***NewsCount***. I also control for factors likely to affect the direction of issued guidance (Billings et al., 2015; Rogers et al., 2009): the average implied volatility in the 30 days prior to each announcement (***IdioVol***), the closing levels of the Chicago Board Option's Exchange volatility index for the three days centred on an announcement (***VIX***) and the firm's guidance pattern (***MultiGuide***), which equals one if the firm has previously issued guidance for earnings that are the subject of the forecast (i.e. quarter t+1); ***GuidePrior***, which equals one if the firm issued guidance for quarter t earnings, during the five-day window centred on the earnings announcement date in the previous quarter (quarter t–1). I also include controls for firm performance in current and prior quarters (***Surprise***, ***Loss*** and ***MeetBeat***), previous stock return

(***PriorRet***), the market value of equity (***MarketCap***), analyst coverage (***AnalystCov***) and their forecast dispersion (***AnalystDisp***), and litigation risk (***LitiRisk***).

3.5 Empirical Results

3.5.1 Descriptive Statistics

Table 2 describes the sample of 25,600 firm-quarter observations from 1,625 distinct firms used in the main analysis. In this sample, 8,679 (34 percent) observations represent positive guidance and 16,921 (66 percent) observations provide either negative or neutral guidance, which is consistent with prior research (e.g., Billings et al., 2015; Rogers et al., 2009). Of the 25,600 observations, 14,882 (58 percent) observations have previously issued guidance for quarter t+1 earnings, and 21,303 (83 percent) cases issued bundled guidance in the quarter immediately preceding the current quarter. Table 3 presents the Pearson correlations and Spearman's rank correlations for key variables.

Table 2 Descriptive statistics—partitioned based on GuidePos

	Full sample (n = 25,600)			***GuidePos* = 1 (n = 8,679)**			***GuidePos* = 0 (n = 16,921)**			**Difference**	
	Mean	Median	Std.dev	Mean	Median	Std.dev.	Mean	Median	Std.dev.	Mean	Median
CTO[0,+1]	0.210	0.058	4.724	0.475	0.193	4.332	−0.243	0.000	4.919	***	***
NPISent	0.988	0.000	2.849	0.982	0.000	2.712	0.990	0.000	2.916		
NewsSent	2.861	1.000	8.136	2.958	1.000	7.980	2.811	1.000	8.214		***
TweetCount	9.382	4.000	92.683	8.386	4.000	72.121	9.893	5.000	101.627		***
NewsCount	14.385	6.000	23.050	14.301	6.000	23.394	14.429	6.000	22.873		***
MktSent	79.021	76.400	10.755	78.429	75.000	10.781	79.324	76.400	10.730	***	***
IdioVol	0.382	0.338	0.186	0.386	0.347	0.181	0.380	0.333	0.189	**	***
MultiGuide	0.581	1.000	0.493	0.553	1.000	0.497	0.596	1.000	0.491	***	***
GuidePrior	0.832	1.000	0.374	0.894	1.000	0.308	0.800	1.000	0.400	***	***
InsideTrade	0.387	0.000	1.027	0.454	0.000	1.149	0.352	0.000	0.956	***	***
InsideTradePost	0.161	0.000	0.607	0.187	0.000	0.665	0.147	0.000	0.574	***	**
VIX	19.784	17.080	8.069	20.206	17.400	8.044	19.568	16.710	8.073	***	***
Surprise	0.001	0.001	0.006	0.002	0.001	0.005	0.001	0.000	0.006	***	***
Loss	0.064	0.000	0.244	0.044	0.000	0.206	0.074	0.000	0.261	***	***
PriorRet	0.049	0.046	0.224	0.053	0.048	0.225	0.047	0.044	0.224	**	*
MarketCap	8.862	2.559	18.057	8.734	2.421	18.240	8.929	2.623	17.962		*
AnalystCov	11.496	10.000	6.890	11.016	9.000	6.716	11.743	10.000	6.965	***	***
MeetBeat	0.672	0.750	0.286	0.726	0.750	0.270	0.644	0.750	0.290	***	***
LitiRisk	0.025	0.025	0.013	0.025	0.024	0.012	0.025	0.025	0.013		**
AnalystDisp	0.028	0.020	0.034	0.029	0.020	0.035	0.028	0.020	0.033	***	***
RetVol	0.039	0.010	0.063	0.020	0.018	0.010	0.023	0.020	0.012	***	***
MtB	3.509	2.483	4.736	3.843	2.674	4.933	3.315	2.379	4.606	***	***

Note: Variables are defined in Appendix B; descriptive statistics for *CTO[0,+1]* are multiplied by 100 for readability; *, **, *** denote instances where the samples differ significantly at the 10%, 5%, 1% level for two-tailed tests.

Table 3 Correlation matrix—NPI sentiment and management guidance

Variable		**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**
GuidePos	(1)	1.00	**0.09**	**0.08**	**0.14**	**0.10**	**0.21**	**0.03**	**−0.09**	**−0.09**	0.01	**0.04**
CTO[0,+1]	(2)	**0.07**	1.00	0.00	0.00	0.00	−0.02	0.00	**0.04**	0.01	0.01	0.08
NPISent	(3)	**0.05**	**−0.02**	1.00	**0.16**	**0.53**	**0.20**	**0.22**	**−0.13**	**0.02**	0.00	**0.06**
NewsSent	(4)	**0.09**	0.00	**0.13**	1.00	**0.22**	**0.36**	**0.11**	**−0.19**	−0.01	**0.02**	**0.05**
TweetCount	(5)	0.01	**−0.03**	**0.85**	**0.09**	1.00	**0.40**	**0.49**	**−0.26**	**0.04**	−0.01	**0.06**
NewsCount	(6)	**0.10**	−0.01	**0.21**	**0.21**	**0.25**	1.00	0.02	**−0.25**	**−0.03**	0.02	**−0.03**
MktSent	(7)	**0.05**	0.00	**0.13**	**0.09**	**0.16**	**−0.03**	1.00	**−0.27**	**0.04**	0.00	**0.03**
IdioVol	(8)	**−0.10**	0.01	**−0.03**	**−0.17**	**−0.03**	**−0.18**	**−0.27**	1.00	**0.12**	**−0.03**	0.02
MultiGuide	(9)	**−0.08**	0.00	**0.04**	−0.01	**0.04**	**−0.04**	**0.04**	**0.09**	1.00	−0.01	**0.07**
GuidePrior	(10)	0.02	0.01	−0.01	0.01	**−0.02**	0.00	0.01	**−0.04**	−0.01	1.00	**0.02**
InsideTrade	(11)	0.01	**0.09**	0.01	0.00	−0.01	**−0.09**	**−0.03**	**0.06**	**0.03**	**0.03**	1.00
InsideTradePost	(12)	0.01	**0.04**	0.00	0.00	−0.01	**−0.06**	**−0.03**	**0.04**	0.01	0.01	**0.56**
VIX	(13)	**−0.03**	0.01	**−0.11**	**−0.15**	**−0.11**	**0.04**	**−0.62**	**0.42**	0.01	**−0.02**	−0.01
Surprise	(14)	**0.05**	**0.18**	−0.01	−0.01	−0.01	**−0.02**	−0.01	**0.09**	**0.02**	−0.01	**0.02**
Loss	(15)	**−0.08**	−0.01	0.01	**−0.05**	0.01	**−0.07**	0.01	**0.29**	**0.04**	**−0.03**	0.00
PriorRet	(16)	0.01	0.01	0.01	0.00	0.01	0.00	0.00	**−0.02**	0.00	−0.01	**0.02**
MarketCap	(17)	**0.09**	−0.01	**0.24**	**0.19**	**0.26**	**0.64**	**0.09**	**−0.32**	**−0.04**	0.00	**−0.11**
AnalystCov	(18)	**0.09**	−0.01	**0.22**	**0.17**	**0.25**	**0.42**	**0.04**	**−0.25**	**0.10**	**0.02**	**−0.06**
MeetBeat	(19)	**0.11**	**0.13**	**0.04**	**0.05**	**0.03**	**0.06**	**−0.03**	**−0.03**	**0.11**	**0.03**	**0.08**
LitiRisk	(20)	**0.06**	−0.03	**0.22**	**0.08**	**0.25**	**0.19**	**0.39**	**−0.09**	**0.09**	0.00	**−0.02**
AnalystDisp	(21)	**0.02**	−0.02	−0.01	0.01	0.00	**0.12**	**0.03**	**−0.04**	**−0.21**	**−0.02**	**−0.06**
RetVol	(22)	**−0.11**	0.01	**−0.02**	**−0.18**	−0.02	**−0.11**	**−0.35**	**0.76**	**0.11**	**−0.06**	**0.06**
MtB	(23)	**0.05**	0.01	**0.07**	**0.07**	**0.06**	**0.04**	**0.11**	−0.09	0.00	0.01	0.07

Variable		**(12)**	**(13)**	**(14)**	**(15)**	**(16)**	**(17)**	**(18)**	**(19)**	**(20)**	**(21)**	**(22)**	**(23)**
GuidePos	(1)	**0.03**	0.00	**0.11**	**−0.07**	0.01	**0.14**	**0.07**	**0.11**	**0.06**	0.01	**−0.08**	**0.08**
CTO[0,+1]	(2)	**0.03**	0.01	**0.29**	0.00	**0.02**	**−0.02**	−0.01	**0.14**	−0.01	**−0.04**	**0.04**	0.01
NPISent	(3)	**0.04**	**−0.18**	**−0.02**	−0.01	0.01	**0.24**	**0.22**	**0.05**	**0.28**	0.00	**−0.10**	**0.17**
NewsSent	(4)	**0.03**	**−0.16**	**−0.02**	**−0.06**	0.01	**0.21**	**0.16**	**0.05**	**0.13**	0.00	**−0.18**	**0.13**
TweetCount	(5)	**0.03**	**−0.39**	**−0.04**	**−0.03**	0.01	**0.47**	**0.43**	**0.08**	**0.59**	**0.05**	**−0.18**	**0.29**
NewsCount	(6)	**−0.04**	**−0.03**	**−0.04**	**−0.08**	0.01	**0.53**	**0.44**	**0.04**	**0.26**	**0.13**	**−0.16**	**0.10**
MktSent	(7)	**0.00**	**−0.53**	**−0.06**	0.01	0.00	**0.11**	**0.02**	**−0.03**	**0.50**	**0.03**	**−0.29**	**0.17**
IdioVol	(8)	**0.04**	**0.36**	**0.19**	**0.25**	**−0.04**	**−0.63**	**−0.27**	0.02	**−0.16**	**−0.15**	**0.85**	**−0.18**
MultiGuide	(9)	**0.04**	**0.06**	**0.04**	**0.05**	0.00	**−0.05**	**0.09**	**0.11**	**0.10**	**−0.30**	**0.15**	**0.03**
GuidePrior	(10)	0.01	**−0.02**	0.00	**−0.03**	0.00	**0.03**	**0.02**	**0.03**	**0.00**	**−0.01**	**−0.04**	**0.02**
InsideTrade	(11)	**0.56**	**−0.04**	**0.03**	0.00	**0.02**	**0.00**	**0.04**	**0.09**	**0.02**	**−0.09**	**0.02**	**0.18**
InsideTradePost	(12)	1.00	**−0.02**	0.01	0.01	0.01	**−0.03**	0.01	**0.07**	0.00	**−0.06**	**0.04**	**0.11**
VIX	(13)	−0.01	1.00	**0.08**	0.00	−0.01	**−0.09**	**−0.02**	**0.03**	**−0.41**	−0.01	**0.37**	**−0.14**
Surprise	(14)	0.00	0.01	1.00	**0.03**	−0.01	**−0.13**	**−0.08**	**0.48**	**−0.02**	**0.02**	**0.17**	**−0.08**
Loss	(15)	0.01	0.02	**−0.04**	1.00	**−0.04**	**−0.26**	**−0.13**	**−0.04**	**−0.04**	0.00	**0.22**	**−0.06**
PriorRet	(16)	0.00	−0.01	0.00	**−0.03**	1.00	**0.04**	0.02	0.01	0.00	0.01	−0.03	0.04
MarketCap	(17)	**−0.08**	**−0.05**	**−0.03**	**−0.10**	0.00	1.00	**0.67**	**0.09**	**0.32**	**0.18**	**−0.53**	**0.30**
AnalystCov	(18)	**−0.06**	**−0.04**	**−0.04**	**−0.13**	0.00	**0.46**	1.00	**0.08**	**0.34**	**0.04**	**−0.18**	**0.26**
MeetBeat	(19)	**0.06**	**0.02**	**0.24**	**−0.05**	0.01	**0.07**	**0.08**	1.00	**0.04**	**−0.04**	**0.03**	**0.12**
LitiRisk	(20)	**−0.02**	**−0.27**	−0.01	−0.02	0.00	**0.17**	**0.31**	**0.03**	1.00	**0.05**	**−0.10**	**0.17**
AnalystDisp	(21)	**−0.05**	0.01	−0.01	**0.03**	0.00	**0.06**	0.00	**−0.06**	**0.07**	1.00	**−0.10**	**−0.15**
RetVol	(22)	**0.04**	**0.53**	**0.05**	**0.25**	−0.02	**−0.27**	**−0.17**	−0.01	**−0.04**	**0.02**	1.00	**−0.15**
MtB	(23)	**0.03**	**−0.09**	**−0.02**	0.01	**0.02**	**0.11**	0.13	**0.03**	**0.07**	**−0.06**	**−0.07**	1.00

Note: Variables are defined in Appendix B; bold typeface indicates significance at the 1% level. Pearson’s correlation coefficients are shown in the lower triangle (shaded), including the diagonal, and Spearman’s rank correlations appear above the diagonal.

3.5.2 The Impact of NPI Sentiment on NPI Reaction to Positive Guidance

NPI reaction to guidance is measured by the overnight return spanning the closing price on the day guidance was issued to the opening price the following day (***CTO[0,+1]***). Table 4 presents the results of tests of NPI reaction to positive guidance, and the conditioning role of NPI sentiment on this relationship, using a sample of firm-quarters where the guidance of any direction was provided. All reported coefficients have been standardised to allow comparison of their relative size. Column I of Table 4 shows the results of regressions based on model 1, but in which ***NPISent*** is omitted.[24] Column II of Table 4 shows the results for the full specification of model 1, which includes ***NPISent*** and its interaction with ***GuidePos***. The coefficient for ***GuidePos*** × ***NPISent*** is positive and significant ($\beta = 0.021$, $p = 0.024$), suggesting that when positive guidance is issued at the time of high NPI sentiment, it garners a stronger overnight reaction than when positive guidance is issued at times of low NPI sentiment. The economic effect of NPI sentiment on the reaction to positive guidance is also meaningful; a one standard deviation increase in ***NPISent*** is associated with an increase in overnight returns of approximately 14.6 basis points. Overall, the results in Table 4 align with my expectation that positive guidance issued following periods of high NPI sentiment exhibits a stronger NPI reaction than similar guidance issued when NPI sentiment is low.[25]

[24] Because I interact ***GuidePos*** with several factors, the main effect for ***GuidePos*** does not have a simple interpretation. If I exclude these interactions, the coefficient on ***GuidePos*** is positive and significant ($\beta = 0.059$, $p < 0.001$), suggesting that positive guidance induces a higher overnight return than negative or neutral guidance.

[25] In Table A1, I include a continuous measure of guidance surprise (***GuideSurp***), measured as guidance estimate minus the prevailing median analysts' estimate, deflated by stock price three trading days prior to the guidance issuance, and its two-way and three-way interactions with ***GuidePos*** and ***NPISent***. The coefficient for ***GuidePos*** x ***NPISent*** remains positive and significant ($\beta = 0.030$, $p = 0.002$), but the effect of ***NPISent*** on reaction to positive guidance does not increase significantly with the magnitude of the guidance surprise ($\beta = -0.011$, $p = 0.329$).

Table 4 NPI reaction to positive guidance

	Dependent variable = *CTO[0,+1]*			
	Column I		Column II	
Variable	coef.	t-stat.	coef.	t-stat.
GuidePos	0.047	0.929	0.051	1.023
NPISent			−0.048	−0.963
GuidePos×NPISent			0.021	2.250**
NewsSent	0.009	1.064	0.010	1.147
GuidePos×NewsSent	−0.010	−1.202	−0.010	−1.229
TweetCount			0.032	0.728
NewsCount	0.001	0.136	0.003	0.305
GuidePos× NewsCount	−0.020	−2.562**	−0.023	−2.954**
MktSent	0.227	1.202	0.227	1.203
GuidePos× MktSent	0.016	0.323	0.011	0.213
Surprise	0.140	11.488***	0.140	11.487***
GuidePos× Surprise	0.035	2.627***	0.035	2.633***
Loss	0.000	−0.006	0.000	0.012
PriorRet	0.013	1.627	0.013	1.635
MarketCap	0.002	0.290	0.003	0.407
AnalystCov	−0.003	−0.357	−0.003	−0.324
MeetBeat	0.092	13.229***	0.092	13.221***
LitRisk	−0.037	−2.603***	−0.036	−2.553**
AnalystDisp	−0.009	−1.361	−0.009	−1.363
RetVol	0.005	0.410	0.006	0.454
MtB	0.005	0.727	0.005	0.725
Industry FE	Yes		Yes	
Yearr-Quarter FE	Yes		Yes	
Observations	23,061		23,061	
F statistics	8,295		8.054	
Prob > F	0.000		0.000	
Adjusted R^2	0.048		0.048	

Note: Table 4 presents standardized coefficients and t-statistics from regressions of NPI reaction on the interaction between positive guidance and NPI sentiment and relevant controls. *CTO[0,+1]*, an NPI reaction proxy, is the overnight stock return (close-to-open price) measured as the natural logarithm of the ratio of closing price on guidance date and opening price one day after guidance date; *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the net number of positive news, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing a firm (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan Consumer Sentiment Index; *Surprise* measures earnings surprise, which equals the actual earnings minus the prevailing median analysts' estimate, deflated by stock price three trading days prior to the guidance issuance; *Loss* is an indicator of loss making that equals 1 if reported earnings are negative, and 0 otherwise; *PriorRet* measures the stock return over the 90-day period ending three trading days prior to the guidance issuance; *MarketCap* is the market value of equity (i.e., stock price multiplied by number of stocks outstanding) measured at three trading days prior to the guidance issuance; *AnalystCov* is measured as the number of analysts with outstanding estimates three trading days prior to the guidance issuance; *MeetBeat* is the proportion of the previous four quarters in which firms' reported earnings met or exceeded analysts' prevailing median consensus estimates; *LitiRisk* is a measure of litigation risk estimated from the probit model by Kim and Skinner (2012); *AnalystDisp* measures the standard deviation of prevailing analysts' estimates for the current period's earnings; *RetVol* measures the standard deviation of daily stock returns over the 90 days prior to the guidance issuance; *MtB* is the market-to-book ratio. Industry-fixed effects based on the Fama–French 48 classification and fiscal quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

3.5.3 The Impact of NPI Sentiment on the Issuance of Positive Guidance

Table 5 presents the results of logistic regressions of the issuance of positive guidance on NPI sentiment and controls for firm-quarters where the guidance of any type is issued. Results in Column I report a positive and significant coefficient for ***NPISent*** ($\beta = 0.024$, $p = 0.008$), consistent with more positive NPI sentiment increasing the likelihood that managers issue positive guidance. The marginal effect of ***NPISent***, holding all covariates constant at their mean (untabulated), indicates that for a one standard deviation increase in NPI sentiment, the likelihood of issuing positive guidance increases by 1.54 percent. The coefficient for ***NewsSent*** is also positive and significant ($\beta = 0.004$, $p = 0.037$), suggesting a positive association between broader firm-specific sentiment and managers' tendency to issue positive guidance. However, the marginal effect of firm-specific news sentiment (one standard deviation increase in news sentiment increases the likelihood of issuing positive guidance by 0.93 percent) is weaker than that of NPI sentiment ($\chi = 4.658$, $p = 0.031$). The coefficient for the market-wide sentiment (***MktSent***) is also positive and significant ($\beta = 0.020$, $p = 0.022$). The marginal effect of market-wide sentiment (one standard deviation increase in market sentiment increases the likelihood of issuing positive guidance by 2.57 percent) is not significantly different from that of NPI sentiment ($\chi = 0.092$, $p = 0.762$). The results in Column I support my hypothesis that the likelihood of managers issuing positive guidance increases as NPI sentiment increases.

Because the provision of guidance is discretionary, it is possible that the main test is influenced by selection bias. To investigate these concerns, I follow the Heckman two-stage approach used in Fang and Peress (2009) and Cheng et al. (2013). In so doing, I employ an initial sample of 64,190 firm-quarters (3,356 distinct firms) with full data for the independent variables. Approximately 40 percent of firm-quarters earnings announcements within this sample are accompanied by guidance, and 60 percent are not. These percentages are consistent with prior studies (e.g., Billings et al., 2015; Rogers & Van Buskrik, 2013). In the first stage regression, I model the likelihood of issuing guidance (***Guide***) as a function of NPI sentiment (***NPISent***), news sentiment, market sentiment, firms' guidance history, information environment, insider trading incentives, earnings news and litigation risk, as shown in model 3. I also include Δ***IdioVolPrior***, measured as the change in implied volatility over the 15-day window preceding the earnings announcement date, which Billings et al. (2015) shows is positively associated with the likelihood of bundled guidance being issued. Importantly, I see no obvious reason why this measure should be associated with the direction of that guidance.[26]

$$\begin{aligned}
\Pr(Guide_{i,t}) \quad = \quad \alpha_0 \quad & + \quad \alpha_1 NPISent_{i,t} + \alpha_2 NewsSent_{i,t} + \alpha_3 TweetCount_{i,t} \\
& + \quad \alpha_4 NewsCount_{i,t} + \alpha_5 MktSent_{i,t} + \alpha_6 \Delta IdioVolPrior_{i,t} \\
& + \quad \alpha_7 IdioVol_{i,t} + \alpha_8 MultiGuide_{i,t} + \alpha_9 GuidePrior_{i,t} \\
& + \quad \alpha_{10} InsiderTrade_{i,t} + \alpha_{11} InsiderTradePost_{i,t} \\
& + \quad \alpha_{12} VIX_{i,t} + \alpha_{13} Surprise_{i,t} + \alpha_{14} Loss_{i,t} \\
& + \quad \alpha_{15} PriorRet_{i,t} + \alpha_{16} MarketCap_{i,t} + \alpha_{17} AnalystCov_{i,t} \\
& + \quad \alpha_{18} MeetBeat_{i,t} + \alpha_{19} LitiRisk_{i,t} + \alpha_{20} AnalystDisp_{i,t} \\
& + \quad \sum \alpha_i Industry_i + \sum \alpha_j FiscalQrt_j + \varepsilon_{i,t}
\end{aligned} \tag{3}$$

I then include the inverse Mills ratio (***IMR***) from the selection regression (model 3) in the structural model (based on model 2) to control for selection effects. Results of the regression of model 3 are reported in Column II of Table 5. The coefficient for my instrument, Δ***IdioVolPrior***, is positive and significant (β = 0.175, $p < 0.001$). The coefficient for ***NPISent*** is insignificant (β = −0.002, $p = 0.527$), and thus there is no evidence NPI sentiment affects the decision to issue guidance. In the second stage, I regress ***GuidePos*** on ***NPISent***, the inverse Mills ratio (***IMR***) obtained from the first-stage logistic regression and control variables. Results for the estimation of my second stage model, which estimates the likelihood of issued guidance

[26] The coefficient for ***ΔIdioVolPrior*** is not significant if included in the second stage regression. Fang and Peress (2009) and Cheng et al. (2013) use analyst following as an instrument when predicting the issuance of guidance. However, this variable is included in our structural model and is significantly associated with the likelihood of issued guidance being positive.

being positive after controlling for selection bias, are reported in Column III. Consistent with the main analyses, the coefficient on ***NPISent*** is positive and significant ($\beta = 0.005$, $p = 0.007$), which suggests that the results are robust to the correction of self-selection bias (***IMR***: $\beta = -0.362$, $p < 0.001$).

Table 5 NPI sentiment and the direction of guidance

			Heckman two-stage model			
	Column I: Logistic regression		Column II: First stage		Column III: Second stage	
	Dependent variable = *GuidePos*		Dependent variable = *Guide*		Dependent variable = *GuidePos*	
Variable	**coef.**	**t-stat.**	**coef.**	**t-stat.**	**coef.**	**t-stat.**
NPISent	0.024	2.640***	–0.002	–0.633	0.005	2.693***
IMR					−0.362	−16.007***
NewsSent	0.004	2.090**	0.003	2.675***	0.000	1.140
TweetCount	−0.004	−1.882	–0.007	–7.303***	0.001	1.465
NewsCount	−0.001	−1.047	0.001	1.734	0.000	−1.595
MktSent	0.020	2.291**	–0.008	–1.619	0.005	2.857***
∆IdioVolPrior			0.175	4.585***		
IdioVol	0.135	1.316	–0.499	–9.854***	0.113	5.231***
MultiGuide	−0.384	−11.870***	0.892	43.035***	−0.195	−19.269***
GuidePrior	0.829	18.583***	1.945	121.887***	−0.228	−8.738***
InsideTrade	0.066	4.227***	0.015	1.643	0.014	3.875***
InsideTradePost	0.018	0.696	0.004	0.256	0.003	0.535
VIX	0.027	2.374**	–0.008	–1.230	0.007	2.754***
Surprise	30.166	7.639***	–0.961	–1.135	4.239	8.078***
Loss	−0.487	−7.096***	–0.301	–11.122***	−0.033	−2.614**
PriorRet	0.088	1.447	–0.018	–0.584	0.020	1.559
MarketCap	0.002	1.876*	0.003	5.065***	0.000	0.408
AnalystCov	−0.023	−8.644***	0.007	5.338***	−0.006	−10.449***
MeetBeat	0.927	17.163***	0.263	9.915***	0.160	15.004***
LitiRisk	5.706	3.414***	1.608	5.122***	0.947	2.842***
AnalystDisp	1.057	2.398**	–2.284	–12.101***	0.712	7.583***
Industry FE	Yes		Yes		Yes	
Year−Quarter FE	Yes		Yes		Yes	
Observations	25,600		64,190		25,600	
LR chi^2 or F statistics	1434.586		25337.202		24.153	
Prob > chi^2 or Prob>F	0.000		0.000		0.000	
Pseudo R^2 or Adjusted R^2	0.055		0.587		0.070	

Note: Table 5 Column I presents results of a logistic regression of the incidence of positive guidance on NPI sentiment and relevant controls. Columns II and III document results from a two-stage Heckman selection regime, in which both stages are estimated by logistic regression. *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *IMR* is the inverse Mills ratio obtained from the first-stage model of the Heckman two-stage model; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the net positive news, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing a firm (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan Consumer Sentiment Index; *ΔIdioVolPrior* is the natural logarithm of the ratio of implied volatility measured at the close of day prior to the report date of quarterly earnings to implied volatility measured 15 days prior to the report date of earnings. *IdioVol* is the average implied volatility (for a 30-day window, at the money option) measured at the release date of guidance issuance; *MultiGuide* is an indicator variable that equals 1 if the firm previously provided earnings guidance for the current quarter's earnings, and 0 otherwise; *GuidePrior* is an indicator variable that equals 1 if the firm issued an earnings guidance during the five-day window centred on the guidance issuance last quarter, and 0 otherwise; *InsiderTrade* measures the total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the current quarter; *InsiderTradePost* measures the total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the 15 days after the earnings announcement; *VIX* is the Chicago Board Option's Exchange Implied Volatility Index (a.k.a. the 'fear index') during the three-day window centred on the guidance issuance; *Surprise* measures earnings surprise that equals the actual earnings minus the prevailing median analysts' estimate, deflated by stock price three trading days prior to the guidance issuance; *Loss* is an indicator of loss making that equals 1 if reported earnings are negative, and 0 otherwise; *PriorRet* measures the stock return over the 90-day period ending three trading days prior to the guidance issuance; *MarketCap* is the market value of equity (i.e., stock price multiplied by number of stocks outstanding) measured at three trading days prior to the guidance issuance; *AnalystCov* is measured as the number of analysts with outstanding estimates three trading days prior to the guidance issuance; *MeetBeat* is the proportion of the previous four quarters in which firms' reported earnings met or exceeded analysts' prevailing median consensus estimates; *LitiRisk* is the measure of litigation risk estimated from the probit model by Kim and Skinner (2012); *AnalystDisp* measures the standard deviation of prevailing analysts' estimates for the current period's earnings. Industry-fixed effects based on the Fama–French 48 classification and fiscal quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

3.6 Additional Analyses

3.6.1 Addressing Endogeneity Concerns

3.6.1.1 Do managers really consider NPI sentiment when issuing earnings guidance? Exploiting an exogenous increase in NPI skepticism

My hypothesis hinges on the assumption that managers consider NPI sentiment when issuing guidance and that NPI sentiment can be inferred from online social network postings. However, it is plausible that the measure of NPI sentiment may be correlated with some other time-varying factor that drives managers' guidance decisions, raising the prospect of an omitted variable threat. To investigate this possibility further, we exploit an (arguably) exogenous shock to the NPIs' scepticism about social media discussions in an attempt to achieve stronger identification.On 8 April 2014, the SEC issued a press release stating that it had filed fraud charges against the operators of a South Florida-based Ponzi scheme targeting investors through YouTube videos and a Honolulu woman posing as an investment banker and soliciting investors through Twitter, Facebook, and other social media.[27] Following these charges, on 25 July 2014, the SEC issued a highly publicised investor alert warning NPIs about fraudsters who attempt to manipulate share prices by spreading rumours on social media. One recent example is Michael M. Beck, who used the Twitter handle @BigMoneyMike6, engaged in the scalping of eight different penny stocks - recommending a stock without disclosing his intent to sell the stock and then selling it at inflated prices to generate profits. This SEC alert was widely disseminated to investors by news media (e.g., the Wall Street Journal and the Reuters) and social media (e.g., Twitter and StockTwits). It thus appears reasonable to assume that following this SEC alert, at least some NPIs would have become increasingly suspicious about the

[27] For further information, see "SEC, Criminal Authorities Halt Florida-Based Ponzi Scheme Targeting Investors through YouTube Videos" https://www.sec.gov/news/press-release/2014-70; and "SEC Announces Charges Against Honolulu Woman Defrauding Investors Through Social Media" https://www.sec.gov/news/press-release/2014-72.

originality of social media discussions and resultant to speculate on price movement based on these discussions. As a result, our social media-based sentiment measure (NPISent) becomes less effective in reflecting the NPIs' propensity to speculate after this SEC alert. Consequently, if our main results do reflect the manager's reliance on social media–based NPI sentiment, then the association between our NPI sentiment and guidance decisions should be weaker in the post-alert period.

To test this contention, I generate a binary variable (***SECAlert***) indicating guidance issued after the date of the SEC fraud charges (8 April 2014) and re-estimate model 2, including ***SECAlert*** and its interaction with ***NPISent***, ***NewsSent*** and ***NewsCount***.[28] Column I of Table 6 presents the results. The coefficient for ***NPISent***, measuring the effect of ***NPISent*** in the pre-alert period, is positive and significant ($\beta = 0.034$ $p = 0.001$). In support of my expectation, the coefficient on ***NPISent*** × ***SECAlert*** is negative and marginally significant ($\beta = -0.018$, $p = 0.073$).[29] This result is consistent with managers considering NPI sentiment in their guidance decisions and that the tone of online social network discussions represents a source that they may have considered.

[28] I do not include the interaction term between ***SECAlert*** and ***TweetCount*** because the correlation between ***NPISent*** and ***TweetCount*** is 0.85

[29] Because our model includes quarter fixed effects, the main effect for ***SECAlert*** does not have a meaningful interpretation.

Table 6 Additional analyses: The effect of an exogenous shock on public awareness of fake tweets and 2SLS regressions

			2SLS regressions			
	Column I: Exogenous shock effects		**Column II: First stage**		**Column III: Second stage**	
	Dependent variable = *GuidePos*		**Dependent variable = *NPISent***		**Dependent variable = *GuidePos***	
Variable	**coef.**	**t-stat.**	**coef.**	**t-stat.**	**coef.**	**t-stat.**
NPISent	0.034	3.192***				
SECAlert	5.687	1.167				
NPISent × SECAlert	−0.018	−1.796*				
NPISentPred					0.120	2.822***
NPISent4Q			0.093	7.527***		
SalesGrowth			0.000	0.000		
RetVol			0.141	0.119		
TradeVolume			0.004	2.885***		
LitiIndustry			0.113	2.717***		
NewsSent	0.004	1.632	0.016	12.937***	–0.001	–1.419
NewsSent × SECAlert	0.001	0.290				
TweetCount	−0.004	−1.918*	0.182	232.551***	–0.022	–2.809***
NewsCount	−0.002	−1.247	–0.003	–5.648***	0.000	0.506
NewsCount × SECAlert	0.079	1.426				
MktSent	−0.063	−1.091	–0.000	–0.001	–0.014	–0.789
MktSent × SECAlert	0.137	1.335				
IdioVol	−0.375	−11.474***	0.030	0.419	0.016	0.683
MultiGuide	0.826	18.534***	–0.040	–1.813*	–0.080	−10.812***
GuidePrior	0.066	4.190***	–0.031	–1.168	0.156	17.930***
InsideTrade	0.019	0.721	0.025	2.313**	0.011	2.907***
InsideTradePost	0.090	1.839*	–0.018	–1.017	0.006	0.996
VIX	29.932	6.705***	–0.001	–0.039	0.005	0.530
Surprise	1.024	0.122	0.204	0.119	4.829	7.503***
Loss	−0.487	−7.093***	0.008	0.205	–0.082	−6.146***
PriorRet	0.088	1.445	0.055	1.321	0.006	0.404
MarketCap	0.002	2.022**	0.005	6.215***	0.000	0.202
AnalystCov	−0.023	−8.711***	–0.004	–2.183**	–0.005	−7.335***
MeetBeat	0.927	17.146***	0.121	3.411***	0.178	14.098***

LitiRisk	5.916	3.567***	–3.172	–1.071	1.238	1.877*
AnalystDisp	1.053	2.386**	–0.482	–1.499	0.395	3.598***
Industry FE	Yes		Yes		Yes	
Year-Quarter FE	Yes		Yes		Yes	
Observations	25,600		24,735		24,735	
LR chi^2 or F statistics	144.752		676.258		15.279	
Prob > chi^2 or Prob>F	0.000		0.000		0.000	
Pseudo R^2 or Adjusted R^2	0.055		0.737		−0.066	
The area under ROC	0.6628					

Note: Column I of Table 6 reports results of a logistic regression examining the effect of an exogenous increase to public awareness of fake tweets. Columns II and III of Table 6 report results of a 2SLS regression approach intended to address reverse causality concerns. *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *SECAlert* is an indicator that equals 1 if an earnings guidance is issued after the date of the SEC investor alert, and 0 otherwise. *NPISent4Q* measures the average of *NPISent* for the prior four quarters. *NPISentPred* is the predicted NPI sentiment from Column II; *SalesGrowth* is measured as the change in quarterly sales deflated by total assets; *RetVol* measures the standard deviation of daily stock returns over the 90 days prior to the guidance issuance; *TradeVolume* is the annual trading volume deflated by the beginning-of-the-year outstanding shares; *LitiIndustry* is an indicator that equals 1 if the firm is in industries with high litigation risk (i.e., four-digit Standard Industrial Classification [SIC] code between 2833 and 2836, 3570 and 3577, 3600 and 3675, 5200 and 5961, and 7370 and 7374), and 0 otherwise. Other variables are defined in Appendix B. Industry-fixed effects based on Fama–French 48 classification and fiscal quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

3.6.1.2 Assessing the Magnitude of the Omitted Variable Threat

Notwithstanding the exogenous shock test results described above, and the fact that my models include a comprehensive set of (observable) covariates and fixed effects to control for unobservable heterogeneity, there remains a possibility that the omission of other unobservable factors may materially bias the findings. To assess the extent of this threat, I conduct a test proposed by Oster (2019) that computes the share of variation of unobservable factors (relative to observables) that is required to 'explain away' the observed effects (measured by the test statistic, δ) and reduce the effect of variable of interest, NPI sentiment (***NPISent***), to zero.[30]

The Oster (2019) test requires us to set a value of R_{max}, which denotes the value of R-squared from a hypothetical regression that consists of both unobservables and observables. Following Oster (2019) and Babenko et al. (2020), I first specify R_{max} as $1.3 \times R_{squared}$ (1.3×0.053), where $R_{squared}$ is the R-squared from an ordinary least squared (OLS) model that regresses ***GuidePos*** on ***NPISent*** and includes all observables (with industry and year-quarter fixed effects).[31] The obtained δ is –14.634, which indicates that the unobservables need to be more than 14 times as significant as the observables to reduce the effect of NPI sentiment to zero, which seems highly unlikely given that my regression model includes many factors known to affect positive guidance issuance and a number of fixed effects.[32] Therefore, it appears unlikely that omitted variables have materially biased the findings.

[30] For example, δ = 2 suggests that the unobservable variables need to be twice as significant as observables for the omitted variable bias to 'explain away' the results and decrease the coefficient of interest to zero (Babenko et al., 2020). Because $\delta\frac{\sigma_1 X}{\sigma_1{}^2}=\frac{\sigma_2 X}{\sigma_2{}^2}$ in Oster (2019), a negative δ means that if the observables are positively correlated with the treatment, the unobservables must be negatively correlated with the treatment to decrease the coefficient of interest to zero.
[31] I use the Stata command *psacalc* provided by Oster (2019) to conduct the test.
[32] The same conclusion is reached if I opt for a more stringent threshold by setting R_{max} as $2 \times R_{squared}$ (2×0.053). The value of δ is -4.392, which suggests that the unobservables need to be more than four times as significant as the observables to reduce the effect of NPI sentiment to zero, which again is unlikely.

3.6.1.3 Addressing reverse causality concerns

Although much less intuitive, a potential alternative explanation for my results could be that NPIs' expectation of forthcoming guidance direction affects their sentiment and that causation is opposite to the direction that this study posits. I believe that this alternative explanation is much less convincing for two reasons. First and most importantly, the issuance, contents and format of earnings guidance is largely at the managers' discretion. Second, while analysts and institutional investors usually anticipate the announcement of future management guidance and acquire private information to maximize the potential gains from trading on that information before the public announcement (Altschuler et al., 2015; Billings et al., 2015), NPIs often search for private information after the public announcement of management guidance but not before (Cho & Kwon, 2014). Nevertheless, I adopt a standard 2SLS approach to address reverse causality concerns similar to mine. In the first stage, I model expected NPI sentiment (***NPISent***) using my existing controls and a vector of instrumental variables comprising the mean NPI sentiment in the previous four quarters (***NPISent4Q***) and firm-specific characteristics that may affect NPIs' attitude towards a firm: quarterly sale growth (***SalesGrowth***), stock return volatility (***RetVol***), trading volume (***TradeVolume***) and an indicator for litigious industries (***LitiIndustry***) (each defined in detail in Appendix B). These candidate instruments collectively satisfy standard tests for weak identification (their F-statistic of 22.176 exceeds the critical value associated with 5 percent relative bias in the Stock-Yogo test) and exogeneity (Sargan-Hansen overidentification test *p-value* = 0.124). I then substitute the predicted NPI sentiment (***NPISentPred***) estimated from the first stage coefficients for actual ***NPISent*** and re-estimate model 2. As shown in Column III of Table 6, the coefficient for ***NPISentPred*** is again positive and significant ($\beta = 0.120$, $p = 0.005$). The coefficients on other variables are also qualitatively similar to those in Table 5, suggesting that reverse causality is not a serious concern.

Another potential issue that may bias my results is the potential for guidance disclosures to influence NPI sentiment, that is, a high NPI sentiment in StockTwits is caused by a positive guidance issued at the earnings announcement in previous quarter. To address this issue, I re-run model while controlling for the disclosure of bundle guidance (***GuidePosPre***) in the previous quarter. ***GuidePosPre*** equals one if a firm disclose a positive guidance at the earnings announcement in previous quarter and zero for neutral or negative guidance. As shown in the untabulated results, the coefficient of ***GuidePosPre*** is positive and significant ($\beta = 0.503$, $p < 0.001$). The coefficient of ***NPISent*** is positive and significant ($\beta = 0.048$, $p < 0.001$), consistent with the main results.

3.6.2 Reconciling Competing Perspectives on Guidance: The Opportunistic, Informative, and Managerial Sentiment Views

3.6.2.1 Tests of the managerial sentiment view

The managerial sentiment perspective on the determinants of guidance behaviour is distinct from the opportunistic and informative views, arguing that managers are susceptible to the same market sentiment as investors and may thus act similarly irrationally in the presence of such sentiment (e.g., Brown et al., 2012; Hurwitz, 2018). The rationale behind this view is that managers may not have access to sufficient resources and macroeconomic expertise to fully understand and analyse market-wide factors (e.g., Brown et al., 2012; Hurwitz, 2018) and separate fundamental effects from hubris. While the managerial sentiment view is plausibly descriptive of guidance behaviour in general, there is no obvious reason to expect that it would bias my tests, given that this phenomenon is argued to be driven by market-wide sentiment rather than firm-specific non-professional investor sentiment (Aboody et al., 2018; Weißofner & Wessels, 2020), and the fact that I control for market sentiment in all regressions.

Nevertheless, I estimate tests to assess the plausibility of a managerial sentiment explanation for my findings. To this end, I allow the impact of NPI sentiment on managers' guidance decisions to vary according to whether NPI sentiment is consistent with market sentiment. If the managerial sentiment view explains the main results, I should not find a significant association between NPI sentiment and the likelihood of issuing positive guidance in cases where NPI sentiment contradicts market sentiment.

The analysis does not support the managerial sentiment view. I first use the industry-quarter median values of NPI sentiment and historical median values of market sentiment to identify disagreement in these sentiment measures. ***Disagree*** is a dichotomous variable equal to one where NPI sentiment is above (below) its industry-quarter median value and market sentiment is below (above) its historical median value. I then re-estimate model 2, including ***Disagree*** and its interaction with ***NPISent***. Results reported in Column I of Table 7 show that the main effect for ***NPISent*** has a positive and significant coefficient ($\beta = 0.036$, $p = 0.008$), indicating that ***NPISent*** has a positive association with ***GuidePos*** when NPI sentiment and market sentiment agree. There is no evidence that this association is weaker when the sentiment measures disagree. The interaction between ***NPISent*** and ***Disagree*** is positive but insignificant ($\beta = 0.002$, $p = 0.852$). These findings are inconsistent with a managerial sentiment explanation for the main findings. I obtain results of similar tenor using a different model specification (Column II) in which ***NPISent*** is interacted with an indicator of the above-median market sentiment (***HighMktSent***). ***NPISent*** remains positive and significant ($\beta = 0.041$, $p = 0.008$), while the interaction term for ***NPISent*** and ***HighMktSent*** is negative and significant ($\beta = –0.022$, $p = 0.049$), suggesting the association between ***NPISent*** and ***GuidePos*** is indeed weaker when market sentiment is high than low. The results provide no evidence that the main results are driven by an alignment between NPI and market sentiment.

Table 7 Distinguishing managerial opportunism from managerial sentiment

	Dependent variable = *GuidePos*			
	Column I: NPI sentiment disagree to market sentiment		**Column II: High *vs* low market sentiment**	
Variable	**coef.**	**t-stat.**	**Coef.**	**t-stat**
NPISent	0.036	3.440***	0.041	3.262***
Disagree	0.075	2.336**		
NPISent ×Disagree	0.002	0.187		
HighMktSent			–1.056	–0.512
NPISent ×HighMktSent			–0.022	–1.971**
NewsSent	0.004	2.114**	0.004	2.143**
TweetCount	–0.005	–2.330**	–0.004	−2.030*
NewsCount	−0.001	−1.629	–0.001	−1.169
MktSent	0.027	2.881***	0.023	2.423**
IdioVol	0.123	1.187	0.132	1.284
MultiGuide	−0.412	−12.600***	–0.385	−11.913***
GuidePrior	0.699	15.491***	0.827	18.540***
InsideTrade	0.067	4.193***	0.066	4.197***
InsideTradePost	0.021	0.775	0.019	0.707
VIX	0.033	2.765**	–0.010	−0.137
Surprise	32.618	7.861***	30.213	7.647***
Loss	−0.462	−6.575***	–0.485	−7.072***
PriorRet	0.076	1.226	0.087	1.429
MarketCap	0.002	1.539	0.002	1.828*
AnalystCov	−0.024	−8.784***	–0.023	−8.662***
MeetBeat	0.894	16.245***	0.927	17.148***
LitiRisk	6.455	3.639***	5.804	3.500***
AnalystDisp	1.000	2.168**	1.052	2.385**
Industry FE	Yes		Yes	
Quarter FE	Yes		Yes	
Observations	25,600		25,600	
LR chi^2	1334.466		1436.863	
Prob > chi^2	0.000		0.000	
Pseudo R^2	0.051		0.055	
The area under ROC	0.6561		0.6626	

Note: Table 7 reports the results of a logistic regression in which *NPISent* is interacted with indicators of disagreement with market sentiment (Column I) and high market sentiment (Column II). *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the net positive news, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan Consumer Sentiment Index; *Disagree* is an indicator that equals 1 when NPI sentiment and market sentiment disagree (i.e., high (low) NPI sentiment and low (high) market sentiment; low NPI sentiment and high market sentiment), and 0 otherwise; *HighMktSent* is an indicator that equals 1 if *MCIS* is above the median value, and 0 otherwise. Other variables are defined in Appendix B. Industry-fixed effects based on Fama–French 48 classification and fiscal quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

3.6.2.2 Distinguishing the opportunistic and informative views of the provision of positive guidance

Both the opportunistic and informative views assume that managers can access information about investor sentiment regarding their firm and use this information in preparing their guidance (Brown et al., 2012; Seybert & Yang, 2012). In my context, the two views differ in managers' motives behind issuing positive guidance according to NPI sentiment—the opportunistic view holds that managers consider NPI sentiment for self-interested purposes; the informative view maintains that firms provide information to avoid misvaluation of the firm's stock.

I contend that if the opportunistic view holds, then managers are more likely to exploit NPI sentiment when they are likely to extract greater benefit from short-term stock price increases. These benefits may, for instance, derive from holdings of exercisable options (Leone et al., 2006; Souder & Bromiley, 2017) or insider trading activities (Billings et al., 2015; Choi & Kim, 2017). In contrast, managers' intentions to correct NPI sentiment should be unrelated to their holdings of exercisable options or insider trading if the informative view prevails.

I conduct a series of tests to examine the role of managerial incentives in explaining my main results. I first generate a measure (***StockOptions***) of the accumulated value of exercisable options held by directors and officers at the end of the previous fiscal year (Souder & Bromiley, 2017), scaled by the market value of equity. I then use the industry-median values to identify sub-samples representing cases of high and low senior management option holdings and re-estimate model 2 for each sub-sample. As shown in columns I and II of Table 8, ***NPISent*** has a positive and significant coefficient in both high ($\beta = 0.065$, $p = 0.001$) and low senior management option holdings subsamples ($\beta = 0.028$, $p = 0.010$). The coefficient of ***NPISent*** is higher in the subsample of high senior management option holdings compared with their counterparts ($p = 0.091$), indicating that firms with senior management option holdings above the industry median respond more strongly to NPI sentiment when making guidance decisions.

Similar inferences are obtained when I conduct equivalent tests based on the level of insider trading among senior management during the 15 days after the guidance issuance date (***InsiderTradePost***) (Billings et al., 2015). The results in Column I of Table A2 show ***NPISent*** has a positive and significant coefficient in the subsample of high (above industry-median) insider trading incentives ($\beta = 0.059$, $p < 0.001$). The coefficient of ***NPISent*** is positive but insignificant ($\beta = 0.019$, $p = 0.136$) in the subsample of low (below industry-median) insider trading incentives (as shown in Column II of Table A2). The results suggest that managers with stronger insider trading incentives are more inclined to reference NPI sentiment when issuing positive guidance than their counterparts ($p = 0.038$). Both sets of results support the opportunistic view of guidance behaviour.

Table 8 Additional analyses: Regressions on sub-samples defined by insider option holdings and institutional shareholding concentration

	Dependent variable = *GuidePos*							
	Distinguishing the opportunistic and informative views				Do StockTwits capture NPI-specific sentiment?			
	Column I: High option holdings		Column II: Low option holdings		Column III: High institutional concentration		Column IV: Low institutional concentration	
Variable	**coef.**	**t-stat.**	**coef.**	**t-stat**	**coef.**	**t-stat.**	**coef.**	**t-stat.**
NPISent	0.065	3.351***	0.028	2.577**	0.024	2.022**	0.067	3.949***
NewsSent	0.009	2.603***	0.015	6.737***	0.011	5.400***	0.019	4.438***
TweetCount	–0.013	−2.924**	–0.004	−1.651*	–0.005	–2.115**	0.002	0.298
NewsCount	0.004	2.833***	0.009	9.006***	0.004	4.066***	0.027	10.147***
MktSent	0.430	1.617	0.052	0.641	0.007	3.058***	–0.026	−0.308
IdioVol	–0.753	−3.107***	–0.504	−4.015***	–0.664	–3.782***	–0.479	−3.196***
MultiGuide	–0.434	−6.307***	–0.462	−12.547***	–0.430	–10.272***	–0.359	−7.409***
GuidePrior	–0.112	−1.297	0.092	1.680*	–0.066	–1.027	0.119	1.774*
InsideTrade	0.055	1.736	0.086	4.908***	0.070	2.607***	0.098	5.228***
InsideTradePost	0.071	1.395	–0.006	–0.206	0.075	1.684	–0.015	−0.473
VIX	0.316	2.530**	0.010	0.286	0.009	2.650***	0.015	0.418
Surprise	37.192	4.464***	21.914	4.929***	31.443	4.763***	20.662	4.084***
Loss	–0.420	−2.589**	–0.441	–5.682***	–0.527	–4.399***	–0.417	−4.667***
PriorRet	–0.033	−0.250	0.073	1.057	0.246	2.664***	–0.123	−1.460
MarketCap	0.002	1.092	–0.005	−3.401***	–0.000	−0.061	–0.009	−1.269
AnalystCov	0.000	0.030	0.003	1.055	0.001	0.201	–0.007	−1.310
MeetBeat	0.692	6.337***	0.704	11.233***	0.662	8.774***	0.758	9.426***
LitiRisk	15.315	4.321***	13.000	6.473***	11.845	5.538***	7.031	2.479**
AnalystDisp	–0.745	–0.729	–1.444	−2.511**	–0.926	−1.262	–1.107	−1.580
Industry FE	Yes		Yes		Yes		Yes	
Quarter FE	Yes		Yes		Yes		Yes	
Observations	6,031		7,636		12,361		11,930	
LR chi^2	423.775		718.965		614.283		832.892	

Prob > chi^2	0.000	0.000	0.000	0.000
Pseudo R^2	0.062	0.063	0.043	0.076
Chi^2 for difference in coefficient of ***NPISent*** between subsamples		2.760*		4.312**

Note: Columns I and II of Table 8 present results from regressing the incidence of positive guidance on NPI sentiment and controls for sub-samples of high and low senior management option holdings, respectively. Sub-samples of high and low senior management option holdings are determined by the industry-median values of the accumulated value of exercisable options held by directors and officers at the end of the previous fiscal year, scaled by the market value of equity (*StockOptions*).Columns III and IV of Table 8 present results from regressing the incidence of positive guidance on NPI sentiment and controls for sub-samples of high and low institutional ownership concentration, respectively. Sub-samples of high and low institutional ownership concentration are determined by the industry-quarter medians of the aggregate percentage stockholdings by the largest five institutional investors (*InstConcent*). *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the net positive news, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing a firm (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan Consumer Sentiment Index. Other variables are defined in Appendix B. Industry-fixed effects based on Fama–French 48 classification and fiscal quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

3.6.3 NPI Sentiment and the Issuance of Standalone Guidance

All analyses so far have focused on bundled guidance (i.e., guidance issued during or soon after earning announcements). In this section, I examine standalone guidance. Standalone guidance differs from bundled guidance in two major ways. First, unlike bundled guidance regulated at earnings announcement, standalone guidance often accompanies non-routine corporate events such as mergers and acquisitions (Milian, 2018) or economy shocks associated with the recent Covid-19 crisis, which may induce their own specific dominant guidance incentives. Importantly, in many such cases, guidance incentives may be less sensitive to NPI sentiment and rather reflect imperatives dictated by the nature of the news to be disclosed. For example, takeover bidders in stock-for-stock mergers issue pessimistic guidance to manage down analyst earnings forecasts prior to earnings releases (He et al., 2020). Such guidance issuance benefits bidders by increasing their own stock prices and saving on acquisition costs; Also, during the recent Covid-19 crisis, some firms provide earnings warnings to enable investors to understand company expectations of the possible impacts of Covid-19 on the business before earnings announcements (Brennan et al., 2022). Second, managers prefer optimistic initial guidance and then walk-down guidance to management down investors' expectations near the earnings announcement date. For example, they can withdraw the initial guidance that is no longer valid or revise the initial guidance down (Marshall & Skinner, 2022). Not in the context of NPIs, there is considerable evidence of guidance walk-down. In particular, managers use negative guidance after the initial guidance to dampen analysts' earnings expectations to avoid a negative surprise at earnings announcement (Baik & Jiang, 2006; Feng & McVay, 2010). Consequently, I expect that the association between NPI sentiment and the likelihood that guidance issued is positive will be weaker in cases of standalone guidance.

To test this contention, I obtain 17,736 firm-quarter observations (1,392 distinct firms) from May 2008 to January 2017, in which firms issued standalone guidance (3,979 positive standalone guidance). Table A3 and Table A4 presents the descriptive statistics and correlations for key variables. In contrast to the bundled guidance sample, the mean NPI sentiment prior to positive standalone guidance (mean = 0.513) is lower ($p < 0.001$) than in cases preceding negative standalone guidance (mean = 0.783). I then re-estimate model 2 and find that, in contrast to the tests using bundled guidance (Table A5), the coefficient for ***NPISent*** is *negative* and marginally significant ($\beta = -0.035$, $p = 0.071$). To dig into this issue further, I re-run the analysts for standalone guidance by partitioning the sample based why the managers disclosed an initial positive bundled guidance. The untabulated results show that, the coefficient for ***NPISent*** is significantly more negative (p = 0.075) in the subsample with initial positive bundled guidance (β = -0.134, p = 0.011) than the subsample with initial neutral or negative bundled guidance (β = -0.012, p = 0.722), as expected.

3.6.4 How Much Space Does NPI Sentiment Occupy in Capital Markets?

In this section, we assess the space for NPI sentiment to affect managers' guidance decisions by partitioning the sample according to the likely preponderance of these investors in the share register. Because NPI data are unavailable, we use the firms' the percentage of institutional holdings (***Pih***) as inverse proxies of the size of the NPI investor base. We first classify observations into high and low percentage of institutional holdings subsamples at the industry-quarter quartiles and re-estimate model 1 for each sub-sample. We expect NPI reaction to positive guidance is more sensitive to NPI sentiment when NPI ownership is high. The results (columns I and II of Table A6 Panel A) show that the coefficient on the interaction between ***NPISent*** and ***GuidePos*** is significantly more positive ($p = 0.056$) in the lower quartile of institutional holdings subsample ($\beta = 0.024$, $p = 0.004$) than in the upper quartile of institutional holdings subsample (β = 0.016, p = 0.424), as expected. We then re-estimate model 2 for each sub-samples. We expect managers' guidance decisions to be more sensitive to NPI sentiment when NPI ownership is high because of a higher NPI response to positive guidance news. The results (columns I and II of Table A6 Panel B) show that the coefficient on ***NPISent*** is significantly more positive (p = 0.018) in the lower quartile of institutional holdings subsample (β = 0.071, p < 0.001) than in the upper quartile of institutional holdings subsample (β = 0.017, p = 0.350), as expected. Taken together, the results show that the space for NPI sentiment to affect managers' guidance decisions is depended on the percentage of NPI ownership and reinforces the opportunistic explanation of our findings.

Because institutional investors differ in their behaviour and incentives, certain groups of institutions influence stock price in a manner different from the influence of overall institutional ownership (An and Zhang 2013). Transient institutional owners hold small stakes in numerous firms and trade frequently in and out of stocks, heavily basing their trades on current earnings (An and Zhang 2013). The other two types of institutional owners are "dedicated" and "quasi-indexers", who provide long-term, stable ownership to firms because they are geared toward longer-term dividend income or capital appreciation. Dedicated institutions are characterized by large average investments in portfolio firms and extremely low turnover, consistent with a "relationship investing" role and a commitment to provide long-term patient capital (An and Zhang 2013). Quasi-indexers are also characterized by low turnover, but they tend to have diversified holdings, consistent with a passive, buy-and-hold strategy of investing portfolio funds in a broad set of firms (Bushee 2001).

We expect that how NPI sentiment affect managers' guidance decisions is more sensitive to transient institutional ownership than dedicated and quasi-indexers for two reasons First, transient institutional investors tend to be short-term focused and are interested in the firm's stock is based on forward-looking information centred on near-term forecasts and news events that present speculative trading opportunities (Bushee 2005).Their trading sensitivity to forward-looking earnings information creates incentives for managers to provide positive earnings guidance to boost the stock price opportunistically. Conversely, because of the longer investment horizons of dedicated and quasi-indexers, they should be less focused on near-term earnings and do not trade actively for short-term profits. Even if these investors have near-term forecasts, it may not be optimal for them to trade on it (Ke and Petroni 2004). Since dedicated and quasi-indexer investors are largely insensitive to short-term earnings news, their presence is unlikely to influence managers' decision to positive earnings guidance. Second, because transient institutional investors have similar investment horizon to NPIs—both groups turn their portfolios over quickly, NPIs mimic transient institutions more than dedicated and quasi indexers. As such, the trading activities of transient institutional investors may amplify sentiment-related mispricing of NPIs, which creates incentives for managers to opportunistically provide positive earnings guidance when NPI sentiment is high.

We classify observations into high and low percentage of transient institutional holdings (***PihTra***) subsamples at the industry-quarter median and re-estimate model 1 for each sub-sample. We expect NPI reaction to positive guidance is more sensitive to NPI sentiment when transient institutional investor ownership is high. The results (columns III and IV of Table A6 Panel A) show that the coefficient on the interaction between ***NPISent*** and ***GuidePos*** is significantly more positive (p = 0.017) in the upper quartile of transient institutional holdings subsample (β = 0.068, p = 0.049) than in the lower quartile of transient institutional holdings subsample (β = -0.009, p = 0751), as expected. We then re-estimate model 2 for each sub-samples. We expect managers' guidance decisions to be more sensitive to NPI sentiment when transient institutional holdings is high because of a higher NPI response to positive guidance news. The results (columns III and IV of Table A6 Panel B) show that the coefficient on ***NPISent*** is significant more positive (p = 0.044) in the upper (β = 0.095, p = 0.001) than in the lower quartile of transient institutional holdings subsample (β = 0.074, p = 0.004). The results show that the space for NPI sentiment to affect managers' guidance decisions is depended on the percentage of transient institutional holdings.

3.7 Sensitivity Analyses

3.7.1 Do StockTwits Tweets Capture NPI-specific Sentiment Reliably?

I next assess how well the social media–based sentiment measure (***NPISent***) captures the sentiment of NPIs, by partitioning the sample according to the likely preponderance of these investors in the share register. I expect managers' guidance decisions to be more sensitive to NPI sentiment when NPI ownership is higher. Because NPI data are unavailable, I use the firms' (1) institutional ownership concentration and (2) institutional shareholdings as inverse proxies of the size of the NPI investor base. Following Buchanan et al. (2018), I define institutional ownership concentration as the aggregate percentage stockholdings by the largest five institutional investors (***InstConcent***), as identified in Thomson Reuters' 13-F database. I classify observations into high and low institutional concentration subsamples at the industry-quarter medians and re-estimate model 2 for each sub-sample. The results (columns III and IV of Table 8) show that while the coefficient on ***NPISent*** remains positive and significant for both high ($\beta = 0.024$, $p = 0.043$) and low ($\beta = 0.067$, $p < 0.001$) institutional concentration subsamples, the ***NPISent*** coefficient in the low institutional concentration subsample is significantly more positive ($p = 0.034$), as expected. The results provide support for the opportunistic explanation of the findings.

I then estimate similar tests using the proportionate institutional shareholdings (***InstHold***), measured by the number of common shares held by institutional investors divided by the total common shares outstanding. The results show that the coefficient ***NPISent*** has a positive and significant coefficient ($\beta = 0.025$, $p = 0.092$) in high (Column III in Table A2) and low ($\beta = 0.044$, $p = 0.001$) institutional shareholdings subsamples (Column IV in Table A2).

However, while the coefficient on ***NPISent*** in the low institutional shareholdings subsample is almost two times larger than that in the high institutional shareholdings sample, the difference in coefficients is not statistically significant (p = 0.348).

3.7.2 Controlling for Expectations of Economic Fundamentals

Although NPI sentiment is a distinct construct, it potentially correlates with macroeconomic fundamentals (Hribar et al., 2017). To mitigate concerns that the earlier results simply reflect uncontrolled correlations with macroeconomic conditions, I follow Lemmon and Portniaguina (2006) and Hribar et al. (2017) and adopt a two-stage approach to isolate a ‘cleaner’ measure of NPI sentiment by stripping out the part of NPI sentiment that can be explained by macroeconomic factors.[33] Such a proxy helps alleviate the concerns that any result is purely driven by a market-wide phenomenon.

Table A7 Panel A describes the macroeconomic factors for the sample of 25,600 firm-quarter observations used in the main analysis. Using the same first-stage model as these authors’, I regress ***NPISent*** on a broad set of macroeconomic variables, as well as their lagged and lead measures (Table A7 Panel B). These include market returns (***MktRet***), bond yield spread (***Def***), the yield on 3-month Treasury bills (***Yld***), GDP growth (***GDP***), personal consumption growth (***Cons***), labour income growth (***Labour***), unemployment rate (***URate***), consumer price index inflation (***CPIQ***) and consumption-to-wealth ratio (***CAY***). In doing so, I dissect NPI sentiment into (i) sentiment that is explained by macroeconomic factors and (ii) unexplained (or residual) sentiment. In the second stage, I replace NPI sentiment (***NPISent***) with residual NPI sentiment (***NPISentResid1***) and re-estimate model 2. Consistent with the main analyses, I find that the

[33] Lemmon and Portniaguina (2006) apply this approach on a market-wide sentiment measure, while Hribar et al. (2017) follow Lemmon and Portniaguina (2006) and apply it on their managerial sentiment measure.

coefficient on ***NPISentResid1*** (Table A7 Panel C) is positive and significant ($\beta = 0.026$, $p = 0.009$).

Unlike prior studies, this paper shows that macroeconomic factors only explain 2.9 percent of the variation in NPI sentiment, a stark difference from the very high R-squared of 0.80 obtained using the market-wide investor sentiment proxy (Lemmon & Portniaguina, 2006) and nearly 0.90 obtained using the firm-level managerial sentiment proxy (Hribar et al., 2017). The low R-squared obtained is consistent with the view that much of NPI sentiment is unjustified by macroeconomic fundamentals, and our NPI sentiment measure contains mostly information that macroeconomic factors cannot explain. Therefore, I do not include market-wide fundamentals in all tests at this stage.

3.7.3 Distinguishing Investor Attention and Investor Sentiment

Prior studies document the effect of investor attention (i.e., NPI and professional investors) on stock return (Liu et al., 2014) and stock market volatility (Audrino et al., 2020), which possibly affects managers' decisions to issue positive guidance and may be correlated with NPI sentiment. To investigate whether the positive association between NPI sentiment and managers' likelihood of issuing positive guidance is driven by investor attention, I adopt another two-stage residual inclusion approach. I use ***TweetCount*** (Audrino et al., 2020) and ***NewsCount*** (Liu et al., 2014) to proxy for investor attention. In the first stage, I regress ***NPISent*** on ***TweetCount***, ***NewsCount***, ***NewsSent*** and ***MktSent***. The coefficient for ***TweetCount*** (Column I Table A8) is positive and significant ($\beta = 0.177$, $p < 0.001$), while the coefficient for ***NewsCount*** is negative and significant ($\beta = -0.005$, $p < 0.001$). I then use the residuals from the first stage as a proxy for NPI sentiment (***NPISentResid2***) that is independent of attention, and re-estimate model 2. In line with the main analyses, the coefficient on ***NPISentResid2*** (Column II Table A8) is positive and significant ($\beta = 0.026$, $p = 0.008$).

3.7.4 Using Alternative Lexicons to Measure NPI Sentiment

In the main analyses, I adopt the VADER lexicon because this dictionary was purposely designed to analyse social media discussions (e.g., it recognises sentiment from emoticons and punctuations) and is among the very top in terms of accuracy (Gilbert, 2014). To test whether the results are sensitive to the choice of the lexicon, I tested four alternate candidates: (i) Renault's (2017) weighted field-specific lexicon (L1), (ii) Renault's (2017) manual field-specific lexicon (L2), (iii) Harvard-IV psychosocial dictionary (Harvard) and (iv) Loughran-McDonald dictionary (LM). The first two alternate measures were recently developed by Renault (2017) specifically for StockTwits. The latter two are commonly used in accounting and finance research. I re-estimated model 2 using each of the four alternative measures of NPI sentiment and found consistent results across all four lexicons. The results (Table A9) show that the coefficients for ***NPISent*** are each positive and significant (L1: $\beta = 0.003$, $p = 0.009$; L2: $\beta = 0.003$, $p = 0.001$; Harvard: $\beta = 0.003$, $p = 0.001$; and LM: $\beta = 0.008$, $p = 0.015$).

3.7.5 Measuring NPI Sentiment Across Different Time Windows

In the main analyses, I follow prior studies (e.g., Hurwitz, 2018; Liu et al., 2014) and measured NPI sentiment using tweets posted during the 30-day window (i.e., $[-31, -1]$) immediately prior to the earnings announcement date. To test the sensitivity of the results to the choice of measurement window, I re-estimated NPI sentiment over a range of more distant windows ($[-91, -61]$ and $[-61, -31]$) and more recent windows ($[-16, -1]$ and $[-8, -1]$) and repeated the main tests. Results in Table A10 demonstrate the robustness of the results to the use of all alternative windows except for the shortest recent window $[-8, -1]$ ($p = 0.166$), which arguably may not capture sufficient tweets to be representative of the sentiment of the broader NPIs. In addition, sentiment expressed so close to the guidance date leaves little time for managers to adjust their guidance strategy should they wish to.

3.8 Chapter Summary

Social media platforms unlock the potential to aggregate a large volume of NPI discussions and facilitate the formation of NPIs' beliefs about a firm. However, because NPIs are frequently sentiment-driven, social media platforms also provide fertile ground for managers' opportunism. The findings highlight how NPI sentiment from social media platforms allows managers who are incentivised by their self-interests to issue positive guidance for strong NPI reaction opportunistically.

I first document that the magnitude of NPI reaction to positive guidance increases with NPI sentiment. I then find that the likelihood of managers issuing positive guidance increases with NPI sentiment. In addition, I show that managers with higher incentives (e.g., exercisable options and insider trading) are more likely to issue positive guidance according to NPI sentiment, which helps eliminate the informative view. Further, managers are likely to provide positive guidance according to NPI sentiment when NPI sentiment conflicts with market sentiment, which helps eliminate the managerial sentiment view. Taken together, the findings are suggestive of managers' opportunistic exploitation of NPI sentiment.

The findings contribute to the literature on how investor sentiment influences managers' disclosure decisions. Prior studies have provided mixed evidence on this, and their evidence is based on market-level sentiment, which cannot differentiate NPIs from professional investors. This is largely because prior to the advent of online social networks, researchers lacked data concerning individuals' opinions about a firm's future prospects. The study complements this line of research by showing that NPI sentiment is not just noise but information that managers take into account when acting opportunistically in their guidance issuance. The findings should be of interest to securities regulators. In an age of rapid technological development such as social sentiment analysis tools, a rapid regulatory response to these new technologies will be critical. This requires regulators with the capabilities (e.g., skills, resources, and authority) to create NPI protection regulation effectively and implement the regulation.

While I believe that the findings provide interesting insights into managers' opportunism affecting guidance decisions and open a new avenue for future research, I recognise several potential endogeneity threats. To address these issues, I (i) investigate the impact of an exogenous shock to managers wary of false or misleading information on social media platforms, (ii) employ a 2SLS approach to isolate NPIs' anticipation of forthcoming guidance and (iii) conduct an unobservable selection analysis to mitigate omitted variables concerns. The inferences from the main results remain robust. I also recognize that a firm's own circumstance may affect managers' guidance decisions and I can only control for measurable effects. To address this, I execute several robustness tests: (iv) using a Heckman two-stage model to control for self-selection of guidance issuance, (v) employing a two-stage approach to strip out macroeconomic factors and investor attention. In addition, I recognize VADER is a general sentiment analysis tool. To address this issue, I conduct analyses using finance-specific lexicon: Renault's (2017) weighted field-specific lexicon; Renault's (2017) manual field-specific lexicon; Harvard-IV psychosocial dictionary; Loughran-McDonald dictionary. These tests broadly support a managerial opportunism explanation for the association between NPI sentiment and the issuance of positive guidance. A related issued is that how managers develop their knowledge of NPIs earnings expectation e.g., derived intuitively or in a calculated analysis from online social media sources. In this study, due of the nature of archival, it is difficult to explore whether and how managers monitor in real-time any messages about their firm and the level of discussion on StockTwits about their stock. In the future, I will consider using surveys or experiments to further explore manager's knowledge of NPI expectation. Finally, there is limitation of using stock returns after market close to measure NPI reactions. While NPIs can place orders after market hours (Aboody et al., 2018; Berkman et al., 2012), they can not easily trade after the market closes. I conduct additional tests to address this concern by partitioning the sample according to the likely preponderance of NPIs in the share register. Overall, the findings support the notion that managers infer NPI sentiment from social media platforms and use this to inform their opportunistic guidance behaviour.

Chapter 4 Crowdsourced Earnings Forecasts: Implications For Sell-side Analysts' Earnings Forecasts Strategies

4.1 Introduction

As information intermediaries, sell-side analysts gather, analyze, and produce information for the benefit of investors (Kothari et al., 2016). Investors seek analyst forecasts just before an earnings announcement as an earnings benchmark, and prefer to purchase the stock of firms that meet or beat analyst forecasts (Brown et al., 2009; Drake et al., 2012; Lawrence et al., 2017). Analysts may help managers to beat earnings benchmarks by walking down their short-term earnings forecasts (hereafter, forecasts) to be easy to beat in exchange for benefits in the form of greater investment banking revenues (Cowen et al., 2006), better access to management (Bradley et al., 2020), and better career prospects (Lourie, 2019). However, there is a risk of issuing easy-to-beat forecasts. The analyst profession thrives because of its reputation for unbiased assessments of a firm's potential (Cote, 2000; Kadous et al., 2009; Meng, 2015). When investors consider easy-to-beat analyst forecasts as biased, the costs of worsening reputation may offset the intended benefits (Jame et al., 2021; Meng, 2015). In the presence of reputational concerns, analysts prefer to approach investors' expectations and avoid standing out to investors (Bissessur & Veenman, 2016; Blasco et al., 2018). Specifically, analysts may walk down their forecasts according to how investors change their expectations about a given firm's perspectives over time. The underlying argument is that when investors have a lower expectation about a given firm's perspectives, they are more tolerant of analysts' walk-downs and likely to attribute forecast errors resulting from a walk-down to situational variables such as the difficulty of the forecasting task rather than an analyst's subjectivity (Kadous et al., 2009; Thayer, 2011). Extending Jame et al. (2021), to better assess investor predictions for a firm's future profitability, some analysts seek information from the Estimize crowdsourcing financial platform, on which investors submit their forward-looking financial estimates. I argue that because of analysts' desire to mitigate reputational costs, they are more likely to walk down their forecasts when investors lower their estimations on crowdsourcing financial platforms than they would otherwise. Also, this walk-down is expected to induce a pessimistic forecast error (i.e., the forecasted earnings are less than the actual earnings).

Crowdsourcing financial platforms connect investors with common interests, which creates a new way for them to produce and disseminate information. Both crowdsourcing and social media platforms allow individuals to share their opinions, but the two types of platform have several subtle differences. First, investment-oriented social media platforms like Seeking Alpha and StockTwits allow individuals to discuss stocks' quotes, price trajectories, and investment strategies with millions of investors and traders. These discussions are indicative of investor trading behavior in stock markets and their influence on market returns. For example, Chen et al. (2014) conduct textual analyses to extract sentiment from investors' posts on Seeking Alpha and find that investor sentiment on Seeking Alpha predicts future stock returns and earnings surprise. Adopting a similar approach, Deng et al. (2018) extract sentiment from investor posts on StockTwits and find a two-way interaction between investor sentiment on StockTwits and intraday stock return. In addition, analysts reference investor sentiment on social media platforms when making their forecasts, and their use of investor sentiment is accompanied by a strong market reaction to their forecasts (Chi et al., 2012). However, discussions on investment-oriented social media platforms are random; that is, investors can share their opinions on any stock-related topics such as product breakthroughs, mergers and acquisitions, and executive turnovers. While these opinions are useful in gathering insights into investors' trading behavior, they are ambiguous in terms of investors' projections of a specific element of future firm fundamentals (Barbier et al., 2012). The use of data from crowdsourcing financial platforms such as Estimize can fill this void.

Crowdsourcing is the practice of gathering input into a task by recruiting the support of a large number of individuals, either paid or unpaid, typically via the internet (Jame et al., 2016). Rather than tracking investors' random ideas as social media platforms do, crowdsourcing financial platforms allow analysts to tap into investor communities for the content of a specific type—in the case of Estimize, expectations of a firm's future earnings. As such, Estimize provides a more refined channel for analysts to access investors' projections of firms' prospects.

Because crowdsourcing financial platforms are a relatively recent advent, the effects of crowdsourced earnings estimates on analysts' forecasting strategies remain to be explored. The two earliest studies, of Adebambo et al. (2016) and Jame et al. (2016), aggregate earnings estimates submitted by individuals on Estimize during the first few years around the launch of Estimize in 2011,[34] and compare the consensus figures with the analyst forecasts for the same firm–quarters. They find that the Estimize consensus is less biased and more representative of investors' opinions of a firm's future profitability than are analyst forecasts. Given this, subsequent studies (Jame et al., 2016, 2021; Schafhäutle & Veenman, 2021) investigate the implications of crowdsourced earnings estimates for analyst forecasts. Schafhäutle and Veenman (2021) argue that analyst forecast bias is more salient to investors when analyst forecasts and investor earnings expectations provide contrasting earnings surprise signals (i.e., a firm meets or beats analyst forecasts but misses crowdsourcing earnings estimates) than when analyst forecasts and investor earnings expectations provide matching earnings surprise signals (i.e., a firm meets or beats both analyst forecasts and crowdsourcing earnings estimates). Consistent with their argument, they find that investors discount positive earnings surprises based on analyst forecasts in the presence of contrasting earnings surprise signals, using data

[34] The sample period is 2012–14 in Adebambo et al. (2016) and 2012–13 in Jame et al. (2016). Estimize recommends only using data after 1 January 2012.

from 2012 to 2018. Because contrasting earnings surprise signals could increase the salience of analyst forecast bias, Jame et al. (2021) argue that Estimize has a disciplining effect on analyst forecasting behavior; that is, analysts worry that when their forecasts do not match crowdsourcing earnings estimates, investors may consider them biased. Comparing firms with Estimize coverage with those without during the period 2012–15 shows that analysts (i) have smaller pessimistic forecast errors for Estimize-covered firms than for non-covered firms, and (ii) shift their following from Estimize-covered firms to non-covered firms.

Although Jame et al. (2016, 2021) document a relatively low level of pessimistic errors in analyst forecasts for Estimize-covered firms compared with non-covered firms, it is challenging to draw inferences about how crowdsourcing earnings estimates on Estimize affect analyst forecast strategies. First, by focusing on the initiation of Estimize coverage, Jame et al. (2021) cannot relate the low level of pessimistic forecast errors to analysts' fear of contradicting investor earnings expectations; that is, whether analysts walk their forecast downward more (less) when investors are bearish (bullish) about the outlook of a firm's earnings. Second, whether a firm gains Estimize coverage is hardly exogenous. Individuals on Estimize may choose firms with, for example, lower information asymmetry, high visibility, and strong monitoring mechanisms for which to make earnings estimates (Sul 2020), which raises the concern that analysts' different treatment of Estimize-covered and non-covered firms is a result of selection bias. Specifically, the level of pessimistic forecast errors for Estimize-covered firms may have been low before the initiation of Estimize coverage and thus may not be due to the effect of Estimize coverage. Consequently, it remains unclear how investors' earnings expectations derived from crowdsourced earnings estimates affect analyst forecast strategies.

The current study adds to previous studies in several ways. First, unlike the Estimize coverage used by Jame et al. (2021), I look into crowdsourced earnings estimates that allow observation of how investors' earnings expectations change over time. I use the changes in investors' earnings expectations to condition for analysts' incentives to avoid contradicting investors' earnings expectations. Another reason I extend the research from Estimize coverage to crowdsourced earnings estimates is that the classification of Estimize-covered and non-covered firms is becoming less useful as fewer firms remain non-covered. Every year more firms gain coverage from Estimize; thus we are rapidly approaching full coverage of the 3,000 most liquid firms on US exchanges, compared with fewer than 1,000 in 2012.[35] Analysts, who no longer have the choice to treat Estimize-covered and non-covered firms differently, are likely to examine crowdsourced earnings estimates to guide their forecast strategies. Further, noting that Jame et al. (2021) examines the level of pessimistic forecast errors, I also investigate how analysts revise their forecasts after observing changes in investors' earnings expectations to dismiss the possibility that differences in the level of pessimistic forecast errors between Estimize-covered and non-covered firms are merely a result of selection bias.

I assess the effect of investors' earnings expectations on analyst forecasting strategies by examining analysts' forecast revisions subsequent to changes in investors' earnings expectations. The analyses are guided by two hypotheses. The first hypothesis (H1) is that analysts' forecast revisions are positively correlated with recent changes in investors' earnings expectations. The second hypothesis (H2) suggests that the probability of analysts' forecast errors being pessimistic is greater when analysts' prior forecast revisions follow downward revisions in investor earnings expectations. I start with H1. I estimate OLS regressions using quarterly forecasts issued during the fiscal quarter, where the dependent variable is analyst

[35] See https://www.estimize.com/community

forecast revision; that is, the difference between the analyst EPS forecast consensus [−30, −1] days before actual earnings announcement for firm *j* in quarter *t*, and the analyst EPS forecast consensus [−60, −31] days before actual earnings announcement for firm *j* in quarter *t*, deflated by the stock price at the end of the previous quarter. The main explanatory variable is the recent change in investors' earnings expectations, measured as the difference between the Estimize EPS forecast consensus [−60, −31] days before actual earnings announcement for firm *j* in quarter *t* and the Estimize EPS forecast consensus [−90, −61] days before the actual earnings announcement for firm *j* in quarter *t*, deflated by the stock price at the end of the previous quarter. To calculate the Estimize EPS forecast consensus, I collect 879,015 'street earnings' estimates submitted by 70,926 participants to Estimize in the period January 2012–September 2018. I then take a straight average of all earnings estimates of a focal firm for a reporting period. All these tests include controls for firm characteristics and analyst characteristics that are likely to affect the association between investors' earnings expectations and analyst forecast decisions (Jame et al., 2016, 2021). The resulting coefficient for the change in Estimize EPS forecast consensus is positive and significant, suggesting that analysts revise their forecasts in line with changes in investors' earnings expectations. I then analyze H2 using logistic regressions to predict the occurrence of pessimistic forecast errors from downward analyst revisions following changes in investors' earnings expectations. After controlling for firm and analyst attributes, I find consistent evidence of a significant positive association between downward analyst revision following a downgrade in investors' earnings expectations and the likelihood of issuing pessimistic forecasts (pessimistic forecast errors). Taken together, these findings are consistent with analysts opportunistically walking down their forecasts following lower investors' earnings expectations, to mitigate reputational penalty.

I perform a series of additional tests to shed more light on the main findings. First, I use individual analyst forecast data to verify that a downward revision in analyst forecast consensus following decreases in investors' earnings expectations is at least partially due to a walk-down at the individual analyst level. I then test how the effect of changes in investors' earnings expectations on analyst forecast revision is likely to induce pessimistic errors at the individual analyst level. Consistent with the main analyses, I find that individual analysts are likely to revise down their forecasts after investors downgrade their expectations for a firm, and these analysts are likely to issue pessimistic forecasts. I verify the argument that investors' earnings expectations impose a disciplining effect on analyst forecast walk-down, using a difference-in-difference model, which compares changes in analyst forecast errors in cases when analysts are more well informed about investors' earnings expectation (e.g., firms with Estimize coverage) and when analysts are less knowledgeable about investors' earnings expectation (e.g., firms without Estimize coverage) (Jame et al., 2021). In support of the argument, I find that analyst forecast errors are less pessimistic for firms with Estimize coverage than for those without. To address the alternative explanation that analysts are learning from investors' earnings expectations (as opposed to analysts being disciplined by investors) and rely on investors' earnings expectations to adjust their forecasts, I adopt another difference-in-difference model to examine the proportion of public information (e.g., investors' earnings expectations) of the total information that analysts use in making their forecasts for firms with (and without) Estimize coverage. In contrast to the alternative explanation, the findings show that the proportion of public to total information in analyst forecasts is lower for firms with Estimize coverage than for those without, suggesting that analysts rely less on public information when forecasting for firms with Estimize coverage.

The study contributes to the literature in three ways. First, it contributes to the emerging literature on the role of crowdsourcing platforms in capital markets. Recent research provides evidence that crowdsourced forecasts can be incrementally useful in predicting firms' future earnings beyond those of analyst forecasts (Adebambo & Bliss, 2015; Jame et al., 2016), and investors take crowdsourced forecasts into account when pricing earnings news (Schafhäutle & Veenman, 2021). I find a positive association between changes in crowdsourced earnings estimates and analyst forecast revision, suggesting crowdsourced forecasts also provide analysts with valuable information that affects their forecast decisions.

Second, this study contributes to our understanding of market forces that constrain analysts' conflicts of interest. Previous research shows that analysts have reputational considerations that potentially constrain them in biasing their forecasts (Fang & Yasuda, 2009; Jame et al., 2021). As investors tend to scrutinize analyst incentives when analyst forecasts are different from their own earnings expectations, analysts are likely to consider investor expectations in providing their forecasts. Building on those findings, I argue that analysts are likely to consider investor earnings expectations as reflected by crowdsourced forecasts, to opportunistically bias their results. My findings confirm the role of reputational considerations in disciplining analysts and show that analysts are likely to opportunistically issue pessimistic forecasts following a downgrade in investors' earnings expectations to mitigate the risk of losing reputation.

Third, this study has implications for investors who incorporate analyst forecasts into their investment decisions. On 30 August 2010, the SEC published for investors a description of analysts' conflicts of interest to enable investors to recognize such conflicts (SEC, 2010). The arrival of Estimize provides a means for investors to detect analysts' conflicts of interest. By benchmarking analyst forecasts with crowdsourcing earnings expectations, investors may assess the extent to which analysts opportunistically bias their forecasts. They can select which analyst forecast to rely on when faced with multiple forecasts from different analysts.

The chapter is organized as follows. Chapter 4.2 reviews the related literature, and Chapter 4.3 develops the hypotheses. Chapter 4.4 discusses the sample selection and research design.

Chapter 4.5 presents the results of hypothesis testing. Chapter 4.6 presents the additional analyses, followed by sensitivity analyses in Chapter 4.7. The last section concludes the paper.

4.2 Literature Review

4.2.1 Meeting or Beating Analyst Earnings Forecasts and Market Response

Rewards for a firm's managers, in terms of both career opportunities (Sul, 2020) and compensation benefits (Tahir et al., 2019), depend on them achieving specific earnings thresholds on their watch—usually, analyst forecast consensus (Carvajal et al., 2017; Herrmann et al., 2011; Jiang, 2008). Meeting or beating analyst forecast consensus has positive market consequences. Firms that consistently report earnings higher than the analyst forecast consensus attract investors (especially NPIs) to buy shares (Shanthikumar, 2012), receive a price premium (Carvajal et al., 2017), and can negotiate for lower cost of debt (Jiang, 2008)

The market appears to strongly respond to earnings announcements around analyst forecast consensus, thus creating incentives to meet or beat this threshold. Firms offering greater incentives, primarily resulting from their ownership structure, are more likely to just meet or beat analyst forecast consensus (Han et al., 2014; Koh, 2007). Managers with high equity incentives such as stock option compensation are more likely to just meet or beat analyst forecast consensus (Tahir et al., 2019), and to sell stocks subsequent to just meeting or beating (Kraft et al., 2014). Managers' intentions to exceed the analyst forecast consensus is also affected by the behavior of peer firms in the same industry (Bratten et al., 2016; Du & Shen, 2018). Specifically, when industry leaders (large firms that are the first to announce earnings) meet or beat analyst forecast consensus, followers perceive that earnings news of the leader will affect investors' earnings expectations for their firms, and attempt to meet analyst forecast consensus (Bratten et al., 2016).

Managers place great importance on exceeding the analyst forecast consensus; hence, they use their influence to guide analyst forecasts downward to improve their chances of meeting or beating these forecasts when earnings are announced. After an analyst makes their initial forecast, the firm's manager—who is in possession of private information about the firm's earnings—issues (possibly biased) public earnings guidance (Kross et al., 2011) and communicates an additional (possibly biased) private earnings signal to the analyst (Versano & Trueman, 2017). The manager's goal in these public and private communications is to guide the analyst to revise their initial forecast down, especially when the initial forecast is optimistic (Versano & Trueman, 2017). Managers of firms that have consistently met or beaten the analyst forecast consensus provide more 'bad news' earnings guidance than do firms with no record of exceeding analyst forecast consensus (Kross et al., 2011; Versano & Trueman, 2017). In response to bad news earnings guidance, analysts lower their initial forecasts and issue final meetable or beatable forecasts wishing to curry favor with management (Cotter et al., 2006; Cowen et al., 2006; Feng & McVay, 2010).

4.2.2 Analysts Curry Favor With Managers

Managers hope to have beatable earnings benchmarks, and thus often create incentives for analysts to issue forecasts lower than the earnings figure in managers' minds (Bradley et al., 2020; Cowen et al., 2006; Lourie, 2019). It is common practice for analysts to issue slightly optimistic forecasts, which gives room for analysts to walk down their forecasts to a level that firms can beat at the official earnings announcement (Jame et al., 2016, 2021). The walk-down to beatable analyst forecasts is most pronounced when managers have an incentive to sell stocks after earnings announcements on the firm's behalf via new equity issuance (Feng & McVay, 2010) or from their personal accounts, through option exercises and stock sales (Contreras & Marcet, 2021). Analysts who walk down their forecasts to be beatable are likely to be rewarded by managers with inside information (Bradley et al., 2020; Feng & McVay, 2010; Ke & Yu, 2006) or future employment opportunities (Horton et al., 2017; Lourie, 2019).

While analysts benefit by walking down their forecasts, they face a potential cost to their reputations and long-term careers. In general, analysts are cautious about their reputations. The forecasts produced by analysts with strong reputations often generate increased trade activity and underwriting business for their brokerage firms (Roger, 2018; Wang, 2009). If an analyst's reputation suffers, this will affect their influence on stock markets directly connected to their bonuses and promotion (Chang & Choi, 2017; Wang, 2009). As analysts build their reputations by issuing unbiased and accurate forecasts, reputation can serve as a constraint on analyst forecast bias. Consistent with the disciplining effect of reputation, studies (Bradley et al., 2012; Fang & Yasuda, 2009) find that analysts who issue biased forecasts tradeoff the access to inside information and future career opportunities against any possible reputation loss. Analysts who are rated as "All-Stars" with strong reputations are unlikely to issue biased forecasts for relatively greater losses to them than to those who are less well-known (Bradley et al., 2012; Fang & Yasuda, 2009). Also, since institutional investors possess more information about firms, they are more capable of detecting bias than are NPIs. Also, institutional investors are well known, and to protect their reputations, they may not want analysts to deviate forecasts too much from actual earnings (Ljungqvist et al., 2007). In addition to concerns about personal reputation, analysts aim to protect their employer's reputation (i.e., a bank or a broker), and this prevents analysts from biasing their forecasts too obviously. Analysts employed by high-reputation banks are more likely to be penalized for forecast bias as a result of more monitoring than are analysts from low-reputation banks (Altınkılıç et al., 2019).

4.3 Hypothesis Development

Broadly, there are two ways in which investor earnings expectations can affect analyst forecast decisions. First, the psychology literature (Bordalo et al., 2020; Fiske & Taylor, 2013; Parr & Friston, 2019) suggests individuals direct their attention to salient signals, especially those contrasting with their own beliefs. The contrast effect comes from an unconscious bias that arises when two options are judged in comparison to one another instead of being assessed individually (Bordalo et al., 2020; Fiske & Taylor, 2013). The contrast effect can occur when the perception of currently viewed stimuli is modulated by previously viewed stimuli: a person may compare an object they are looking at with one they saw before (Fiske & Taylor, 2013). The very different object compared with yesterday's generates a surprise (Bordalo et al., 2020). For example, positive earnings news today generates a larger surprise when yesterday's earnings news was negative, than it otherwise would. This surprise translate into a higher investor response (i.e., stock return) to today's earnings news (Hartzmark & Shue, 2018). Regarding analyst forecasts, the contrast effect occurs when investors compare analyst forecasts with their previous earnings expectations (Schafhäutle & Veenman, 2021). The surprise is relatively large and salient for investors when analyst forecasts are sufficiently far from investors' previous earnings expectations. Investors pay much attention to the surprise, which may lead them to anticipate bias (if any) in analyst forecasts. To elaborate, investors generally pay more attention to evaluating an analyst's incentives to curry favor with management when the analyst's forecast bias is visible and easy to identify, such as when earnings exactly meet or just beat the analyst's expectations (Keung et al., 2010) and when the analyst delivers excessively volatile forecasts (Lundholm & Rogo, 2020). Investors discount earnings news accordingly, which leads to lower stock returns around the time of earnings announcements (Keung et al., 2010; Veenman & Verwijmeren, 2018). In contrast, investors are less concerned when evaluating an analyst's incentive to curry favor with management when the analyst's forecast bias is hidden and difficult to spot; for example, when firms have hard-to-forecast earnings (Grinblatt et al., 2018) and when the analyst provides consistent forecasts (Byun & Roland, 2021; Hilary & Hsu, 2013). Investors do not fully adjust earnings news for the bias, and there remains a strong stock return around the time of earnings announcements (Veenman & Verwijmeren, 2018). As such, when the surprise resulting from contrasting the analyst's forecasts relative to their previous earnings expectations is sufficiently salient to engage investor attention, investors will scrutinize analyst incentives and are likely to conclude that the analyst is biased. Specifically, Schafhäutle and Veenman (2021) find that investors' responses to earnings news are substantially attenuated in situations in which investors' previous earnings expectations and analyst forecasts lead to contrasting earnings surprise signals (i.e., firms meet or beat analyst forecasts, but miss the other). Given that analysts have reputation concerns, investor earnings expectations can affect forecast decisions by exposing and penalizing biased analysts.

Second, the motivated reasoning theory (Kunda, 1990) suggests that individuals motivated to arrive at a certain conclusion attempt to justify their conclusion to a dispassionate observer, and draw the desired conclusion only if they can muster the evidence to support it. According to the motivated reasoning theory, analysts motivated to walk down their forecasts to be easy to beat would try to construct a justification for their forecast bias, and would only do so if they could maintain the illusion of objectivity (Jame et al., 2021). Analysts are more likely to bias their forecasts for firms with higher earnings volatility or reported losses because analysts can attribute their forecast bias to the difficulty of the forecast task, instead of an intention of bias forecasts to curry favor with management (Bradshaw et al., 2016). When their forecasts are far from investors' expectations, analysts have a hard time justifying this discrepancy and choosing to reduce it (Jame et al., 2021). As such, investor earnings expectations can affect analyst forecast decisions even in the absence of reputational concerns.

Taken together, I posit that analysts consider investors' recent earnings expectations when deciding on their forecasts. While analysts are motived to issue easy-to-beat forecasts, they prefer to issue forecasts close to investors recent earnings expectations to avoid being perceived as biased. As a result, analysts are likely to revise down their forecasts when investors recently adjust their earnings expectations down, and in doing so, they are likely to make pessimistic forecast errors. I propose the following hypotheses.

> HYPOTHESIS 1. *Analyst forecast revisions are positively correlated with recent changes in investors' earnings expectations.*

> HYPOTHESIS 2. *The probability of analysts' forecast being pessimistic is greater when analysts' prior forecast revisions follow downward revisions in investor earnings expectations.*

4.4 Sample, Data and Research Design

4.4.1 Sample and Data

My analysis focuses on quarterly analyst forecasts issued during the fiscal quarter, for two reasons. First, analysts' incentives to understate forecasts are most common in short-run forecasts (Ke & Yu, 2006; Walther & Willis, 2013). Motivations to understate forecasts are rooted in analysts' relationships with managers. Pessimistic forecasts made just before an earnings announcement help managers meet or beat earnings expectations, contributing to analysts' relationship building with managers (Bissessur & Veenman, 2016). In contrast, optimistic short-run forecasts may increase the likelihood of managers missing earnings expectations (Bissessur & Veenman, 2016). Second, investors' demand for accounting information such as analyst forecasts increase before earnings announcements, as investors establish the benchmark for firm profitability (Drake et al., 2012; Lawrence et al., 2017). Also, investors perceive that analyst forecasts made close to earnings announcements have a timing advantage compared with previous analyst forecasts. In other words, by delaying their forecasts, analysts can observe other analysts' forecasts issued earlier, as well as other firm disclosures, to utilize more information as the earnings announcement date draws near (Kim et al., 2011; Shroff et al., 2014).

Table 9 details the sample selection criteria. The study period extends from January 2012 (Estimize recommends only using data after 1 January 2012) to 31 December 2018. I first collect a dataset of 'street earnings' daily consensus estimates from Estimize. These daily consensus estimates are calculated as the statistical mean of all estimates (excluding estimates flagged as unreliable). The initial sample consists of 3,896,909 observations. After excluding consensus estimates issued more than 90 days prior to the earnings announcement and consensus estimates issued after earnings are announced, the sample is reduced to 1,584,446 (38,240 distinct firm–quarters). As shown in Table 9, Estimize consensus estimates are more concentrated in the period immediately prior to an earnings announcement (i.e., days [–30, –1] and [0]) than in earlier periods (days [–90, –61] and [–60, –31]).

Next, I obtain a sample comprised of the intersection of realized quarterly 'street earnings' on the I/B/E/S Actuals file and quarterly analyst forecast data (Forecast Period Indicator = 6) from the I/B/E/S Detail file (1,210,847 observations). After excluding forecasts issued more than 90 days prior to the earnings announcement, those issued after earnings are announced, and those that deviate from those of the majority of analysts, the sample is reduced to 1,024,545. After restricting the sample to the most recent forecast by an analyst in each 30-day forecast window, the sample is further reduced to 951,068 (112,610 distinct firm–quarters), Similarly to Estimize data, analyst forecasts are more concentrated in the period immediately prior to earnings announcements (i.e., days [–30, –1] and [0]) than in earlier periods (days [–90, –61] and [–60, –31]).

I then create an Estimize–I/B/E/S matched sample by requiring that (1) an Estimize firm–quarter includes at least one I/B/E/S forecast, and (2) Estimize and I/B/E/S report actual EPS that matches to two decimal places. After accommodating these data requirements, and those relating to control variables acquired from Compustat and the Center for Research in Security Prices (CRSP), the final Estimize–I/B/E/S matched sample includes 6,510 firm–quarter observations for 1,041 firms.

Table 9 Sample selection—investor earnings expectations and analyst forecasts

Sample Selection	No. of observations	No. of firm-quarters	No. of distinct firms
Estimize Daily Consensus (1 Jan 2012 to 31 Dec 2018)	3,896,909	40,466	2729
Less: consensus made outside [−90, 0] period prior to the earnings announcement	(2,311,463)		(0)
Number of observations remaining	1,585,446	38,240	2729
[−90,−61]	414,321	15,120	1900
[−60, −31]	478,631	18,139	2102
[–30, –1]	655,533	34,406	2669
[0]	36,961	36,961	2729
Keep one obs. per firm quarter	(1,548,485)	(0)	(0)
Number of observations remaining	38,240	38,240	2729
I/B/E/S (FPI=6) Forecasts Consensus (1 Jan 2012 to 31 Dec 2018)	1,210,847	118,602	7531
Less: excluded forecasts	(62,974)	(1765)	(115)
Less: forecast made outside [−90, 0] period prior to earnings announcement	(123,328)	(4227)	(113)
Number of observations remaining	1,024,545	112,610	7303
Keep the most recent forecast by analyst per window	(73,477)	(0)	(0)
Number of observations remaining	951,068	112,610	7303
[−90,−61]	269,545	79304	6679
[−60, −31]	210,079	74,849	6730
[−30, −1]	450,868	90,859	6802
[0]	20,576	15,245	3981
Keep one obs. per firm quarter	(838,458)	(0)	(0)
Number of observations remaining	112,610	112,610	7303
Estimize and I/B/E/S merged	26,418	26,418	2065
Less: Estimize and I/B/E/S report actual EPS does not match to two decimal places	(4262)	(4262)	(113)
Number of observations remaining	22,156	22,156	1952
Less: observations with missing control variables			
Compustat – fundamentals	(2274)	(2274)	(208)
CRSP – daily stock	(33)	(33)	(4)
Number of observations remaining	19,849	19,849	1740
[−90,−61]	7612	7612	1175
[−60, −31]	9212	9212	1317

[−30, −1]	18,228	18,228	1709
[0]	1526	1526	865
Total useable observations	6510	6510	1041

Note: Estimize and I/B/E/S data are merged using Cusip, Year−quarter, and Actual announcement date.

4.4.2 The Impact of Changes in Investors' Earnings Expectations on Analyst Forecast Revision

H1 investigates how recent changes in investors' earnings expectations affect analyst forecast revisions. Following the literature (Kim & Song, 2015; Zhang, 2006), I measure analyst forecast revision as the difference between the analyst forecast consensus issued $[-30, -1]$ days before the actual earnings announcement for firm *j* in quarter *t*, and the analyst forecast consensus issued [−60, −31] days before the actual earnings announcement, deflated by the stock price at the end of the previous quarter:

$$Revision_{j,t} \quad = \quad \frac{Consensus[-30,-1]_{j,t} - Consensus[-60,-31]_{j,t}}{StockPrice_{j,t}}$$

where *Consensus[−60, −31]* (or *Consensus[−30, −1]*) is measured as the average of all analysts' forecasts issued in the [−60, −31] (or [−30, −1]) days before the actual earnings announcement.

In a similar vein, I measure changes in investors' earnings expectations as the difference between the Estimize estimate consensus issued in the [−60, −31] days before the actual earnings announcement for firm *j* in quarter *t* and the Estimize estimates consensus issued in the [−90, −61] days before the actual earnings announcement, deflated by the stock price at the end of the previous quarter:

$$EstmzRev_{j,t} \quad = \quad \frac{EstmzConsensus[-60,-31]_{j,t} - EstmzConsensus[-90,-61]_{j,t}}{StockPrice_{j,t}}$$

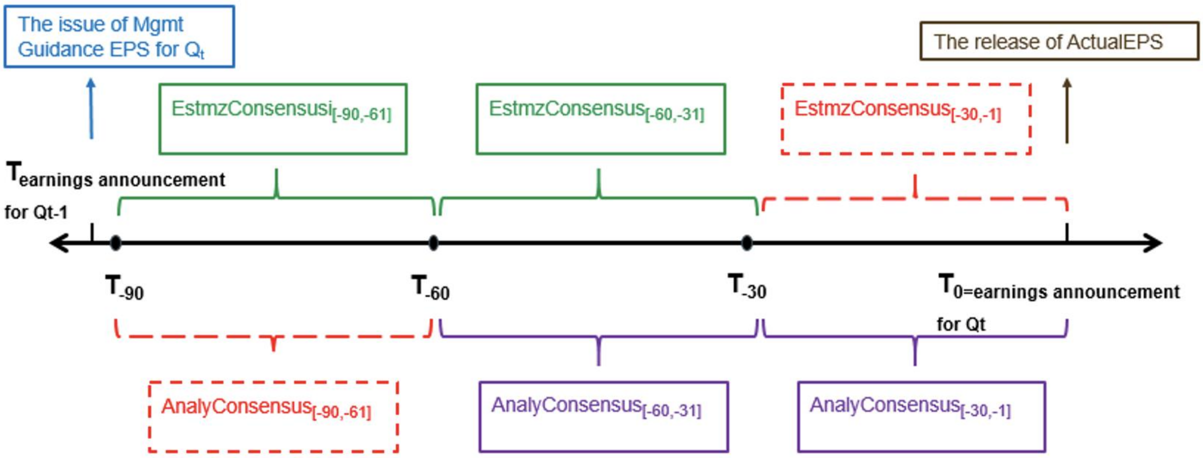
where *EstmzConsensus [−90, −61]* (or *EstmzConsensus [−60, −31]*) is measured as the average of the daily Estimize estimates consensus for [−90, −61] ([or −60, −31]) days before the actual earnings announcement.

Model 4 is estimated using OLS, with standard errors adjusted for clustering at the firm level:

$$
\begin{aligned}
Revision_{j.t} = \alpha_0 &+ \alpha_1 EstmzRev_{j,t} + \alpha_2 Follow_{j,t} + \alpha_3 Dispersion_{j,t} \\
&+ \alpha_4 FirmSize_{j,t} + \alpha_5 BtM_{i,j,t} + \alpha_6 Loss_{j,t} \\
&+ \alpha_7 Leverage_{j,t} + \alpha_8 Turnover_{j,t} + \alpha_9 RetVol_{j,t} \\
&+ \alpha_{10} Guidance_{j,t} \\
&+ \sum \alpha_j Firm_j + \sum \alpha_t YearQrt_t + \varepsilon_{i,j,t}
\end{aligned} \tag{4}
$$

In model 4, the dependent variable is the analyst forecast revision (***Revison***$_{j,t}$), The independent variable is the investors' earnings expectations (***EstmzRev***$_{j,t}$). I include controls for firm characteristics that are likely to affect analyst forecast decisions (Jame et al., 2016, 2021). Firm size (***FirmSize***$_{j,t}$) is the natural logarithm of market capitalization for firm *j* computed as the share price times the total shares outstanding as of the end of the fiscal year prior to the earnings announcement date for firm *j*. Book-to-market (***BtM***$_{j,t}$) is the book value of equity for the most recent fiscal year prior to the earnings announcement date, scaled by market capitalization as of the end of the same fiscal year for firm *j*. Average daily turnover (***Turnover***$_{j,t}$) is defined as share volume scaled by shares outstanding in the calendar year prior to the earnings announcement date for firm *j*. Loss making (***Loss***$_{j,t}$) equals 1 if reported earnings in the most recent fiscal year prior to the earnings announcement date are negative, and 0 otherwise. Leverage (***Leverage***$_{j,t}$) is the book value of long-term debt at the most recent fiscal year prior to the earnings announcement date, deflated by total assets. Volatility (***RetVol***$_{j,t}$) is measured as the standard deviation of daily returns over the calendar year prior to the earnings announcement date for firm *j*. I also control for the issuance of management guidance for firm *j* in quarter *t* (***Guidance***$_{j,t}$) and the number of analysts following firm *j* in quarter *t* (***Follow***$_{j,t}$). All variables used in analyses are defined in detail in Appendix C. Figure 5 shows the timeline and setup for the study.

Figure 5 Timeline and setup—investor earnings expectations and analyst forecasts



4.4.3 The Impact of Analysts' Prior Forecast Revisions on Pessimistic Forecast Errors

I then use model 5 to investigate whether analysts' prior forecast revisions are likely to generate pessimistic forecast errors, using a logistic regression with standard errors adjusted for clustering by firm and year–quarter:

$$\begin{aligned} Pessimism_{j,t} = \alpha_0 &+ \alpha_1 DownFollow_{j,t} + \alpha_2 DownAgainst_{j,t} \\ &+ \alpha_3 EstmzRev_{j,t} + \alpha_4 Follow_{j,t} + \alpha_5 Dispersion_{j,t} \\ &+ \alpha_6 FirmSize_{j,t} + \alpha_7 BtM_{j,t} + \alpha_8 Loss_{j,t} \\ &+ \alpha_9 Leverage_{j,t} + \alpha_{10} Turnover_{j,t} + \alpha_{11} RetVol_{j,t} \\ &+ \alpha_{12} Guidance_{j,t} \\ &+ \sum \alpha_j Firm_j + \sum \alpha_t YearQrt_t + \varepsilon_{i,j,t} \end{aligned} \tag{5}$$

In model 5, the dependent variable is a binary variable (***Pessimism**$_{j,t}$*) that equals 1 if analyst forecast errors (i.e., the actual EPS minus analyst forecast consensus issued within 30 days before the actual earnings announcement, deflated by the stock price at the end of the previous quarter) are greater than 0, and 0 otherwise. I test whether the analyst downward revision that follows a decrease in investors' earnings expectations has a different effect on cases where the analyst downward revision following an increase (or no change) in investors' earnings expectations. To this end, I construct two alternative variables: ***DownFollow**$_{j,t}$*, which equals 1 if both ***Revision*** $_{j,t}$ and ***EstmzRev*** $_{j,t}$ are less than 0, and 0 otherwise; and ***DownAgainst*** $_{j,t}$, which equals 1 if ***Revision*** $_{j,t}$ is less than 0 while ***EstmzRev*** $_{j,t}$ is greater than (or equal to) 0, and 0 otherwise. I include the same control variables as in model 4.

4.5 Empirical Results

4.5.1 Descriptive Statistics

Table 10 describes the sample of 6,510 observations from 1,041 distinct firms used in the main analysis. Table 11 presents the Pearson's and Spearman's correlations for key variables. The mean (median) for ***Revision*** $_{j,t}$ is 0.000 (0.000), suggesting analyst forecast revisions from [−60,−31] to [−30,−1] days before the earnings announcement are close to zero on average. The mean (median) for ***EstmzRev*** $_{j,t}$ is 0.002 (0.000), suggesting investors are likely to adjust their earnings expectations down from [−90,−61] to [−60,−31] days before the earnings announcement. The mean (median) values for ***Pessimism*** $_{j,t}$ is 0.7203 (1.000), suggesting that around 72% of the analyst forecast consensus issued within 30 days before an earnings announcement are pessimistic. The errors in these analyst forecast consensus (***Error*** $_{j,t}$) have a mean (median) of 0.009 (0,005). Other firm- and analyst-level variables are generally consistent with previous studies.

Table 10 Descriptive statistics—investor earnings expectations and analyst forecasts

Variable	N	Mean	SD	Min	P25	P50	P75	Max
Revision $_{j,t}$	6510	0.0000	0.0004	−0.0027	0.0000	0.0000	0.0000	0.0019
Error $_{,t}$	6510	−0.0009	0.0041	−0.0188	−0.0018	−0.0005	0.0001	0.0169
Pessimism $_{,t}$	6510	0.7203	0.4489	0.0000	0.0000	1.0000	1.0000	1.0000
PessimismMag $_{j,t}$	6510	0.0016	0.0030	0.0000	0.0000	0.0005	0.0018	0.0188
RevDown $_{j,t}$	6510	0.0788	0.2694	0.0000	0.0000	0.0000	0.0000	1.0000
DownFollow $_{j,t}$	6510	0.0384	0.1922	0.0000	0.0000	0.0000	0.0000	1.0000
DownAgainst $_{j,t}$	6510	0.0404	0.1969	0.0000	0.0000	0.0000	0.0000	1.0000
DownFollowMag $_{j,t}$	6510	0.0032	0.0236	0.0000	0.0000	0.0000	0.0000	0.2702
DownAgainstMag $_{j,t}$	6510	0.0034	0.0247	0.0000	0.0000	0.0000	0.0000	0.2702
EstmzRev $_{j,t}$	6510	−0.0002	0.0015	−0.0084	−0.0002	0.0000	0.0001	0.0050
Follow $_{,t}$	6510	4.3207	3.3015	1.0000	2.0000	3.0000	6.0000	17.0000
Dispersion $_{j,t}$	6510	0.0488	0.0653	0.0000	0.0106	0.0250	0.0576	0.3788
FirmSize $_{j,t}$	6510	8.7605	1.6300	5.1052	7.6063	8.6329	9.8507	12.5532
BtM $_{j,t}$	6510	0.3332	0.3256	−0.2338	0.1396	0.2620	0.4316	2.0587
Loss $_{j,t}$	6510	0.2605	0.4390	0.0000	0.0000	0.0000	1.0000	1.0000
Leverage $_{j,t}$	6510	0.2678	0.2075	0.0000	0.1058	0.2545	0.3848	0.9294
Turnover $_{,t}$	6510	12.4957	10.0064	2.9023	6.4855	9.3875	14.5378	61.6052
RetVol $_{j,t}$	6510	0.0214	0.0095	0.0084	0.0145	0.0190	0.0261	0.0551
Guidance $_{j,t}$	6510	0.2478	0.4318	0.0000	0.0000	0.0000	0.0000	1.0000

Note: All variables are defined in Appendix C. For presentation, *DownFollowMag* $_{j,t}$ and *DownFollowMag* $_{j,t}$ are multiplied by 100.

Table 11 Correlation matrix—investor earnings expectations and analyst forecasts

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)	***Revision$_{j,t}$***	1.00	**0.03**	−0.03	**−0.04**	**−0.74**	**−0.51**	**−0.52**	**−0.51**	**−0.52**	0.01
(2)	***Error$_{j,t}$***	0.01	1.00	**−0.78**	**−0.99**	**−0.04**	−0.03	−0.02	−0.03	−0.02	**−0.06**
(3)	***Pessimism$_{j,t}$***	−0.02	**−0.52**	1.00	**0.79**	0.03	0.02	0.03	0.02	0.02	**0.09**
(4)	***PessimismMag$_{j,t}$***	−0.02	**−0.82**	**0.34**	1.00	0.04	0.03	0.02	**0.03**	0.02	0.05
(5)	***RevDown$_{j,t}$***	**−0.57**	−0.02	0.03	0.03	1.00	**0.68**	**0.70**	**0.68**	**0.70**	−0.01
(6)	***DownFollow$_{j,t}$***	**−0.38**	−0.01	0.02	0.02	**0.68**	1.00	**−0.04**	**1.00**	**−0.04**	**−0.17**
(7)	***DownAgainst$_{j,t}$***	**−0.40**	−0.01	0.03	0.02	**0.70**	**−0.04**	1.00	**−0.04**	**1.00**	**0.15**
(8)	***DownFollowMag$_{j,t}$***	**−0.57**	−0.02	−0.01	**0.07**	**0.46**	**0.67**	−0.03	1.00	**−0.04**	**−0.17**
(9)	***DownAgainstMag$_{j,t}$***	**−0.60**	−0.03	0.00	**0.08**	**0.47**	−0.03	**0.67**	−0.02	1.00	**0.15**
(10)	***EstmzRev$_{j,t}$***	0.00	**−0.04**	**0.06**	−0.05	0.00	**−0.10**	**0.10**	−0.15	0.11	1.00
(11)	***Follow$_{j,t}$***	0.01	−0.01	−0.01	−0.02	**0.07**	**0.06**	**0.03**	−0.01	−0.02	0.00
(12)	***Dispersion$_{j,t}$***	0.02	0.00	**−0.10**	0.13	**−0.08**	**−0.06**	**−0.06**	−0.02	−0.01	**−0.05**
(13)	***FirmSize$_{j,t}$***	**0.06**	**0.07**	**0.05**	−0.25	**−0.17**	**−0.10**	**−0.13**	−0.14	−0.14	**0.07**
(14)	***BtM$_{j,t}$***	0.00	**−0.10**	**−0.04**	0.24	0.00	0.01	−0.01	0.04	0.01	**−0.07**
(15)	***Loss$_{j,t}$***	−0.02	**−0.07**	−0.01	0.22	0.01	0.01	0.01	0.06	0.05	**−0.05**
(16)	***Leverage$_{j,t}$***	0.01	0.02	**−0.06**	0.04	**−0.06**	**−0.04**	**−0.03**	−0.03	−0.01	−0.02
(17)	***Turnover$_{j,t}$***	−0.02	**−0.07**	**−0.04**	0.21	−0.02	−0.01	−0.02	0.02	0.02	**−0.11**
(18)	***RetVol$_{j,t}$***	**−0.07**	**−0.09**	**−0.07**	0.34	**0.06**	**0.04**	**0.04**	0.12	0.12	**−0.12**
(19)	***Guidance$_{j,t}$***	0.02	−0.02	**0.12**	−0.07	0.00	0.00	**0.00**	−0.03	−0.04	0.02

		(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
(1)	***Revision***$_{j,t}$	0.01	0.00	0.03	−0.01	−0.01	0.01	−0.01	−0.03	0.01
(2)	***Error***$_{j,t}$	0.02	−0.01	**0.11**	**−0.06**	**−0.11**	**0.05**	**−0.08**	**−0.13**	**0.03**
(3)	***Pessimism***$_{j,t}$	0.00	**−0.12**	**0.04**	**−0.04**	−0.01	**−0.06**	**−0.05**	**−0.06**	**0.12**
(4)	***PessimismMag***$_{j,t}$	−0.03	0.03	**−0.14**	**0.08**	**0.14**	**−0.04**	0.10	0.16	0.02
(5)	***RevDown***$_{j,t}$	**0.06**	**−0.08**	**−0.17**	0.00	0.01	**−0.07**	−0.02	**0.06**	0.00
(6)	***DownFollow***$_{j,t}$	**0.06**	**−0.05**	**−0.11**	0.01	0.01	**−0.05**	−0.01	**0.04**	0.00
(7)	***DownAgainst***$_{j,t}$	0.02	**−0.05**	**−0.13**	0.00	0.01	**−0.04**	−0.03	**0.04**	0.00
(8)	***DownFollowMag***$_{j,t}$	**0.06**	**−0.05**	**−0.11**	0.01	0.01	**−0.05**	0.00	**0.04**	0.00
(9)	***DownAgainstMag***$_{j,t}$	0.02	**−0.05**	**−0.13**	0.00	0.01	**−0.04**	−0.02	**0.04**	0.00
(10)	***EstmzRev***$_{j,t}$	−0.02	**−0.03**	0.02	**−0.04**	−0.01	−0.02	**−0.07**	**−0.05**	0.01
(11)	***Follow***$_{j,t}$	1.00	**0.29**	**0.36**	**0.05**	−0.03	**0.08**	**0.05**	**−0.09**	**−0.09**
(12)	***Dispersion***$_{j,t}$	**0.22**	1.00	**0.19**	**0.08**	0.03	**0.17**	**0.14**	**0.07**	**−0.36**
(13)	***FirmSize***$_{j,t}$	**0.32**	**0.15**	1.00	**−0.14**	**−0.34**	**0.16**	**−0.36**	**−0.61**	−0.03
(14)	***BtM***$_{j,t}$	**0.09**	**0.09**	**−0.18**	1.00	0.00	**−0.13**	0.00	0.03	**−0.05**
(15)	***Loss***$_{j,t}$	0.01	**0.07**	**−0.34**	**0.07**	1.00	**−0.04**	**0.32**	**0.52**	**−0.05**
(16)	***Leverage***$_{j,t}$	**0.04**	**0.10**	**0.09**	**−0.12**	−0.02	1.00	**−0.04**	**−0.14**	**−0.09**
(17)	***Turnover***$_{j,t}$	**0.06**	**0.13**	**−0.28**	**0.09**	**0.30**	−0.02	1.00	**0.69**	**−0.06**
(18)	***RetVol***$_{j,t}$	**−0.06**	**0.12**	**−0.57**	**0.17**	**0.54**	**−0.06**	**0.65**	1.00	**−0.07**
(19)	***Guidance***$_{j,t}$	**−0.09**	**−0.25**	**−0.04**	**−0.09**	**−0.05**	**−0.09**	**−0.08**	**−0.09**	1.00

Note: All variables are defined in Appendix C; bold typeface indicates significance at the 1% level. Pearson's correlation coefficients are shown in the lower triangle (shaded), including the diagonal; and Spearman's rank correlations appear above the diagonal.

4.5.2 The Impact of Changes in Investors' Earnings Expectations on Analyst Forecast Revision

Table 12 presents the results from tests of the association between changes in investors' earnings expectation and analyst forecast revision (H1). All reported coefficients have been standardized to allow comparison of their relative size. Column I of Table 12 shows the results of regressions based on model 4. The coefficient for $\boldsymbol{EstmzRev_{j,t}}$ is positive and significant ($\beta = 0.023$, $p = 0.006$), suggesting that a larger decrease in investors' earnings expectations increases the level of downward revision in analyst forecasts. The economic effect of changes in investors' earnings expectations on analyst forecast revision is also meaningful; a one-standard-deviation decrease in $\boldsymbol{EstmzRev_{j,t}}$ is associated with an increase in the level of downward revision in analyst forecasts of approximately 0.449 basis points (holding all covariates constant at their mean). As suggested in previous studies (Kim & Song, 2015; Kross et al., 2011), managers may issue earnings guidance to guide down analyst forecasts. To better attenuate the effect of management guidance on analyst forecast revision, I restrict the sample to observations without management guidance for year quarter t announced during the [−90, −1]-day period before the actual earnings announcement date. Column II of Table 12 shows the results of model 1 regressions using the restricted sample. The coefficient for $\boldsymbol{EstmzRev_{j,t}}$ is still positive and significant ($\beta = 0.024$, $p = 0.006$).

Table 12 The impact of changes in investor earnings expectations on analyst forecast revision

	DV = *Revision*$_{j,t}$			
	Column I: Full sample		**Column II: Exclude standalone guidance**	
Variable	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**
EstmzRev$_{j,t}$	0.023	2.730***	0.024	2.741***
Follow$_{j,t}$	−0.015	−1.446	−0.015	−1.411
Dispersion$_{,t}$	0.052	1.426	0.052	1.424
FirmSize$_{j,t}$	−0.025	−0.173	−0.030	−0.202
BtM$_{j,t}$	0.083	1.734*	0.085	1.760*
Loss$_{j,t}$	0.001	0.052	0.002	0.073
Leverage$_{j,t}$	−0.026	−0.369	−0.030	−0.394
Turnover$_{j,t}$	−0.006	−0.293	−0.006	−0.300
RetVol$_{j,t}$	−0.067	−0.766	−0.071	−0.768
Guidance$_{j,t}$	0.002	0.382	−0.001	−0.134
Regression Type	OLS		OLS	
No. of obs.	6510		6036	
Year-Quarter FE	Yes		Yes	
Firm FE	Yes		Yes	
Adjust R Square	0.192		0.183	

Note: Column I reports the regression results using all 6510 observations. Column II reports the regression results using observations without management guidance for year quarter t announced during the [−90, −1] period for actual earnings announcement date. *Revision*$_{j,t}$ is measured as analyst EPS forecast consensus [−30, −1] days before actual earnings announcement for firm j in quarter t minus analyst EPS forecast consensus [−60, −31] days before actual earnings announcement for firm j in quarter t, deflated by the stock price at the end of the previous quarter; *EstmzRev*$_{j,t}$ is measured as Estimize EPS forecast consensus [−60, −31] days before actual earnings announcement for firm j in quarter t minus Estimize EPS forecast [−90, −61] days before actual earnings announcement for firm j in quarter t, deflated by the stock price at the end of the previous quarter; *Follow*$_{j,t}$ is the number of analysts following the firm j in quarter in [−30,−1] days before the earnings announcement date; *Dispersion*$_{j,t}$ is the standard deviation in the analyst forecasts for firm j in quarter t in [−30,−1] days before the earnings announcement date; *FirmSize*$_{j,t}$ is measured as the the natural logarithm of market capitalization for firm j computed as share price times total shares outstanding as of the end of the fiscal year before the earnings announcement date for firm j; *BtM*$_{j,t}$ is the book value of equity for the most recent fiscal year before the earnings announcement date, scaled by market capitalization on as of the end of the same fiscal year for firm j; *Loss*$_{j,t}$ is a dummy for firm j, set to one if firm j generated a net loss in the most recent fiscal year before the earnings announcement date; zero otherwise; *Leverage*$_{j,t}$ is the leverage for firm j, defined as the book value of long−term debt at the most recent fiscal year before the earnings announcement date, deflated by total assets; *Turnover*$_{j,t}$ is defined as share volume scaled by shares outstanding in the calendar year before the earnings announcement date for firm j; *RetVol*$_{j,t}$ is the standard deviation of daily returns over the calendar year prior to the earnings announcement date for firm j; *Guidance*$_{j,t}$ is a binary variable equal to one if firm j issues earnings guidance for the quarter t. Firm fixed effects fiscal year-quarter fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

4.5.3 The Impact of Analysts' Prior Forecast Revisions on Pessimistic Forecast Errors

Table 13 presents the results of testing how analysts' prior forecast revisions are likely to induce pessimistic forecast errors (H2). Column I shows the results of logistic regressions based on model 5. The coefficient for $\boldsymbol{DownFollow_{j,t}}$ is positive and significant (β = 0.191, p = 0.031), suggesting that analysts who revise their forecasts down following decreased investors' earnings expectations are more likely to have pessimistic forecast errors. The marginal effect of $\boldsymbol{DownFollow_{j,t}}$ suggests that these analysts are around 8.2% more likely to end up with pessimistic forecast errors. The coefficient for $\boldsymbol{DownAgainst_{j,t}}$ is positive but insignificant (β = 0.148, p = 0.109). Similar results are obtained using only observations without management guidance for year quarter t announced during the [−90, −1]-day period prior to the actual earnings announcement date. As shown in Column II of Table 13, the coefficient for $\boldsymbol{DownFollow_{j,t}}$ is positive and marginally significant (β = 0.154, p = 0.094) but the coefficient for $\boldsymbol{DownAgainst_{j,t}}$ is insignificant (β = 0.123, p = 0.214). One concern is the fact that the difference between the coefficients for $\boldsymbol{DownFollow_{j,t}}$ and $\boldsymbol{DownAgainst_{j,t}}$ is insignificant (p = 0.733 in Column I; p = 0.804 in Column II). To address this concern, I construct alternative variables ($\boldsymbol{DownFollowMag_{j,t}}$ and $\boldsymbol{DownFollowMag_{j,t}}$) to capture the effect of the magnitude of analyst revision following changes in investors' earnings expectation on the likelihood of pessimistic forecast errors. $\boldsymbol{DownFollowMag_{j,t}}$ **(**$\boldsymbol{DownAgainstMag_{j,t}}$**)** is the magnitude of downward revision in analyst forecast consensus when $\boldsymbol{EstmzRev_{j,t}}$ is less than (greater or equal to) 0, measured as $\boldsymbol{DownFollow_{j,t}}$ ($\boldsymbol{DownAgainst_{j,t}}$) times the absolute value of $\boldsymbol{Revision_{j,t}}$. As reported in Column III of Table 13, the coefficient of $\boldsymbol{DownFollowMag_{j,t}}$ is positive and significant (β = 0.374, p = 0.026), suggesting that the greater the level of downward revision in analyst forecast consensus following a decrease in investors' earnings expectations, the greater the likelihood of pessimistic forecast errors. The marginal effect of $\boldsymbol{DownFollowMag_{j,t}}$ suggests that a one-standard-deviation increase in analysts' downward revision following decreased investors' earnings expectations increases the likelihood of ending up with pessimistic forecast errors by around 1.6%. The coefficient of ***DownAgainstMag***$_{j,t}$ is positive but insignificant ($\beta = 0.004$, $p = 0.879$). The difference between the coefficient for ***DownFollowMag***$_{j,t}$ and that for ***DownAgainstMag***$_{j,t}$ is significant ($p = 0.031$), providing support for H2. Similar results are obtained using only observations without management guidance for year quarter t announced during the [−90, −1]-day period before the actual earnings announcement date (Column IV of Table 13).

Table 13 The probability of analysts' forecast errors being pessimistic is greater when analysts' prior forecast revisions follow downward revisions in investor earnings expectations

	DV = ***Pessimism***$_{j,t}$							
	Column I: Full sample		**Column II: Exclude standalone guidance**		**Column III: Full sample**		**Column IV: Exclude standalone guidance**	
Variable	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**
DownFollow$_{j,t}$	0.191	2.152**	0.154	1.677*				
DownAgainst$_{j,t}$	0.148	1.602	0.123	1.249				
DownFollowMag$_{j,t}$					0.374	2.232**	0.379	2.237**
DownAgainstMag$_{j,t}$					0.004	0.152	0.005	0.183
EstmzRev$_{j,t}$	−0.093	−2.587**	−0.103	−2.767***	−0.007	−0.972	−0.007	−0.966
Follow$_{j,t}$	0.001	0.011	−0.013	−0.114	0.002	0.18	0.002	0.134
Dispersion$_{,t}$	−0.442	−3.087***	−0.433	−2.968***	0.007	0.275	0.006	0.236
FirmSize$_{j,t}$	−1.246	−2.072**	−1.255	−2.057**	−0.517	−3.561***	−0.460	−3.300***
BtM$_{j,t}$	−0.086	−0.773	−0.089	−0.738	−0.092	−0.952	−0.077	−0.891
Loss$_{j,t}$	0.254	1.645	0.238	1.498	0.032	1.219	0.035	1.302
Leverage$_{j,t}$	0.165	0.513	0.291	0.846	−0.003	−0.097	0.009	0.245
Turnover$_{j,t}$	0.209	1.758*	0.270	2.300**	−0.054	−1.884*	−0.051	−1.817*
RetVol$_{j,t}$	−0.020	−0.080	0.047	0.184	0.164	3.054***	0.153	3.046***
Guidance$_{j,t}$	0.002	0.013	0.214	1.227	−0.021	−1.321	−0.003	−0.225
Regression Type	Logistic		Logistic		Logistic		Logistic	
No. of obs.	5185		4741		5185		4741	
Year-Quarter FE	Yes		Yes		Yes		Yes	
Firm FE	Yes		Yes		Yes		Yes	
Cluster by firm and year-quarter	Yes		Yes		Yes		Yes	
Pseudo R Square	0.143		0.139		0.143		0.139	
	_b[DownFollow]=_b[DownAgainst]				_b[DownFollowMag]=_b[DownAgainstMag]			
	Pr > chi2 = 0.7329		Pr > chi2 = 0.8035		Pr > chi2 = 0.0309		Pr > chi2 = 0.0313	

Note: Columns I and III report the regression results using all 5,185 observations. Columns II and IV report the regression results using observations without management guidance for year quarter *t* announced during the [−90, −1]-day period before the actual earnings announcement date. $Pessmism_{j,t}$ is a binary variable equal to 1if the error of analyst EPS forecast consensus is less than 0; and 0 otherwise; $RevDown_{j,t}$ is a binary variable equal to 0 if $Revision_{j,t}$ is less than 0; and 0 otherwise; $DownFollow_{j,t}$ is a binary variable equal to 1 if both $Revision_{j,t}$ and $EstmzRev_{j,t}$ are less than 0; and 0 otherwise; $DownAgainst_{j,t}$ is a binary variable that equals 1 if $Revision_{j,t}$ is less than 0 while $EstmzRev_{j,t}$ is greater than (or equal to) 0; and 0 otherwise; $EstmzRev_{j,t}$ is measured as Estimize EPS forecast consensus [−60, −31] days before the actual earnings announcement for firm *j* in quarter *t*, minus the Estimize EPS forecast [−90, −61] days before the actual earnings announcement for firm *j* in quarter *t*, deflated by the stock price at the end of the previous quarter; $Follow_{j,t}$ is the number of analysts following firm *j* in quarter *t* [−30,−1] days before the earnings announcement date; $Dispersion_{j,t}$ is the standard deviation in the analyst forecasts for firm *j* in quarter *t* [−30,−1] days before the earnings announcement date; $FirmSize_{j,t}$ is measured as the natural logarithm of market capitalization for firm *j* computed as share price times total shares outstanding as of the end of the fiscal year before the earnings announcement date for firm *j*; $BtM_{j,t}$ is the book value of equity for the most recent fiscal year before the earnings announcement date, scaled by market capitalization on as of the end of the same fiscal year for firm *j*; $Loss_{j,t}$ is a dummy for firm *j*, set to 1 if firm *j* generated a net loss in the most recent fiscal year before the earnings announcement date; 0 otherwise; $Leverage_{j,t}$ is the leverage for firm *j*, defined as the book value of long-term debt at the most recent fiscal year before the earnings announcement date, deflated by total assets; $Turnover_{j,t}$ is defined as share volume scaled by shares outstanding in the calendar year before the earnings announcement date for firm *j*; $RetVol_{j,t}$ is the standard deviation of daily returns over the calendar year prior to the earnings announcement date for firm *j*; $Guidance_{j,t}$ is a binary variable equal to 1 if firm *j* issues earnings guidance for quarter *t*. Firm- and fiscal-year-quarter-fixed effects are included; *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

4.6 Additional Analysis

4.6.1 Individual Analyst Forecast

4.6.1.1 The Impact of Changes in Investors' Earnings Expectations on Individual Analyst Forecast Revision

In this section, I use individual analyst forecast data to understand better the effect of investors' earnings expectations on analyst forecast decisions. I begin by verifying that the evidence for downward revision in analyst forecast consensus following decreases in investors' earnings expectations is at least partially due to a walk-down at the individual analyst level. Table A11 details the sample selection criteria. I obtain an Estimize–I/B/E/S matched sample by requiring that (1) an Estimize firm–quarter includes at least one I/B/E/S forecast, and (2) Estimize and I/B/E/S report actual EPS that match to two decimal places. After accommodating these data requirements, and those relating to control variables acquired from Compustat and CRSP, the final Estimize–I/B/E/S matched sample includes 10,361 observations from 686 firms. I estimate the following OLS model:

$$\begin{aligned} Revision_{i.j,t} = \alpha_0 &+ \alpha_1 EstmzRev_{j,t} + \alpha_2 Follow_{j,t} + \alpha_3 Dispersion_{j,t} \\ &+ \alpha_4 AccScore_{i.j,t} + \alpha_5 BoldScore_{i.j,t} + \alpha_6 Bias\ Score_{i.j,t} \\ &+ \alpha_7 FirmExp_{i.j,t} + \alpha_8 IndusExp_{i.j,t} + \alpha_9 GenExp_{i,t} \\ &+ \alpha_{10} BrokerSize_{i.j,t} + \alpha_{11} FirmSize_{j,t} + \alpha_{12} BtM_{j,t} \\ &+ \alpha_{13} Loss_{j,t} + \alpha_{14} Leverage_{j,t} + \alpha_{15} Turnover_{j,t} \\ &+ \alpha_{16} RetVol_{j,t} + \alpha_{17} Guidance_{j,t} \\ &+ \sum \alpha_j Firm_j + \sum \alpha_t YearQrt_t + \varepsilon_{i,j,t} \end{aligned} \tag{6}$$

In model 6, the dependent variable is analyst *i*'s forecast revision (***Revision**$_{,j,t}$*) measured as analyst *i*'s EPS forecast [–30, –1] days before the actual earnings announcement for firm *j* in quarter *t* minus analyst *i*'s EPS forecast [–60, –31] days before the actual earnings announcement for firm *j* in quarter *t*, deflated by the stock price at the end of the previous quarter. The independent variable is the investors' earnings expectations (***EstmzRev**$_{j,t}$*).

Following the literature (Byun & Roland, 2021; Demmer et al., 2019; Kim et al., 2011), I control for analyst characteristics likely to affect analyst forecast decisions: the relative forecast accuracy for analyst *i*'s EPS forecast [–60, –31] days before the actual earnings announcement for firm *j* in quarter *t* ($\boldsymbol{AccScore_{i,j,t}}$); the relative forecast boldness for analyst *i*'s EPS forecast [–60, –31] days before the actual earnings announcement for firm *j* in quarter *t* ($\boldsymbol{BoldScore_{i,j,t}}$); the relative forecast bias for analyst *i*'s EPS forecast [–60, –31] days before the actual earnings announcement for firm *j* in quarter *t* ($\boldsymbol{BiasScore_{i,j,t}}$); the number of analysts employed by the analyst *i*'s brokerage house in the year ($\boldsymbol{BrokerSize_{i,j,t}}$); the number of quarters to date for which analyst *i* has followed firm *j* ($\boldsymbol{FirmExp_{i,j,t}}$); the number of quarters to date for which analyst *i* has followed the industry to which firm *j* belongs ($\boldsymbol{IndusExp_{i,j,t}}$); and the number of quarters to date during which analyst *i* has issued forecasts for this or any other firm ($\boldsymbol{GenlExp_{i,t}}$). I also include controls for firm characteristics as in model 4. All variables used in analyses are defined in detail in Appendix C. Table A12 describes the sample of 10,361 observations, and Table A13 presents the correlation matrix.

Column I of Table 14 shows the results of regressions based on model 6 using all 10,361 observations. All reported coefficients have been standardized to allow comparison of their relative size. The coefficient for $\boldsymbol{EstmzRev_{j,t}}$ is positive and significant ($\beta = 0.070$, $p = 0.036$), suggesting that a greater decrease in investors' earnings expectations increases the level of downward revision in individual analyst forecasts. The economic effect of changes in investors' earnings expectations on analyst forecast revision is also meaningful; a one-standard-deviation decrease in $\boldsymbol{EstmzRev_{j,t}}$ is associated with an increase in the level of downward revision in analyst forecasts, by approximately 3.486 basis points (holding all covariates constant at their mean). Column II of Table 14 shows the results of regressions based on model 6 using only observations without management guidance for year quarter *t* announced during the [–90, –1]-day period before the actual earnings announcement date. The coefficient for ***EstmzRev**$_{j,t}$* remains positive and significant (β = 0.066, *p* = 0.0390).

Table 14 The impact of changes in investor earnings expectations on analyst forecast revision at the individual analyst level

	DV = *Revision*$_{i,j,t}$			
	Column I: Full sample		**Column II: Exclude standalone guidance**	
Variable	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**
***EstmzRev**$_{j,t}$*	0.070	2.105**	0.066	2.069**
***Follow**$_{j,t}$*	0.001	0.090	0.001	0.083
***Dispersion**$_{,t}$*	−0.032	−0.664	−0.035	−0.696
***AccScore**$_{i,j,t}$*	0.004	0.351	0.006	0.485
***BoldScore**$_{i,j,t}$*	−0.068	−4.324***	−0.067	−4.164***
***BiasScore**$_{i,j,t}$*	0.237	11.330***	0.243	11.162***
***FirmExp**$_{i,j,t}$*	−0.014	−1.250	−0.014	−1.245
***IndusExp**$_{i,j,t}$*	0.026	1.323	0.026	1.267
***GenExp**$_{i,j,t}$*	−0.026	−1.530	−0.027	−1.463
***BrokerSize**$_{i,t}$*	−0.006	−0.622	−0.007	−0.745
***FirmSize**$_{j,t}$*	0.202	0.689	0.264	0.861
***BtM**$_{j,t}$*	0.238	1.362	0.258	1.408
***Loss**$_{j,t}$*	0.108	3.150***	0.086	2.587**
***Leverage**$_{j,t}$*	−0.049	−0.860	−0.031	−0.529
***Turnover**$_{j,t}$*	0.193	2.229**	0.127	1.642
***RetVol**$_{j,t}$*	−0.323	−2.663***	−0.275	−2.212**
***Guidance**$_{j,t}$*	−0.067	−1.691*	−0.017	−0.535
Regression Type	OLS		OLS	
No. of obs.	10361		9933	
Year-Quarter FE	Yes		Yes	
Firm FE	Yes		Yes	
Adjust R Square	0.126		0.100	

Note: Column I reports the regression results using all 10,361 observations. Column II reports the regression results using observations without management guidance for year quarter *t* announced during the [−90, −1]-day period before the actual earnings announcement date. *Revision*$_{,j,t}$ is measured as analyst *i*'s EPS forecast [−30, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus analyst *i*'s EPS forecast [−60, −31] days before the actual earnings announcement for firm *j* in quarter *t*, deflated by the stock price at the end of the previous quarter; *EstmzRev*$_{j,t}$ is measured as the Estimize EPS forecast consensus [−60, −31] days before the actual earnings announcement for firm *j* in quarter *t* minus the Estimize EPS forecast [−90, −61] days before the actual earnings announcement for firm *j* in quarter *t*, deflated by the stock price at the end of the previous quarter. All variables are defined in Appendix C; *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

4.6.1.2 The Effect in H1 is Likely to Generate Pessimistic Forecast Errors on Individual Analyst Forecast Revisions

I then test whether the effect of changes in investors' earnings expectations on analyst forecast revision is likely to induce pessimistic errors at the individual analyst level. I estimate the following logistic regression model:

$$
\begin{aligned}
Pessimism_{i.j,t} = \alpha_0 &+ \alpha_1 RevDown_{j,t} + \alpha_2 EstmzRev_{j,t} + \alpha_3 Follow_{j,t} \\
&+ \alpha_4 Dispersion_{j,t} + \alpha_5 AccScore_{i.j,t} + \alpha_6 BoldScore_{i.j,t} \\
&+ \alpha_7 Bias\ Score_{i.j,t} + \alpha_8 FirmExp_{i.j,t} + \alpha_9 IndusExp_{i.j,t} \\
&+ \alpha_{10} GenExp_{i,j,t} + \alpha_{11} BrokerSize_{i.j,t} + \alpha_{12} FirmSize_{j,t} \\
&+ \alpha_{13} BtM_{j,t} + \alpha_{14} Loss_{j,t} + \alpha_{15} Leverage_{j,t} \\
&+ \alpha_{16} Turnover_{j,t} + \alpha_{17} RetVol_{j,t} + \alpha_{18} Guidance_{j,t} \\
&+ \sum \alpha_j Firm_j + \sum \alpha_t YearQrt_t + \varepsilon_{i,j,t} \qquad (7)
\end{aligned}
$$

In model 7, the dependent variable is a binary variable (***Pessimism***$_{i,j,t}$) that equals 1 if the error of analyst *i*'s EPS forecast (i.e., analyst *i*'s EPS forecast [–30, –1] days before the actual earnings announcement for firm *j* in quarter *t* minus the actual EPS, deflated by the stock price at the end of the previous quarter) is greater than 0, and 0 otherwise. The independent variable is a binary (***RevDown***$_{i,j,t}$) that equals 1 if ***Revision***$_{,j,t}$ is less than 0; 0 otherwise. I include the same control variables as in model 6.

As shown in Column I of Table 15, the coefficient for ***RevDown***$_{i,j,t}$ is positive and significant ($\beta = 0.509$, $p = 0.000$), suggesting that analysts are more likely to have pessimistic forecast errors after revising down their forecasts. The marginal effect of ***RevDown***$_{i,j,t}$ suggests that analysts who revise down their forecasts are around 7.7% more likely to end up with pessimistic forecast errors. Column II of Table 15 shows the results of model 6 regressions using only observations without management guidance for year quarter *t* announced during the [−90, −1]-day period before the actual earnings announcement date. The coefficient for ***RevDown***$_{i,j,t}$ is still positive and significant ($\beta = 0.528$, $p = 0.000$). Similar to the tests of analyst forecast consensus, I construct alternative variables ***DownFollow***$_{i,j,t}$, which equals 1 if both ***Revision***$_{,j,t}$ and ***EstmzRev***$_{j,t}$ are negative and 0 otherwise; and ***DownAgainst***$_{i,j,t,}$ which equals 1 if ***Revision***$_{,j,t}$ is negative while ***EstmzRev***$_{j,t}$ is non-negative and 0 otherwise. I then re-run model 6 replacing ***RevDown***$_{i,j,t}$ with ***DownFollow***$_{i,j,t}$ and ***DownAgainst***$_{i,j,t}$. Column III of Table 15 reports the results. The coefficients for ***DownFollow***$_{i,j,t}$ ($\beta = 0.429$, $p = 0.000$) and ***DownAgainst***$_{i,j,t}$ ($\beta = 0.492$, $p = 0.000$) are both positive and significant. However, the difference between the two coefficients is not statistically significant ($p = 0.382$). Similar results are obtained using only observations without management guidance for year quarter t announced during the [−90, −1]-day period before the actual earnings announcement date. As shown in Column IV of Table 15, the coefficients for ***DownFollow***$_{i,j,t}$ ($\beta = 0.442$, $p = 0.000$) and ***DownAgainst***$_{i,j,t}$ ($\beta = 0.512$, $p = 0.000$) are again both positive and significant.

Table 15 The probability of analysts' forecast errors (at the individual analyst level) being pessimistic is greater when analysts' prior forecast revisions follow downward revisions in investor earnings expectations

	DV = *Pessimism*$_{i,j,t}$							
	Column I: Full sample		**Column II: Exclude guidance**	**standalone**	**Column III: Full sample**		**Column IV: Exclude guidance**	**standalone**
Variable	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**
RevDown$_{i,j,t}$	0.509	4.947***	0.528	4.921***				
DownFollow$_{i,j,t}$					0.429	3.730***	0.442	3.727***
DownAganist$_{i,j,t}$					0.492	4.562***	0.512	4.575***
EstmzRev$_{j,t}$	−0.025	−0.294	−0.022	−0.246	−0.046	−0.541	−0.045	−0.501
Follow$_{j,t}$	0.055	0.609	0.058	0.640	0.059	0.659	0.062	0.690
Dispersion$_{,t}$	−0.372	−2.528**	−0.351	−2.320**	−0.372	−2.535**	−0.351	−2.328**
AccScore$_{i,j,t}$	0.151	2.529**	0.162	2.611***	0.150	2.517**	0.161	2.603***
BoldScore$_{i,j,t}$	−0.086	−0.951	−0.085	−0.911	−0.083	−0.919	−0.081	−0.878
BiasScore$_{i,j,t}$	1.238	11.332***	1.276	11.358***	1.237	11.336***	1.274	11.361***
FirmExp$_{i,j,t}$	0.041	0.467	0.056	0.617	0.042	0.481	0.057	0.629
IndusExp$_{i,j,t}$	−0.177	−1.646	−0.166	−1.521	−0.177	−1.648	−0.166	−1.524
GenExp$_{i,j,t}$	0.232	2.403**	0.204	2.081**	0.232	2.401**	0.204	2.078**
BrokerSize$_{i,t}$	0.020	0.327	0.014	0.229	0.019	0.312	0.013	0.209
FirmSize$_{j,t}$	0.312	0.468	0.156	0.224	0.325	0.490	0.171	0.246
BtM$_{j,t}$	0.022	0.114	0.057	0.268	0.032	0.165	0.068	0.320
Loss$_{j,t}$	0.392	2.125**	0.384	2.017**	0.395	2.151**	0.389	2.049**
Leverage$_{j,t}$	0.315	0.716	0.467	1.011	0.317	0.721	0.468	1.014
Turnover$_{j,t}$	−0.059	−0.261	−0.059	−0.255	−0.056	−0.249	−0.055	−0.237
RetVol$_{j,t}$	0.718	2.234**	0.687	2.025**	0.709	2.211**	0.676	2.002**
Guidance$_{j,t}$	0.046	0.244	0.137	0.885	0.050	0.264	0.139	0.897
Regression Type	Logistic		Logistic		Logistic		Logistic	
No. of obs.	9239		8814		9293		8814	
Year-Quarter FE	Yes		Yes		Yes		Yes	
Firm FE	Yes		Yes		Yes		Yes	

Pseudo R Square	0.176	0.182	0.176	0.182
_b[DownFollow]=_b[DownAgainst]			Pr > chi2 = 0.3819	Pr > chi2 = 0.3620

Note: Column I reports the regression results using all 9,239 observations. Column II reports the regression results using observations without management guidance for year quarter *t* announced during the [−90, −1]-day period before the actual earnings announcement date. $Pessimism_{i,j,t}$ is a binary variable that equals 1 if $Error_{i,j,t}$ is less than 0; 0 otherwise; $RevDown_{i,j,t}$ is a binary variable that equals 1 if $Revision_{,j,t}$ is less than 0; 0 otherwise; $DownFollow_{i,j,t}$ is a binary variable that equals 1 if both $Revision_{,j,t}$ and $EstmzRev_{j,t}$ are less than 0; 0 otherwise; $DownAgainst_{i,j,t}$ is a binary variable that equals 1 if $Revision_{,j,t}$ is less than 0 while $EstmzRev_{j,t}$ is greater than (or equal to) 0; 0 otherwise. All variables are defined in Appendix C; *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

4.6.2 The Impact of Estimize Coverage on Analyst Forecast Decisions

My H1 hinges on the argument that investors' earnings expectations impose a disciplining effect on analyst forecast walk-down. If the argument holds, analysts would be more cautious to walk down their forecasts for firms with Estimize coverage than for those without, because Estimize coverage makes investors' earnings expectations more visible to analysts. To test this, I follow the difference-in-difference approach employed by Jame et al. (2021), to compare changes in analyst forecast errors for treatment and control firms around the initiation of Estimize coverage.

I define treated firms as those added to Estimize in 2012 and 2013. These firms experience significantly greater activity on the Estimize platform than do firms added in later years. Control firms are those not added to Estimize by the end of 2015. I define the pre-event periods as the three years prior to the launch of Estimize (i.e., 2009–11) and the post-event period as the five years after the launch of Estimize (i.e., 2012–18). I choose a long post-event window because it may take time for Estimize to prove its viability and influence analysts' forecasting behavior. Table A14 shows the sample selection. I estimate the following model with standard errors adjusted for clustering at the firm level:

$$
\begin{aligned}
Error90d_{j,t} = \alpha_0 &+ \alpha_1 Post_t + \alpha_2 Treat_j + \alpha_3 Post_t \times Treat_j \\
&+ \alpha_4 FirmSize_{j,t} + \alpha_5 Return_{j,t} + \alpha_6 Follow_{j,t} \\
&+ \alpha_7 BtM_{j,t} + \alpha_8 Turnover_{j,t} + \alpha_9 Dispersion_{j,t} \\
&+ \alpha_{10} Guidance_{j,t} \\
&+ \sum \alpha_j Industry + \sum \alpha_t YearQrt_t + \varepsilon_{j,t} \qquad (8)
\end{aligned}
$$

The dependent variable is the error in the analyst EPS forecast consensus, measured as analyst EPS forecast consensus [−90, 0] days before the actual earnings announcement for firm *j* in quarter *t* minus the actual EPS, deflated by the stock price at the end of the previous quarter (***Error90d$_{j,t}$***). The independent variable of interest is the interaction of ***Post*** (equal to 1 for the period 2012–18, and 0 for 2009–11) and ***Treat$_j$*** (equal to 1 for treated firms and 0 for control firms). All variables used in analyses are defined in detail in Appendix C. One concern is that any difference in analyst forecast errors may result from systematic differences in covariates between treated and non-treated firms. To alleviate this concern, I control for confounding covariates by implementing the propensity score matching (PSM) method. Specifically, I estimate a logistic regression in which the dependent variable is 1 for treated firms and 0 for control firms. Covariates include four firm characteristics (***FirmSize$_{j,t}$***, ***BtM$_{j,t}$***, ***Follow$_{,t}$***, ***Turnover$_{j,t}$***) and ***Error90d$_{j,t}$***.

Column I of Table 16 reports the results of model 8 prior to undertaking PSM. The coefficient on the interaction of ***Post*** and ***Treat$_j$*** is positive and significant ($\beta = 0.029$, $p = 0.002$), suggesting a reduction in pessimistic forecast errors for treated firms in the post-event period. Column III of Table 16 reports the results of model 8 for the PSM sample. The coefficient on the interaction of ***Pos$_t$*** and ***Treat$_j$*** is again positive and significant ($\beta = 0.047$, $p = 0.003$). Taken together, the results support the disciplining effect of investors' earnings expectations and suggest that analysts are less likely to issue pessimistic forecasts (i.e., pessimistic forecast errors) for firms covered by Estimize than for those without Estimize coverage.

Table 16 The impact of Estimize coverage on analyst forecast decision

	Unmatched Sample		**PSM Sample**			
	Column I: DV = $\boldsymbol{Error90d_{j,t}}$		**Column II: DV = $\boldsymbol{Treat_j}$**		**Column III: DV = $\boldsymbol{Error90d_{j,t}}$**	
Variable	**Coef.**	**t-stat**	**Coef.**	**t-stat**	**Coef.**	**t-stat**
Post	0.005	0.145			0.016	0.307
$\boldsymbol{Treat_y}$	−0.049	−4.283***			0.008	0.465
$\boldsymbol{Post_t \times Treat_j}$	0.029	2.492**			0.047	2.721***
$\boldsymbol{FirmSize_{j,t}}$	0.034	3.245**	2.437	45.967***	0.024	1.691
$\boldsymbol{Return_{j,t}}$	−0.020	−0.824			−0.035	−0.886
$\boldsymbol{Follow_{j,t}}$	−0.030	−3.223***	1.765	39.204***	−0.028	−2.218**
$\boldsymbol{BtM_{j,t}}$	−0.017	−2.315**	−0.925	−25.467***	−0.006	−0.520
$\boldsymbol{Turnover_{j,t}}$	−0.007	−1.185	0.742	19.930***	−0.014	−1.418
$\boldsymbol{Dispersion_{,t}}$	0.046	7.061***			−0.014	−1.312
$\boldsymbol{Guidance_{j,t}}$	−0.024	−3.903***			−0.036	−3.494***
$\boldsymbol{Error90d_{j,t}}$			−0.123	−3.748***		
Industry FE	Yes		Yes		Yes	
Year-Quarter FE	Yes		Yes		Yes	
Observations	39,814		39,814		12,968	
F statistics	4.913				3.633	
Prob > F	0.000				0.000	
Adj. R2	0.004				0.007	
LR chi2			20766.640			
Prob > chi2			0.000			
Pseudo R2			0.444			

Note: Column I reports the regression results of model 7 using all 39,814 observations. Column II reports the PSM matching process. Column III reported the regression results of model 7 using the PSM sample. $Error90d_{j,t}$ is the error of analyst EPS forecast consensus, measured as analyst EPS forecast consensus [−90, 0] days before the actual earnings announcement for firm *j* in quarter *t* minus the actual EPS, deflated by the stock price at the end of the previous quarter; Pos_t is a binary variable that equals 1 for the period 2012–18 and 0 for 2009–11; $Treat_j$ is a binary variable that equals 1 for treated firms and 0 for control firms. All variables are defined in Appendix C; *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

4.6.3 The Impact of Estimize Coverage on Analysts' Reliance on Public Information

A possible alternative explanation for the positive association between changes in investors' earnings expectations and analyst forecast revision is that analysts learn from investors' earnings expectations and use them to adjust their forecasts. To assess this alternative explanation, I examine the impact of Estimize coverage on the proportion of public information that analysts use in making their forecasts. If the alternative explanation holds, the proportion of public information in analyst forecasts is likely to increase because they incorporate investors' earnings expectations into their forecasts—especially for firms with Estimize coverage—because Estimize coverage makes investors' earnings expectations public.

I follow the literature (Barron et al., 1998; Gleason et al., 2020) to measure the proportion of public information to total information in analyst forecasts (***PublicPrc****$_{j,t,}$*):

$$PublicPrc_{j,t} = \frac{(SE_{j,t} - \frac{D_{j,t}}{N_{j,t}})}{(SE_{j,t} - \frac{D_{j,t}}{N_{j,t}} + D_{j,t})}$$

where $D_{j,t}$ is the realized forecast dispersion, $N_{j,t}$ is the number of analysts issuing forecasts, and $SE_{j,t}$ is the squared errors of the mean forecast, scaled by the absolute value of the actual EPS. $D_{j,t}$ and $SE_{j,t}$ are measured using EPS forecasts activated within 90 days after the earnings announcement date. I estimate the following model with standard errors adjusted for clustering at the firm level:

$$\begin{aligned} PublicPrc_{j,t} = \alpha_0 &+ \alpha_1 Post_t + \alpha_2 Treat_j + \alpha_3 Post_t \times Treat_j \\ &+ \alpha_4 FirmSize_{j,t} + \alpha_5 Return_{i,j,t} + \alpha_6 Leverage_{j,t} \\ &+ \alpha_7 Follow_{j,t} + \alpha_8 GdpChg_{j,t} + \alpha_9 Loss \\ &+ \alpha_{10} Miss_{j,t} + \alpha_{11} BtM_{j,t} \\ &+ \alpha_{12} TotalInfo_{j,t} \\ &+ \alpha_{13} PrivateInfo_{j,t} + \alpha_{14} PublicInfo_{j,t} \\ &+ \sum \alpha_j Industry + \sum \alpha_t YearQrt_t + \varepsilon_{i,j,t} \end{aligned} \tag{9}$$

The dependent variable is the proportion of public information to total information in analyst forecasts (***PublicPrc***$_{j,t,}$). The independent variable of interest is the interaction of ***Post*** (equals 1 for the period 2012–18 and 0 for 2009–11) and ***Treat***$_j$ (equals 1 for treated firms and 0 for control firms). All variables used in analyses are defined in detail in Appendix C. I also control for confounding covariates by implementing the PSM method. Specifically, I estimate a logistic regression in which the dependent variable is 1 for treated firms and 0 for control firms. Covariates include four firm characteristics (***FirmSize***$_{j,t}$, ***ROA***$_{j,t}$, ***Follow***$_{j,t}$, ***Leverage***$_{j,t}$, ***Miss***$_{j,t}$) and analysts' information environments (***TotalInfo***$_{j,t}$, ***PrivateInfo***$_{j,t}$, ***PublicInfo***$_{j,t}$).

Column I of Table 17 shows the results of model 9 prior to undertaking PSM. The coefficient on the interaction of ***Post*** and ***Treat***$_j$ is negative and significant ($\beta = -0.044$, $p = 0.000$), suggesting a decrease in the proportion of public information to total information used by analysts in forecasting earnings for treated firms in the post-event period. Column III of Table 17 reports the results of model 9 for the PSM sample. The coefficient on the interaction of ***Pos***$_t$ and ***Treat***$_j$ is again negative and significant ($\beta = -0.038$, $p = 0.004$). Taken together, the results suggest that analysts are less likely to rely on public information to make forecasts for firms covered by Estimize than for firms without Estimize coverage, which does not support the alternative explanation.

Table 17 The impact of Estimize coverage on analysts' reliance on public information

	Unmatched Sample		PSM Sample			
	Column I: DV = $PublicPrc_{j,t}$		Column II: DV = $Treat_j$		Column III: DV = $PublicPrc_{j,t}$	
Variable	**coef.**	**t-stat**	**coef.**	**t-stat**	**coef.**	**t-stat**
$Post_{j,t}$	0.027	1.246			0.057	1.475
$Treat_y$	0.030	4.159***			0.013	0.996
$Post_t \times Treat_j$	−0.044	−4.121***			−0.038	−2.613***
$FirmSize_{j,t}$	−0.027	−3.580***	3.062	43.838***	−0.067	−4.869***
$ROA_{j,t}$	0.001	0.282	0.664	12.795***	0.003	0.252
$Leverage_{j,t}$	−0.041	−8.157***	−0.187	−4.841***	−0.043	−4.152***
$Follow_{j,t}$	0.006	0.912	2.376	28.388***	0.002	0.176
$GdpChg_{j,t}$	−0.004	−0.097			0.041	0.561
$Loss_{j,t}$	0.003	0.498			0.000	0.029
$Miss_{,t}$	−0.073	−16.110***	−0.200	−5.165***	−0.065	−7.389***
$BtM_{j,t}$	−0.017	−3.322***	−0.296	−6.474***	−0.020	−1.891*
$TotalInfo_{j,t}$	0.481	104.929***	−0.282	−2.723***	0.437	48.555***
$PrivateInfo_{j,t}$			0.369	4.017***		
$PublicInfo_{j,t}$			0.340	5.354***		
Industry FE	Yes		Yes		Yes	
Year-Quarter FE	Yes		Yes		Yes	
Observations	36,905		36,905		10,198	
F statistics	242.041				54.555	
Prob > F	0.000				0.000	
Adj. R2	0.242				0.202	
LR chi2			11584.160			
Prob > chi2			0.000			
Pseudo R2			0.322			

Note: Column I reports the regression results of model 8 using all 36,905 observations. Column II reports the PSM matching process. Column III reports the regression results of model 7 using the PSM sample. $PublicPrc_{j,t}$ is the proportion of public information to total information in analyst forecasts; Pos_t is a binary variable that equals 1 for the period 2012–18 and 0 for 2009–11; $Treat_j$ is a binary variable that equals 1 for treated firms and 0 for control firms. All variables are defined in Appendix C; *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

4.7 Sensitivity Analysis

4.7.1 Alternative Measures

4.7.1.1 Analyst Forecast Consensus

This section focuses on analysts' issuance of *slightly* understated forecasts that generate small positive earnings surprises. To elaborate, managers prefer small positive earnings surprises to large positive earnings surprises because the latter negatively influences investor perceptions of the predictability of firm performance and increases the cost of capital (Bissessur & Veenman, 2016; Graham et al., 2005). Reporting on results of a survey, Graham et al. (2005, 43) notes:

> when asked about whether they would prefer to meet or to beat the earnings target, several CFOs [chief financial officers] say they would rather meet (or slightly beat) the earnings target rather than positively surprising the market in a big way every quarter.

A manager's preference for small positive earnings surprises motives analysts to issue slightly understated forecasts.

Specifically, I construct ***MeetBeat***$_{j,t}$ to capture the issuance of slightly understated forecasts (Bissessur & Veenman, 2016; Griffin & Lont, 2021); this indicator variable equals 1 for zero to two cents earnings surprises, and 0 for all other surprises. I re-run model 5, replacing ***Pessimism***$_{j,t}$ with ***MeetBeat***$_{j,t}$. Column I of Table 18 shows the results of the logistic regression. In line with the main analyses, the coefficient for ***RevDown***$_{j,t}$ is positive and significant ($\beta = 0.226$, $p = 0.013$), suggesting that analysts are more likely to issue forecasts that generate small positive earnings surprises after revising down their forecasts. The marginal effect of ***RevDown***$_{j,t}$ suggests that analysts who revise down their forecasts are around 7.2% more likely to end up with a small positive earnings surprise. Column II of Table 18 shows the results using only observations without management guidance for year quarter *t* announced during the [–90,

–1]-day period before the actual earnings announcement date. The coefficient for ***$RevDown_{j,t}$*** is still positive and marginally significant (β = 0.182, *p* = 0.058). Column III of Table 18 reports the results replacing ***$RevDown_{j,t}$*** with ***$DownFollow_{j,t}$***, and ***$DownAgainst_{j,t}$***. The coefficient for ***$DownFollow_{j,t}$*** is positive and significant (β = 0.193, *p* = 0.027), suggesting that analysts who revise their forecasts down following decreased investors' earnings expectations are more likely to generate small positive earnings surprises. The coefficient for ***$DownAgainst_{j,t}$*** is positive, but insignificant (β = 0.133, *p* = 0.141). Similar results are obtained using only observations without management guidance for year quarter *t* announced during the [–90, –1]-day period before the actual earnings announcement date. As shown in Column IV of Table 18, the coefficient for ***$DownFollow_{j,t}$*** is positive and marginally significant (β = 0.156, *p* = 0.085) and the coefficient for ***$DownAgainst_{j,t}$*** is insignificant (β = 0.106, *p* = 0.268).

Table 18 The probability of analysts' forecast errors being slightly pessimistic is greater when analysts' prior forecast revisions follow downward revisions in investor earnings expectations

	DV = $MeetBeat_{j,t}$							
	Column I: Full sample		**Column II: Exclude standalone guidance**		**Column III: Full sample**		**Column IV: Exclude standalone guidance**	
Variable	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**
$RevDown_{j,t}$	0.226	2.490**	0.182	1.899*				
$DownFollow_{j,t}$					0.193	2.207**	0.156	1.721*
$DownAgainst_{j,t}$					0.133	1.474	0.106	1.107
$EstmzRev_{j,t}$	−0.096	−2.710***	−0.104	−2.842***	−0.094	−2.637**	−0.103	−2.773***
$Follow_{j,t}$	−0.024	−0.227	−0.040	−0.351	−0.026	−0.241	−0.041	−0.361
$Dispersion_{,t}$	−0.460	−3.200***	−0.448	−3.068***	−0.460	−3.199**	−0.448	−3.066***
$FirmSize_{j,t}$	−0.565	−0.879	−0.644	−0.989	−0.572	−0.889	−0.650	−0.998
$BtM_{j,t}$	0.011	0.055	−0.022	−0.104	0.010	0.048	−0.024	−0.109
$Loss_{j,t}$	0.258	1.678*	0.245	1.542	0.260	1.687*	0.246	1.548
$Leverage_{j,t}$	0.194	0.602	0.308	0.893	0.194	0.601	0.307	0.892
$Turnover_{j,t}$	0.220	1.875*	0.286	2.357**	0.221	1.881*	0.287	2.364**
$RetVol_{j,t}$	−0.242	−0.996	−0.167	−0.668	−0.243	−1.000	−0.168	−0.673
$Guidance_{j,t}$	−0.002	−0.010	0.220	1.269	−0.003	−0.016	0.218	1.260
Regression Type	Logistic		Logistic		Logistic		Logistic	
No. of obs.	5210		4766		5210		4766	
Year-Quarter FE	Yes		Yes		Yes		Yes	
Firm FE	Yes		Yes		Yes		Yes	
Pseudo R Square	0.148		0.143		0.148		0.143	
_b[DownFollow]=_b[DownAgainst]					Pr > chi2 = 0.6185		Pr > chi2 = 0.6851	

Note: Columns I and III report the regression results using all 5,210 observations. Columns II and IV report the regression results using observations without management guidance for year quarter *t* announced during the [−90, −1]-day period before the actual earnings announcement date. $MeetBeat_{j,t}$ is a binary variable that equals 1 if the raw forecast error (i.e., analyst *i*'s EPS forecast [−30, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus the actual EPS) is [−0.02, 0]; 0 otherwise. $RevDown_{j,t}$ is a binary variable that equals 1 if $Revision_{j,t}$ is less than 0; 0 otherwise; $DownFollow_{j,t}$ is binary variable that equals 1 if both $Revision_{j,t}$ and $EstmzRev_{j,t}$ are less than 0; 0 otherwise; $DownAgainst_{j,t}$ is a binary variable that equals 1 if $Revision_{j,t}$ is less than 0 while $EstmzRev_{j,t}$ is greater than (or equal to) 0; 0 otherwise. All variables are defined in Appendix C; *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

4.7.1.2 Individual Analyst Forecast

In this section, I use individual analyst forecast data to examine the effect of analyst forecast revision on the occurrence of small positive earnings surprises at the individual analyst level. I first re-run model 7, replacing $\boldsymbol{Pessimism_{,j,t}}$ with $\boldsymbol{MeetBeat_{i,j,t}}$, which is an indicator variable that equals 1 for zero to two cents earnings surprises, and 0 for all other surprises. The coefficient for $\boldsymbol{RevDown_{i,j,t}}$ is negative but insignificant ($\beta = -0.160$, $p = 0.223$; as shown in Table A15). I then restrict the sample to within the just-beat and just-miss range (i.e., [−2 cents, 2 cents]) and re-run model 7. As shown in Column I of Table 19, the coefficient for $\boldsymbol{RevDown_{i,j,t}}$ is positive and significant ($\beta = 0.547$, $p = 0.002$), which provides partial evidence that analysts who downgrade their forecasts are more likely to generate small positive earnings surprise. Columns III and IV of Table 19 report the results replacing $\boldsymbol{RevDown_{i,j,t}}$ with $\boldsymbol{DownFollow_{i,j,t}}$ and $\boldsymbol{DownAgainst_{i,j,t}}$. As shown in Column III of Table 19 the coefficients for $\boldsymbol{DownFollow_{i,j,t}}$ ($\beta = 0.538$, $p = 0.004$) and $\boldsymbol{DownAgainst_{i,j,t}}$ ($\beta = 0.447$, $p = 0.019$) are positive and significant. However, the difference between the two coefficients is not statistically significant ($p = 0.966$). Similar results are obtained using only observations without management guidance for year quarter t announced during the [−90, −1]-day period before the actual earnings announcement date (Column IV of Table 19).

Table 19 The probability of analysts' forecast errors (at the individual analyst level) being slightly pessimistic is greater when analysts' prior forecast revisions follow downward revisions in investor earnings expectations

	DV = $MeetBeat_{i,j,t}$							
	Column I: Full sample		**Column II: Exclude standalone guidance**		**Column III: Full sample**		**Column IV: Exclude standalone guidance**	
Variable	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**
$RevDown_{i,j,t}$	0.547	3.136***	0.511	2.792**				
$DownFollow_{i,j,t}$					0.538	2.892***	0.453	2.324**
$DownAganist_{i,j,t}$					0.447	2.342**	0.483	2.447**
$EstmzRev_{j,t}$	0.012	0.096	0.057	0.432	0.014	0.108	0.034	0.244
$Follow_{j,t}$	0.149	0.879	0.071	0.407	0.148	0.876	0.075	0.426
$Dispersion_{,t}$	−0.118	−0.501	−0.158	−0.637	−0.118	−0.501	−0.159	−0.642
$AccScore_{i,j,t}$	0.049	0.438	0.098	0.826	0.050	0.436	0.094	0.789
$BoldScore_{i,j,t}$	−0.145	−0.773	−0.201	−1.026	−0.145	−0.771	−0.200	−1.020
$BiasScore_{i,j,t}$	0.567	2.848***	0.586	2.822**	0.566	2.855***	0.591	2.855***
$FirmExp_{i,j,t}$	0.000	−0.002	−0.071	−0.380	−0.001	−0.003	−0.071	−0.374
$IndusExp_{i,j,t}$	−0.075	−0.346	0.033	0.144	−0.075	−0.346	0.033	0.147
$GenExp_{i,j,t}$	0.172	0.810	0.127	0.575	0.172	0.809	0.127	0.572
$BrokerSize_{i,t}$	−0.191	−1.376	−0.149	−1.015	−0.191	−1.375	−0.149	−1.010
$FirmSize_{j,t}$	1.859	1.467	1.477	1.098	1.859	1.467	1.474	1.100
$BtM_{j,t}$	0.184	0.573	0.202	0.613	0.182	0.566	0.223	0.673
$Loss_{j,t}$	0.337	1.441	0.343	1.414	0.336	1.413	0.353	1.434
$Leverage_{j,t}$	−0.161	−0.226	0.139	0.243	−0.161	−0.226	0.138	0.240
$Turnover_{j,t}$	−1.267	−1.636	−1.074	−1.347	−1.266	−1.635	−1.082	−1.358
$RetVol_{j,t}$	1.476	2.541**	1.226	2.098**	1.477	2.546**	1.220	2.080**
$Guidance_{j,t}$	0.049	0.155	−0.191	−0.642	0.049	0.155	−0.189	−0.635
Regression Type	Logistic		Logistic		Logistic		Logistic	
No. of obs.	1717		1543		1717		1543	

Year-Quarter FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Pseudo R Square	0.136	0.136	0.136	0.136
b[DownFollow]=_b[DownAgainst]				
			Pr > chi2 = 0.9655	Pr > chi2 = 0.6227

Note: Table 19 reports the logistic regression results using observations within the just-beat and just-miss range (i.e., [−2 cents, 2 cents]). $MeetBeat_{j,t}$ is a binary variable that equals 1 if the raw forecast error (i.e., analyst *i*'s EPS forecast [−30, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus the actual EPS) is [−0.02, 0]; 0 otherwise; $RevDown_{j,t}$ is a binary variable that equals 1 if $Revision_{,j,t}$ is less than 0; 0 otherwise. All variables are defined in Appendix C; *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

4.7.2 Assessing the Magnitude of the Omitted Variable Threat

Even though my models include a comprehensive set of (observable) covariates and fixed effects to control for unobservable heterogeneity, there remains the possibility that the omission of other unobservable factors may materially bias the findings. To assess the extent of this threat, I first conduct a test proposed by Oster (2019) to compute the share of variation of unobservable factors (relative to observables) that is required to 'explain away' the observed effects (measured by the test statistic, δ), and reduce to zero the effect of a change in investors' earnings expectations (***EstmzRev*$_{j,t}$**) on analyst forecast revision (***Revison*$_{j,t}$**) (H1).[36] The Oster (2019) test requires setting a value of R_{max}, which denotes the value of R-squared from a hypothetical regression that consists of both unobservables and observables. Following Oster (2019) and Babenko et al. (2020), I specify R_{max} as 1.3 × $R_{\text{-squared}}$ (1.3 × 0.326), where $R_{\text{-squared}}$ is the R-squared from an OLS model that regresses ***Revison*$_{j,t}$** on ***EstmzRev*$_{j}$**, and includes all observables (with firm- and year-quarter-fixed effects).[37] The obtained δ is 5.757, which indicates that the unobservables need to be more than five times as significant as the observables to reduce the effect of ***EstmzRev*$_{j,t}$** to zero, which seems highly unlikely given that my regression model includes many factors known to affect analyst forecast revision and a number of fixed effects.

I then conduct another test that computes the share of variation of unobservable factors that is required to 'explain away' the observed effects and reduce to zero the effect of analyst forecast revision following downgrade in investors' earnings expectations (***DownFollow*$_{j,t}$**) on the occurrence of pessimistic forecast errors (***Pessimism*$_{j,t}$**) (H2). I specify R_{max} as 1.3 × $R_{\text{-squared}}$

[36] For example, δ = 2 suggests that the unobservable variables need to be twice as significant as observables for the omitted variable bias to 'explain away' the results and decrease the coefficient of interest to zero (Babenko et al., 2020). Because $\delta\frac{\sigma_1 X}{{\sigma_1}^2} = \frac{\sigma_2 X}{{\sigma_2}^2}$ in Oster (2019), a negative δ means that if the observables are positively correlated with the treatment, the unobservables have to be negatively correlated with the treatment to decrease the coefficient of interest to zero.

[37] I use the Stata command *psacalc* provided by Oster (2019) to conduct the test.

(1.3 × 0.291), where $R_{-squared}$ is the R-squared from an OLS model that regresses $\boldsymbol{Pessimism_{j,t}}$ on $\boldsymbol{DownFollow_{j,t}}$ and includes all observables (with firm- and year-quarter-fixed effects). The obtained δ is −5.678, which indicates that the unobservables need to be more than five times as significant as the observables to reduce the effect of $\boldsymbol{DownFollow_{j,t}}$ to zero, which again seems highly unlikely given that my regression model includes many factors known to affect analyst forecast revision and a number of fixed effects. Taken together, it appears unlikely that omitted variables have materially biased the findings.

4.8 Chapter Summary

Crowdsourcing platforms unlock the potential to aggregate a large volume of investors' discussions and facilitate the formation of investors' beliefs about a firm's future earnings. Crowdsourced forecasts are unbiased on average and thus make analyst forecast bias more salient. The findings highlight how investors' earnings expectations from crowdsourcing platforms affect analysts who are incentivized by self-interest to walk down their forecasts to be easy to beat.

I first document that the magnitude of analysts' walk-downs increases with investors' downgrading of their earnings expectations. I then find that the likelihood of analysts issuing pessimistic forecasts increases with analysts' walk-down. In addition, I show that although analysts do not rely more on investors' earnings expectations in making their forecasts when they are knowledgeable about investors' forecasts (i.e., when firms are covered by Estimize), they tend to be less pessimistic in their forecasts. Taken together, the findings are suggestive of a disciplining effect of investors' earnings expectations on analysts' walk-down to beatable forecasts.

The findings contribute to research on how crowdsourcing earnings estimates influence analysts' forecasts decisions. Earlier studies suggest that analyst forecast bias is more salient to investors than are crowdsourcing earnings estimates (Schafhäutle & Veenman, 2021), and provide some evidence that crowdsourcing earnings estimates have a disciplining effect on analyst forecasting behavior (Jame et al., 2021). However, this evidence is based on the initiation of Estimize coverage, which cannot relate analysts' forecasting behavior to their fear of contradicting investor earnings expectations. The study complements this line of research by showing that crowdsourcing earnings estimates (unlike the initiation of Estimize coverage) are information that analysts consider when acting opportunistically in their forecast revisions. These findings should be of interest to securities regulators. Online crowdsourcing technologies gives investor individuals a forum to obtain information at low cost and disseminate their opinions to a vast audience. At the same time, crowdsourcing earnings expectations can disrupt the market for traditional earnings forecast providers by increasing the salience of analyst forecast bias. Changes can potentially reshape how investors access and share information and introduce an alternative source of earnings forecast to analysts. A rapid regulatory response to these new crowdsourcing technologies will be critical. There are some areas that regulators could consider: the lack of understanding or conceptual clarity about crowdsourcing; ensuring good quality of submissions from the crowd; difficulties associated with outsourcing to the crowd in traditional 'line' industries; challenges to talent and organizational culture; keeping hold of confidential information and intellectual property, and additional risks associated with using the crowd. This requires regulators with capabilities (e.g., skills, resources, and authority) to create investor protection regulations with effectiveness, and implement these regulations.

While I believe that the findings provide interesting insights into how crowdsourced earnings estimates affect analyst forecast decisions and open a new avenue for future research, I recognize several potential endogeneity threats. To address these issues, I (i) verify the argument that investors' earnings expectations impose a disciplining effect on analyst forecast walk-down, using a difference-in-difference model; and (ii) address the alternative explanation that analysts are learning from investors' earnings expectations, rather than being disciplined by investors. The inferences from the main results remain robust. I also execute several robustness tests by (iii) using a restricted sample to control for the effect of management guidance and (iv) adopting alternative measures of analyst forecast revisions. These tests broadly support the disciplining effect of investors' earnings expectations, which explains the association between changes in crowdsourced earnings estimates and analyst forecast revisions. Overall, the findings support the notion that analysts infer investors' earnings expectations from crowdsourcing platforms and use this to inform their opportunistic forecast behavior.

Chapter 5 Discussion

5.1 Summary of Research Questions and Findings

This thesis is comprised of two studies investigating the role of investor sentiment and earnings expectation in capital market phenomena. The advent of online social networks means that investors can publish and publicly share their opinions about firm prospects. In this way, management and analysts can more easily access and aggregate individual opinions regarding a firm.

The first study focuses on NPI sentiment and investigates whether managers use optimistic earnings guidance to exploit NPI sentiment regarding their firms in an attempt to induce desirable market reactions. I infer firm-level NPI sentiment from social media discussions on StockTwits, using the 17 million tweets posted by 118,685 users concerning 3,212 distinct firms between May 2008 and January 2017. Using the NPI reaction proxy of Aboody et al. (2018), I find that NPI reaction to positive guidance is stronger when NPI sentiment is high, and that managers are more likely to issue positive guidance at these times. This association between the likelihood of issuing positive guidance and NPI sentiment is stronger in firms in which NPIs have greater proportionate shareholdings and where managers' equity incentives are highly contingent on short-term stock price increases. The findings are consistent with managers opportunistically manipulating guidance to exploit the sentiment of NPIs.

The second study focuses on investors' earnings expectations and investigates whether analysts exploit investors' earnings expectations about a firm to walk down their forecasts to be beatable. I infer investors' earnings expectation from crowdsourced earnings estimates on Estimize, which publishes 879,015 'street earnings' estimates submitted by 70,926 participants in the period January 2012–September 2018. I find that the level of revision in the analyst forecast consensus increases the level of change in investor earnings expectation, and the likelihood of analysts issuing forecasts that generate pessimistic errors is higher when analysts revise down their forecasts. These findings are consistent with analysts opportunistically manipulating forecasts to exploit investors' expectations of future earnings.

5.2 Implications

The contributions of the thesis include the following. The findings in relation to the first objective of the thesis contribute to the literature in three ways. First, while much of the literature regards NPI sentiment as 'noise' and dedicates little attention to its effects on managers' decision making, this study suggests that NPI sentiment is informative, rather than pure 'noise,' and may help explain the factors affecting managers' decisions to issue positive guidance. Second, prior to the advent of social network platforms, observing and robustly aggregating individuals' opinions about a particular firm was problematic for researchers. Social network platforms have become a popular communication tool among investors in recent years. NPIs keep up with the latest news and trends in the finance world. They submit tweets about firm performance, and foster discussion. Using social media data to infer NPI sentiment opens up future research to collect NPIs' opinions of individual firms in a natural environment. Third, in April 2019, the SEC issued an investor bulletin expressing its concern regarding the ethical use of social-sentiment-investing tools by firms; primarily, how these tools can be used to manipulate a stock's price.[38] Responding to the SEC's concern, this study shows that managers can exploit social media discussions to infer NPI sentiment and trigger NPIs to react to positive guidance in the way that managers hope for.

[38] For further information, see *Investor bulletin: Social sentiment investing tools—think twice before trading based on social media*. https://www.sec.gov/oiea/investor-alerts-and-bulletins/ib_sentimentinvesting

The findings stemming from the second objective of the thesis contribute to the literature in three ways. First, recent research provides evidence that crowdsourced forecasts can be useful for investors in predicting firms' future earnings (Adebambo & Bliss, 2015; Jame et al., 2016) and pricing earnings news (Schafhäutle & Veenman 2021). Adding to this research, I find a positive association between changes in crowdsourced earnings estimates and analyst forecast revision, suggesting crowdsourced forecasts also provide analysts valuable information that affects their forecast decisions. Second, this study contributes to understanding of the market forces that constrain analysts' conflicts of interest. My findings confirm the role of reputational considerations in disciplining analysts and show that analysts are likely to opportunistically issue pessimistic forecasts following a downgrade in investors' earnings expectations to mitigate the risk of losing reputation. Third, on 30 August 2010, the SEC published an investor publication that describes analysts' conflicts of interest and enable investors to recognize such conflicts. The arrival of Estimize provides a means for investors to assess the extent to which analysts opportunistically bias their forecasts. Investors can select which analyst forecast to rely on when faced with multiple forecasts from different analysts.

5.3 Limitations

My thesis is subject to limitations inherent in the study design, particularly in regard to potential endogeneity threats. First, the hypothesis in the first aspect of the thesis (regarding NPIs and management) hinges on the assumption that managers consider NPI sentiment when issuing guidance and that NPI sentiment can be inferred from social media postings. However, it is plausible that the measure of NPI sentiment is correlated with some other time-varying factor that drives managers' guidance decisions, raising the prospect of an omitted variable threat. To address this potential threat, I investigate the impact of an exogenous shock on managers wary of false or misleading information on online social networks. Second, a potential alternative explanation for the positive association between NPI sentiment and the likelihood of managers issuing positive guidance is that NPIs' expectations of forthcoming guidance direction affect their sentiment, which is the opposite to the direction posited in this study. To address this, I employ a 2SLS approach to isolate NPIs' anticipation of forthcoming guidance. Also, there remains a possibility that the omission of other unobservable factors may materially bias the findings. To assess the extent of this threat and mitigate omitted variables concerns, I conduct an unobservable selection analysis.

In addition, H1, in the second aspect of the thesis (regarding NPIs and analysts) hinges on the argument that investors' earnings expectations impose a disciplining effect on analyst forecast walk-down. To address this issue, I adopt a difference-in-difference model to compare changes in analyst forecast bias for treatment and control firms around the initiation of Estimize coverage. Further, a potential alternative explanation for the positive association between analyst forecast revision and a change in investors' earnings expectations is that analysts learn from investors' earnings expectations and use them to adjust their forecasts. To address this alternative explanation, I examine the impact of Estimize coverage on the proportion of public information used by analysts in making their forecasts. I also conduct an unobservable selection analysis to mitigate the omitted variables concern. Overall, these limitations are not expected to substantially affect the inferences.

5.4 Future Research Opportunities

My thesis leads to several suggestions for future research. First, concerning how the likelihood of a manager issuing positive guidance is associated with the sentiment of NPIs toward their firm, future research may consider using other popular conventions on social media platform to improve the sentiment analysis. These conventions include 'retweet', a practice of sharing another user's tweets to one's network; 'like', a mechanism to show appreciation and agreement for another user's tweets; and 'reply', a form of addressing and direct communication with other users. These potentially capture how individuals form ties with people with similar opinions versus dissimilar others; how their sentiment is spread via these ties; and how the strength of these ties affects sentiment dissemination.

Second, concerning how analysts' walk-down to beatable forecasts is associated with investors' earnings expectations toward a firm, future research may consider using other features of crowds formed on online social networks to better gauge investors' earnings expectations. These features include crowd size—the number of individuals working together; demographic diversity—the extent to which a crowd is heterogeneous with respect to various attributes (e.g., education and working experience); and cognitive diversity—the different ways that people see the world, interpret problems in it, and make predictions. These features potentially affect how well a crowd represents investors' opinions about a firm. Also, in this study, I use archival data. Due to the nature of archival data, it is difficult to rule out the alternative argument that the information in Estimize forecasts and IBES forecasts overlaps as both types of forecasts are about the same firm and argue for a causal relationship based on archival data. .In the future, I will consider using surveys or experiments to explore further the relation between analyst forecast revisions and investor earnings expectations.

Third, future research may consider explore the ability of individual investors in nowadays capital markets. Do they become more subjective to the manipulative attempts of other market participants like managers and analysts since social media makes their sentiment more visible? Alternatively, do they become more resistant to the manipulative attempts of other market participants thanks to the crowdsourced wisdom?

Chapter 6 Conclusion

The sentiments and expectations of investors about a firm have important implications for managers and analysts. The thesis first explores whether and how the likelihood of a manager issuing positive guidance is associated with the sentiment of NPIs about their firm and provides evidence on managers opportunistically manipulating guidance to exploit the sentiment of NPIs. The thesis then investigates how analysts' walk-down to beatable forecasts is associated with investors' expectations of the firm's future earnings, and provides evidence of analysts opportunistically manipulating forecasts to exploit investors' expectations of future earnings.

As discussed in Chapter 5.2, this thesis contributes to the literature and provides researchers and regulators with insights on how rapid technology development transforms information sharing in the capital market. The arrival of online social networks allows investors to connect and share information about stocks that facilitate their investment processes. Further, these networks also offer opportunities for managers and analysts to exploit investors' opinions regarding a firm for their own interest. Rapid regulatory responses to these technologies will be critical. This requires regulators with the necessary capabilities (e.g., skills, resources, and authority) to create and implement effective investor protection regulations.

References

Aboody, D., Even-Tov, O., Lehavy, R., & Trueman, B. (2018). Overnight returns and firm-specific investor sentiment. *Journal of Financial & Quantitative Analysis*, *53*(2), 485–505.

Adebambo, B., Bliss, B., & Kumar, A. (2016). *Geography, diversity, and accuracy of crowdsourced earnings forecasts* [Working paper]. University of San Diego. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2579402

Adebambo, B. N., & Bliss, B. (2015). *The value of crowdsourcing: Evidence from earnings forecasts* [Working paper]. University of San Diego. https://academic.oup.com/rfs/article/30/3/801/2682978?login=true#114137589

Ajinkya, B. B., & Gift, M. J. (1984). Corporate managers' earnings forecasts and symmetrical adjustments of market expectations. *Journal of Accounting Research*, *22*(2), 425–444.

Altınkılıç, O., Balashov, V. S., & Hansen, R. S. (2019). Investment bank monitoring and bonding of security analysts' research. *Journal of Accounting & Economics*, *67*(1), 98–119.

Altschuler, D., Chen, G., & Zhou, J. (2015). Anticipation of management forecasts and analysts' private information search. *Review of Accounting Studies*, *20*(2), 803–838.

An, H., & Zhang, T. (2013). Stock price synchronicity, crash risk, and institutional investors. *Journal of Corporate Finance*, *21*, 1-15.

Anilowski, C., Feng, M., & Skinner, D. J. (2007). Does earnings guidance affect market returns? The nature and information content of aggregate earnings guidance. *Journal of Accounting and Economics*, *44*(1–2), 36–63.

Asay, H. S., Libby, R., & Rennekamp, K. M. (2018). Do features that associate managers with a message magnify investors' reactions to narrative disclosures? *Accounting, Organizations & Society*, *68*, 1–14.

Audrino, F., Sigrist, F., & Ballinari, D. (2020). The impact of sentiment and attention measures on stock market volatility. *International Journal of Forecasting*, *36*(2), 334–357.

Babenko, I., Fedaseyeu, V., & Zhang, S. (2020). Do CEOs affect employees' political choices? *The Review of Financial Studies*, *33*(4), 1781–1817.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, *61*(4), 1645–1680.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *The Journal of Economic Perspectives, 21*(2), 129–151.

Balakrishnan, K., Billings, M. B., Kelly, B., & Ljungqvist, A. (2014). Shaping liquidity: On the causal effects of voluntary disclosure. *The Journal of Finance, 69*(5), 2237–2278.

Barberis, N., Greenwood, R., Jin, L., & Shleifer, A. (2018). Extrapolation and bubbles. *Journal of Financial Economics, 129*(2), 203–227.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*(3), 307–343.

Barbier, G., Zafarani, R., Gao, H., Fung, G., & Liu, H. (2012). Maximizing benefits from crowdsourced data. *Computational & Mathematical Organization Theory, 18*(3), 257–279.

Barron, O. E., Kim, O., Lim, S. C., & Stevens, D. E. (1998). Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review, 73*(4), 421–433.

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

Beaver, W. H., McNichols, M. F., & Wang, Z. Z. (2020). Increased market response to earnings announcements in the 21st century: An empirical investigation. *Journal of Accounting & Economics, 69*(1), 101244.

Behrendt, S., & Schmidt, A. (2018). The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility. *Journal of Banking & Finance, 96*, 355–367.

Berkman, H., Koch, P. D., Tuttle, L., & Zhang, Y. J. (2012). Paying attention: Overnight returns and the hidden cost of buying at the open. *Journal of Financial & Quantitative Analysis, 47*(4), 715–741.

Beyer, A., Cohen, D. A., Lys, T. Z., & Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting & Economics, 50*(2), 296–343.

Bhardwaj, A., & Imam, S. (2019). The tone and readability of the media during the financial crisis: Evidence from pre-IPO media coverage. *International Review of Financial Analysis, 63*, 40–48.

Bhattacharya, U., Galpin, N., Ray, R., & Yu, X. (2009). The role of the media in the internet IPO bubble. *Journal of Financial & Quantitative Analysis, 44*(3), 657–682.

Billings, M. B., Jennings, R., & Lev, B. (2015). On guidance and volatility. *Journal of Accounting & Economics*, *60*(2–3), 161–180.

Bissessur, S. W., & Veenman, D. (2016). Analyst information precision and small earnings surprises. *Review of Accounting Studies, 21*(4), 1327–1360.

Blasco, N., Corredor, P., & Ferrer, E. (2018). Analysts herding: When does sentiment matter? *Applied Economics, 50*(51), 5495–5509.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2020). Memory, attention, and choice. *The Quarterly Journal of Economics, 135*(3), 1399–1442.

Boritz, J. E., Hayes, L., & Timoshenko, L. M. (2020). How understandable are SOX 404 auditors' reports? *International Journal of Accounting Information Systems, 39*, 100486.

Bradley, D., Clarke, J., & Cooney Jr, J. (2012). The impact of reputation on analysts' conflicts of interest: Hot versus cold markets. *Journal of Banking & Finance, 36*(8), 2190–2202.

Bradley, D., Jame, R., & Williams, J. (2020). *Non-deal roadshows, investor welfare, and analyst conflicts of interest*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3302687

Bradshaw, M. T., Lee, L. F., & Peterson, K. (2016). The interactive role of difficulty and incentives in explaining the annual earnings forecast walkdown. *The Accounting Review*, *91*(4), 995–1021.

Brandon, D. M., Long, J. H., Loraas, T. M., Mueller-Phillips, J., & Vansant, B. (2014). Online instrument delivery and participant recruitment services: Emerging opportunities for behavioral accounting research. *Behavioral Research in Accounting*, *26*(1), 1–23.

Bratten, B., Payne, J. L., & Thomas, W. B. (2016). Earnings management: Do firms play 'follow the leader'? *Contemporary Accounting Research, 33*(2), 616–643.

Brave, S., & Nass, C. (2007). Emotion in human-computer interaction. In J. Jacko & A. Sears (Eds.), *The human-computer interaction handbook* (pp. 103–118) New York: Lawrence Erlbaum Associates.

Brochet, F., Faurel, L., & McVay, S. (2011). Manager-specific effects on earnings guidance: An analysis of top executive turnovers. *Journal of Accounting Research, 49*(5), 1123–1162.

Brown, L. D., Call, A. C., Clement, M. B., & Sharp, N. Y. (2015). Inside the 'black box' of sell-side financial analysts. *Journal of Accounting Research, 53*(1), 1–47.

Brown, L. D., & Kim, K.-J. (1991). Timely aggregate analyst forecasts as better proxies for market earnings expectations. *Journal of Accounting Research*, *29*(2), 382–385.

Brown, L. D., Richardson, G. D., & Schwager, S. J. (1987). An information interpretation of financial analyst superiority in forecasting earnings. *Journal of Accounting Research*, *25*(1), 49–67.

Brown, N. C., Christensen, T. E., Elliott, W. B., & Mergenthaler, R. D. (2012). Investor sentiment and pro forma earnings disclosures. *Journal of Accounting Research, 50*(1), 1–40.

Brown, S., Hillegeist, S. A., & Lo, K. (2009). The effect of earnings surprises on information asymmetry. *Journal of Accounting & Economics, 47*(3), 208–225.

Buchanan, B., Cao, C. X., & Chen, C. (2018). Corporate social responsibility, firm value, and influential institutional ownership. *Journal of Corporate Finance, 52*, 73–95.

Bushee, B. J. (1998). The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review*, *73*(3), 305-333.

Bushee, B. J. (2001). Do institutional investors prefer near-term earnings over long-run value? *Contemporary Accounting research*, *18*(2), 207-246.

Byun, S., & Roland, K. (2021). Analyst bias and forecast consistency. *Accounting & Finance*

Cade, N. L. (2018). Corporate social media: How two-way disclosure channels influence investors. *Accounting, Organizations & Society, 68*, 63–79.

Cahill, D., Wee, M., & Yang, J. W. (2017). Media sentiment and trading strategies of different types of traders. *Pacific-Basin Finance Journal, 44*, 160–172.

Campbell, J. L., DeAngelis, M. D., & Moon, J. R. (2019). Skin in the game: Personal stock holdings and investors' response to stock analysis on social media. *Review of Accounting Studies*, *24*(3), 731–779.

Carvajal, M., Coulton, J. J., & Jackson, A. B. (2017). Earnings benchmark hierarchy. *Accounting & Finance, 57*(1), 87–111.

Chang, J. W., & Choi, H. M. (2017). Analyst optimism and incentives under market uncertainty. *Financial Review, 52*(3), 307–345.

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

Cheng, Q., & Lo, K. (2006). Insider trading and voluntary disclosures. *Journal of Accounting Research, 44*(5), 815–848.

Cheng, Q., Luo, T., & Yue, H. (2013). Managerial incentives and management forecast precision. *The Accounting Review, 88*(5), 1575–1602.

Chi, L., Zhang, G., Zhuang, X., & Song, D. (2012). Investor sentiment indicators and stock markets: a study based on the extended Kalman filter method. *Journal of Management in Engineering, 126*, 132–169.

Cho, M., & Kwon, S. Y. (2014). Trading volume and investor disagreement around management forecast disclosures. *Journal of Accounting, Auditing & Finance, 29*(1), 3–30.

Choi, B., & Kim, J. B. (2017). The effect of CEO stock-based compensation on the pricing of future earnings. *European Accounting Review, 26*(4), 651–679.

Clement, M. B. (1999). Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting & Economics, 27*(3), 285–303.

Coakley, J., & Lazos, A. (2021). New developments in equity crowdfunding: A review. *Review of Corporate Finance, 1*(3–4), 341–405.

Contreras, H., & Marcet, F. (2021). Sell-side analyst heterogeneity and insider trading. *Journal of Corporate Finance, 66*, 101778.

Cookson, J. A., Fos, V., & Niessner, M. (2021, 12 January). *Does disagreement facilitate informed trading? Evidence from Activist Investors.* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3765092

Cordeiro, J. J., & Tewari, M. (2015). Firm characteristics, industry context, and investor reactions to environmental CSR: A stakeholder theory approach. *Journal of Business Ethics, 130*(4), 833–849.

Cote, J. (2000). Analyst credibility: The investor's perspective. *Journal of Managerial Issues, 12*(3), 352–362.

Cotter, J., Tuna, I., & Wysocki, P. D. (2006). Expectations management and beatable targets: How do analysts react to explicit earnings guidance? *Contemporary Accounting Research, 23*(3), 593–624.

Cowen, A., Groysberg, B., & Healy, P. (2006). Which types of analyst firms are more optimistic? *Journal of Accounting & Economics, 41*(1–2), 119–146.

Crawford, S., Gray, W., Johnson, B. R., & Price III, R. A. (2018). What motivates buy-side analysts to share recommendations online? *Management Science, 64*(6), 2574–2589.

D'Augusta, C. (2022). Does accounting conservatism make good news forecasts more credible and bad news forecasts less alarming? *Journal of Accounting, Auditing & Finance, 37*(1), 77–113.

Da, Z., & Huang, X. (2020). Harnessing the wisdom of crowds. *Management Science, 66*(5), 1847–1867.

Daniel, K., & Hirshleifer, D. (2015). Overconfident investors, predictable returns, and excessive trading. *Journal of Economic Perspectives, 29*(4), 61–88.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *The Journal of Finance, 53*(6), 1839–1885.

Demmer, M., Pronobis, P., & Yohn, T. L. (2019). Mandatory IFRS adoption and analyst forecast accuracy: The role of financial statement-based forecasts and analyst characteristics. *Review of Accounting Studies, 24*(3), 1022–1065.

Deng, S., Huang, Z. J., Sinha, A. P., & Zhao, H. (2018). The interaction between microblog sentiment and stock return: An empirical examination. *MIS Quarterly, 42*(3), 895–918.

Donoughue, P., & Chau, D. (2021). *GameStop stock has surged thanks to enthusiastic WallStreetBets Reddit users. Here's what you need to know*. ABC News. https://www.abc.net.au/news/2021-01-28/gamestop-reddit-users-on-wallstreetbets-explained/13097982

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

Du, Q., & Shen, R. (2018). Peer performance and earnings management. *Journal of Banking & Finance, 89*, 125–137.

Dunham, L. M., & Garcia, J. (2020). Measuring the effect of investor sentiment on liquidity. *Managerial Finance, 47*(1), 59–85.

Dzieliński, M. (2017, 2 September). *Do news agencies help clarify corporate disclosure?* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2266173

Eddy, A., & Seifert, B. (1992). An examination of hypotheses concerning earnings forecast errors. *Quarterly journal of Business and Economics, 31*(2), 22–37.

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance, 64*(5), 2023–2052.

Fang, L., & Yasuda, A. (2009). The effectiveness of reputation as a disciplinary mechanism in sell-side research. *The Review of Financial Studies, 22*(9), 3735–3777.

Farrell, A. M., Grenier, J. H., & Leiby, J. (2017). Scoundrels or stars? Theory and evidence on the quality of workers in online labor markets. *The Accounting Review, 92*(1), 93–114.

Feng, M., Li, C., & McVay, S. (2009). Internal control and management guidance. *Journal of Accounting & Economics, 48*(2–3), 190–209.

Feng, M., & McVay, S. (2010). Analysts' incentives to overweight management guidance when revising their short-term earnings forecasts. *The Accounting Review, 85*(5), 1617–1646.

Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture*. Sage.

Foster, G. (1973). Stock market reaction to estimates of earnings per share by company officials. *Journal of Accounting Research*, *11*(1), 25–37.

Frankel, R., McNichols, M., & Wilson, G. P. (1995). Discretionary disclosure and external financing. *The Accounting Review*, *70*(1), 135–150.

Fried, D., & Givoly, D. (1982). Financial analysts' forecasts of earnings: A better surrogate for market expectations. *Journal of Accounting & Economics, 4*(2), 85–107.

Frijda, N. H. (1994). Varieties of affect: Emotions and episodes, moods, and sentiments. In P. Ekman, R.J. Davidson (Eds.), The nature emotion: Fundamental questions (pp. 59–67) New York: Oxford University Press

Giannini, R., Irvine, P., & Shu, T. (2019). The convergence and divergence of investors' opinions around earnings news: Evidence from a social network. *Journal of Financial Markets, 42*, 94–120.

Gilbert, C. H. E. (2014 , 1-4 June). *Vader: A parsimonious rule-based model for sentiment analysis of social media text* [Paper presentation]. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Michigan, USA. http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf

Givoly, D., & Lakonishok, J. (1979). The information content of financial analysts' forecasts of earnings: Some evidence on semi-strong inefficiency. *Journal of Accounting & Economics*, *1*(3), 165–185.

Givoly, D., & Lakonishok, J. (1980). Financial analysts' forecasts of earnings: Their value to investors. *Journal of Banking & Finance, 4*(3), 221–233.

Gleason, C., Ling, Z., & Zhao, R. (2020). Selective disclosure and the role of Form 8-K in the post-Reg FD era. *Journal of Business Finance & Accounting, 47*(3–4), 365–396.

Gleason, C. A., & Lee, C. M. (2003). Analyst forecast revisions and market price discovery. *The Accounting Review, 78*(1), 193–225.

Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting & Economics, 40*(1–3), 3–73.

Griffin, P. A., & Lont, D. H. (2021). Evidence of an increasing trend in earnings surprises over the past two decades: The role of positive manager-initiated non-GAAP adjustments. *Journal of Business Finance & Accounting, 48*, 1525–1559.

Griffith, J., Najand, M., & Shen, J. (2020). Emotions in the stock market. *Journal of Behavioral Finance, 21*(1), 42–56.

Grinblatt, M., Jostova, G., & Philipov, A. (2018, 4 February). *Explaining (some) anomalies: The role of analyst bias* [Working paper]. UCLA Anderson School of Management. https://anderson-review.ucla.edu/wp-content/uploads/2021/03/Grinblatt_SSRN-id2653666.pdf

Groysberg, B., Healy, P. M., & Maber, D. A. (2011). What drives sell-side analyst compensation at high-status investment banks? *Journal of Accounting Research, 49*(4), 969–1000.

Hales, J., Moon Jr, J. R., & Swenson, L. A. (2018). A new era of voluntary disclosure? Empirical evidence on how employee postings on social media relate to future corporate disclosures. *Accounting, Organizations & Society, 68*, 88–108.

Hamadi, R., Ghazzai, H., Besbes, H., & Massoud, Y. (2020). Financial advisor recruitment: A smart crowdsourcing-assisted approach. *IEEE Transactions on Computational Social Systems, 8*(3), 682–688.

Han, J., & Tan, H. T. (2010). Investors' reactions to management earnings guidance: The joint effect of investment position, news valence, and guidance form. *Journal of Accounting Research, 48*(1), 81–104.

Han, S., Jin, J. Y., Kang, T., & Lobo, G. (2014). Managerial ownership and financial analysts' information environment. *Journal of Business Finance & Accounting, 41*(3–4), 328–362.

Hartzmark, S. M., & Shue, K. (2018). A tough act to follow: Contrast effects in financial markets. *The Journal of Finance, 73*(4), 1567–1613.

Hendershott, T., Livdan, D., & Schürhoff, N. (2015). Are institutions informed about news? *Journal of Financial Economics, 117*(2), 249–287.

Herrmann, D., Hope, O. K., Payne, J. L., & Thomas, W. B. (2011). The market's reaction to unexpected earnings thresholds. *Journal of Business Finance & Accounting, 38*(1–2), 34–57.

Hilary, G., & Hsu, C. (2013). Analyst forecast consistency. *The Journal of Finance, 68*(1), 271–297.

Hilbert, M. (2012). Toward a synthesis of cognitive biases: how noisy information processing can bias human decision making. *Psychological Bulletin, 138*(2), 211.

Hirshleifer, D. (2001). Investor psychology and asset pricing. *The Journal of Finance, 56*(4), 1533–1597.

Hirshleifer, D., & Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance, 58*(3), 1009–1032.

Hirst, D. E., Koonce, L., & Venkataraman, S. (2008). Management earnings forecasts: A review and framework. *Accounting Horizons, 22*(3), 315–338.

Horton, J., Serafeim, G., & Wu, S. (2017). Career concerns of banking analysts. *Journal of Accounting & Economics*, *63*(2–3), 231–252.

Houston, J. F., Lev, B., & Tucker, J. W. (2010). To guide or not to guide? Causes and consequences of stopping quarterly earnings guidance. *Contemporary Accounting Research, 27*(1), 143–185.

Houston, J. F., Lin, C., Liu, S., & Wei, L. (2019). Litigation risk and voluntary disclosure: Evidence from legal changes. *The Accounting Review*, *94*(5), 247–272.

Hribar, P., Melessa, S. J., Small, R. C., & Wilde, J. H. (2017). Does managerial sentiment affect accrual estimates? Evidence from the banking industry. *Journal of Accounting & Economics*, *63*(1), 26–50.

Hribar, P., & Yang, H. (2016). CEO overconfidence and management forecasting. *Contemporary Accounting Research, 33*(1), 204–227.

Hurwitz, H. (2018). Investor sentiment and management earnings forecast bias. *Journal of Business Finance & Accounting, 45*(1–2), 166–183.

Jame, R., Johnston, R., Markov, S., & Wolfe, M. C. (2016). The value of crowdsourced earnings forecasts. *Journal of Accounting Research, 54*(4), 1077–1110.

Jame, R., Markov, S., & Wolfe, M. (2021). Can FinTech competition improve sell-side research quality? *The Accounting Review*. Advance online publication. https://doi.org/10.2308/TAR-2019-0266

Janssen, D.-J., Füllbrunn, S., & Weitzel, U. (2019). Individual speculative behavior and overpricing in experimental asset markets. *Experimental Economics, 22*(3), 653–675.

Jennings, R. (1987). Unsystematic security price movements, management earnings forecasts, and revisions in consensus analyst earnings forecasts. *Journal of Accounting Research, 25*(1), 90–110.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*(4), 305-360.

Jiang, J. (2008). Beating earnings benchmarks and the cost of debt. *The Accounting Review*, *83*(2), 377–416.

Kadous, K., Mercer, M., & Thayer, J. (2009). Is there safety in numbers? The effects of forecast accuracy and forecast boldness on financial analysts' credibility with investors. *Contemporary Accounting Research, 26*(3), 933–968.

Kaplan, S. E., Samuels, J. A., & Cohen, J. (2015). An examination of the effect of CEO social ties and CEO reputation on nonprofessional investors' say-on-pay judgments. *Journal of Business Ethics, 126*(1), 103–117.

Karamanou, I. (2011). On the determinants of optimism in financial analyst earnings forecasts: The effect of the market's ability to adjust for the bias. *Abacus, 47*(1), 1–26.

Ke, B., & Petroni, K. (2004). How informed are actively trading institutional investors? Evidence from their trading behavior before a break in a string of consecutive earnings increases. *Journal of Accounting Research*, *42*(5), 895-927.

Ke, B., & Yu, Y. (2006). The effect of issuing biased earnings forecasts on analysts' access to management and survival. *Journal of Accounting Research, 44*(5), 965–999.

Keung, E., Lin, Z. X., & Shih, M. (2010). Does the stock market see a zero or small positive earnings surprise as a red flag? *Journal of Accounting Research, 48*(1), 105–136.

Kim, I., & Skinner, D. J. (2012). Measuring securities litigation risk. *Journal of Accounting and Economics*, *53*(1–2), 290–310.

Kim, Y., Lobo, G. J., & Song, M. (2011). Analyst characteristics, timing of forecast revisions, and analyst forecasting ability. *Journal of Banking & Finance*, *35*(8), 2158–2168.

Kim, Y., & Song, M. (2015). Management earnings forecasts and value of analyst forecast revisions. *Management Science, 61*(7), 1663–1683.

Kimbrough, M. D., & Louis, H. (2011). Voluntary disclosure to influence investor reactions to merger announcements: An examination of conference calls. *The Accounting Review, 86*(2), 637–667.

Koh, P. S. (2007). Institutional investor type, earnings management and benchmark beaters. *Journal of Accounting & Public Policy*, *26*(3), 267–299.

Kothari, S. P., Shu, S., & Wysocki, P. D. (2009). Do managers withhold bad news? *Journal of Accounting Research*, *47*(1), 241–276.

Kothari, S. P., So, E., & Verdi, R. (2016). Analysts' forecasts and asset pricing: A survey. *Annual Review of Financial Economics, 8,* 197–219.

Kraft, A., Lee, B. S., & Lopatta, K. (2014). Management earnings forecasts, insider trading, and information asymmetry. *Journal of Corporate Finance, 26*, 96–123.

Kross, W. J., Ro, B. T., & Suk, I. (2011). Consistency in meeting or beating earnings expectations and management earnings forecasts. *Journal of Accounting & Economics, 51*(1–2), 37–57.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*(3), 480.

Kurov, A. (2008). Investor sentiment, trading behavior and informational efficiency in index futures markets. *Financial Review, 43*(1), 107–127.

Lachmann, M., Wöhrmann, A., & Wömpener, A. (2011). Acquisition and integration of fair value information on liabilities into investors' judgments. *Review of Accounting & Finance, 10*(4), 385–410.

Lawrence, A., Ryans, J. P., & Sun, E. Y. (2017). Investor demand for sell-side research. *The Accounting Review, 92*(2), 123–149.

Leone, A. J., Wu, J. S., & Zimmerman, J. L. (2006). Asymmetric sensitivity of CEO cash compensation to stock returns. *Journal of Accounting & Economics, 42*(1–2), 167–192.

Leuz, C., & Schrand, C. (2009, April). *Disclosure and the cost of capital: Evidence from firms. responses to the Enron shock* [Working paper]. National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w14897/w14897.pdf

Liu, L. X., Sherman, A. E., & Zhang, Y. (2014). The long-run role of the media: Evidence from initial public offerings. *Management Science, 60*(8), 1945–1964.

Ljungqvist, A., Marston, F., Starks, L. T., Wei, K. D., & Yan, H. (2007). Conflicts of interest in sell-side research and the moderating role of institutional investors. *Journal of Financial Economics, 85*(2), 420–456.

Lourie, B. (2019). The revolving door of sell-side analysts. *The Accounting Review, 94*(1), 249–270.

Lundholm, R., & Rogo, R. (2020). Do excessively volatile forecasts impact investors? *Review of Accounting Studies, 25*(2), 636–671.

Maber, D. A., Groysberg, B., & Healy, P. M. (2020). An empirical examination of sell-side brokerage analysts' published research, concierge services, and high-touch services. *European Accounting Review, 30*(4), 1–27.

Meng, X. (2015). Analyst reputation, communication, and information acquisition. *Journal of Accounting Research, 53*(1), 119–173.

Milian, J. A. (2018). The information content of guidance and earnings. *European Accounting Review, 27*(1), 105–128.

Munezero, M., Montero, C. S., Sutinen, E., & Pajunen, J. (2014). Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE Transactions on Affective Computing*, *5*(2), 101–111.

New York Stock Exchange. (2016, 2 November). *Nonprofessional subscriber policy*. https://www.nyse.com/publicdocs/nyse/data/Policy-Non-ProfessionalSubscribers_PDP.pdf

Ng, J., Tuna, I., & Verdi, R. (2013). Management forecast credibility and underreaction to news. *Review of Accounting Studies*, *18*(4), 956–986.

Ng, J., Vasvari, F. P., & Wittenberg-Moerman, R. (2016). Media coverage and the stock market valuation of TARP participating banks. *European Accounting Review*, *25*(2), 347–371.

Nichols, D. R., & Tsay, J. J. (1979). Security price reactions to long-range executive earnings forecasts. *Journal of Accounting Research*, *17*(1), 140–155.

O'Brien, P. C. (1988). Analysts' forecasts as earnings expectations. *Journal of Accounting & Economics*, *10*(1), 53–83.

Obeng, V. A., Ahmed, K., & Cahan, S. F. (2020). Integrated reporting and agency costs: International evidence from voluntary adopters. *European Accounting Review*, *30*(4), 645–674.

Oliveira, N., Cortez, P., & Areal, N. (2013). *On the predictability of stock market behavior using stockTwits sentiment and posting volume.* Portuguese Conference on Artificial Intelligence. Berlin, Heidelberg. https://link.springer.com/chapter/10.1007/978-3-642-40669-0_31

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204.

Palter, R. N., Rehm, W., & Shih, J. (2008). Communicating with the right investors. *McKinsey Quarterly*, *2*, 64.

Parr, T., & Friston, K. J. (2019). Attention or salience? *Current Opinion in Psychology*, *29*, 1–5.

Patell, J. M. (1976). Corporate forecasts of earnings per share and stock price behavior: Empirical test. *Journal of Accounting Research*, *14*(2), 246–276.

Penman, S. H. (1980). An empirical investigation of the voluntary disclosure of corporate earnings forecasts. *Journal of Accounting Research*, *18*(1), 132–160.

Reavis, C. (2012). The global financial crisis of 2008: The role of greed, fear, and oligarchs. *MIT Sloan Management Review*, *16*, 1–22.

Renault, T. (2017). Intraday online investor sentiment and return patterns in the US stock market. *Journal of Banking & Finance*, *84*, 25–40.

Roger, T. (2018). The coverage assignments of financial analysts. *Accounting & Business Research, 48*(6), 651–673.

Rogers, J. L., Skinner, D. J., & Van Buskirk, A. (2009). Earnings guidance and market uncertainty. *Journal of Accounting & Economics, 48*(1), 90–109.

Rogers, J. L., & Stocken, P. C. (2005). Credibility of management forecasts. *The Accounting Review, 80*(4), 1233–1260.

Rogers, J. L., & Van Buskirk, A. (2013). Bundled forecasts in empirical accounting research. *Journal of Accounting & Economics, 55*(1), 43–65.

Santos, L. R., & Rosati, A. G. (2015). The evolutionary roots of human decision making. *Annual Review of Psychology, 66*, 321.

Schafhäutle, S., & Veenman, D. (2021, 29 January). *Crowdsourced earnings expectations and the market reaction to street earnings surprises* [Working paper]. University of Amsterdam Business School. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3444144

Schmeling, M. (2009). Investor sentiment and stock returns: Some international evidence. *Journal of Empirical Finance, 16*(3), 394–408.

Securities and Exchange Commission. (2010). *Analyzing analyst recommendations*. https://www.sec.gov/tm/reportspubs/investor-publications/investorpubsanalystshtm.html

Seybert, N., & Yang, H. I. (2012). The party's over: The role of earnings guidance in resolving sentiment-driven overvaluation. *Management Science, 58*(2), 308–319.

Shanthikumar, D. M. (2012). Consecutive earnings surprises: Small and large trader reactions. *The Accounting Review, 87*(5), 1709–1736.

Shefrin, H. (2002). *Beyond greed and fear: Understanding behavioral finance and the psychology of investing*. Oxford University Press on Demand.

Shroff, P. K., Venkataraman, R., & Xin, B. (2014). Timeliness of analysts' forecasts: The information content of delayed forecasts. *Contemporary Accounting Research, 31*(1), 202–229.

Simpson, A. (2013). Does investor sentiment affect earnings management? *Journal of Business Finance & Accounting, 40*(7–8), 869–900.

Souder, D., & Bromiley, P. (2017). Timing for dollars: How option exercisability influences resource allocation. *Journal of Management, 43*(8), 2555–2579.

Stickel, S. E. (1990). Predicting individual analyst earnings forecasts. *Journal of Accounting Research, 28*(2), 409–417.

Stickel, S. E. (1991). Common stock returns surrounding earnings forecast revisions: More puzzling evidence. *The Accounting Review*, *66*(2), 402–416.

Stol, K.-J., Caglayan, B., & Fitzgerald, B. (2017). Competition-based crowdsourcing software development: A multi-method study from a customer perspective. *IEEE Transactions on Software Engineering, 45*(3), 237–260.

Stuart, A. C., Bedard, J. C., & Clark, C. E. (2021). Corporate social responsibility disclosures and investor judgments in difficult times: The role of ethical culture and assurance. *Journal of Business Ethics, 171*(3), 565–582.

Sul, E. (2020, 22 November). *Takeover threats, job security concerns, and earnings management* [Working paper]. George Washington University. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3034948

Tadesse, A. F., & Murthy, U. S. (2018). Nonprofessional investor perceptions of the partial remediation of IT and non-IT control weaknesses: An experimental investigation. *International Journal of Accounting Information Systems, 28*, 14–30.

Tahir, M., Ibrahim, S., & Nurullah, M. (2019). Getting compensation right—the choice of performance measures in CEO bonus contracts and earnings management. *The British Accounting Review, 51*(2), 148–169.

Tang, V. W. (2018). Wisdom of crowds: Cross-sectional variation in the informativeness of third-party-generated product information on Twitter. *Journal of Accounting Research, 56*(3), 989–1034.

Tauni, M. Z., Yousaf, S., & Ahsan, T. (2020). Investor-advisor Big Five personality similarity and stock trading performance. *Journal of Business Research, 109*, 49–63.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance, 62*(3), 1139–1168.

Thayer, J. (2011). Determinants of investors' information acquisition: Credibility and confirmation. *The Accounting Review, 86*(1), 1–22.

Twedt, B. (2016). Spreading the word: Price discovery and newswire dissemination of management earnings guidance. *The Accounting Review, 91*(1), 317–346.

Veenman, D., & Verwijmeren, P. (2018). Do investors fully unravel persistent pessimism in analysts' earnings forecasts? *The Accounting Review, 93*(3), 349–377.

Versano, T., & Trueman, B. (2017). Expectations management. *The Accounting Review, 92*(5), 227–246.

Walther, B. R., & Willis, R. H. (2013). Do investor expectations affect sell-side analysts' forecast bias and forecast accuracy? *Review of Accounting Studies, 18*(1), 207–227.

Wang, H. (2009). Reputation acquisition of underwriter analysts–theory and evidence. *Journal of Applied Economics, 12*(2), 331–363.

Waymire, G. (1984). Additional evidence on the information content of management earnings forecasts. *Journal of Accounting Research*, *22*(2), 703–718.

Weißofner, F., & Wessels, U. (2020). Overnight returns: An international sentiment measure. *Journal of Behavioral Finance, 21*(2), 205–217.

Xiong, X., Meng, Y., Joseph, N. L., & Shen, D. (2020). Stock mispricing, hard-to-value stocks and the influence of internet stock message boards. *International Review of Financial Analysis, 72*, 101576.

Yang, H. I. (2012). Capital market consequences of managers' voluntary disclosure styles. *Journal of Accounting & Economics, 53*(1–2), 167–184.

Zhang, L. (2012). The effect of ex ante management forecast accuracy on the post-earnings-announcement drift. *The Accounting Review, 87*(5), 1791–1818.

Zhang, X. F. (2006). Information uncertainty and analyst forecast behavior. *Contemporary Accounting Research*, *23*(2), 565–590.

Appendices

Appendix A StockTwits – Examples of Messages and Their VADER Sentiment Scores

Date & time	Tweet content	Negative	Neutral	Positive	Compound
2017-01-27 21:38:52	$X otherwise why they wouldnt keep the shares if they felt X is going to 40, or even 50 and 60s.	0.199	0.881	0.000	–0.224
2017-01-27 21:38:58	$FEYE Mandia had a nervous breakdown!!! Sign of weakness	0.490	0.510	0.000	–0.698
2017-01-27 21:39:10	Sensex Trades on a Positive Note; Power Stocks Gain $INDY $EPI $PIN $INDA	0.000	0.611	0.389	0.791
2017-01-27 21:39:11	GDPhriday - Trump Declares War (Trade) on Mexico! $SPX Also $CVX $AMZN $NFLX	0.259	0.741	0.000	–0.636
2017-01-27 21:39:16	$MHLD on the move. < it looks great + 4 days of short intr	0.000	0.760	0.240	0.625
2017-01-27 21:39:22	$SFM: New Insider Filing on Director JAMES DOUGLAS SANDERS:	0.000	1.000	0.000	0.000
2017-01-27 21:39:46	$feye low of day and week...tanking in afterhours	0.231	0.769	0.000	–0.273
2017-01-27 21:40:00	$VRX yeah just fine	0.000	0.333	0.667	0.459
2017-01-27 21:40:14	$VRX Short play like BK next month.Relax,hold your @ and laugh them in the face,no BK till 2020,after good […]	0.000	0.505	0.495	0.599
2017-01-27 21:40:20	$PCLN Yeahhhhhh baby closed with 20k in 2 days congrats to everyone!!!!!	0.000	0.686	0.314	0.678
2017-01-27 21:41:20	$ASNA There are no more normal, holding sharehlders in this stock. Company CEO has lost $140 million stock value since 2015 [...]	0.306	0.694	0.000	–0.296
2017-01-27 21:41:27	$VRX big drop in the IV on February calls today	0.189	0.811	0.000	–0.273
2017-01-27 21:42:27	$ENDP TPG will take this private IMO	0.000	1.000	0.000	0.000
2017-01-27 21:42:29	$GPRO We need ONE good quarter to take us to $15-$16	0.000	0.775	0.225	0.440
2017-01-27 21:43:22	$ASNA This is the world's most crazy stock. Today another example of the way the stock is jacked […]	0.236	0.667	0.097	–0.636
2017-01-27 21:43:28	$VRX wow beers are still not satisfied there will be time for bulls but charts is also against us.	0.079	0.807	0.114	0.186

Note: This table shows the tweets posted on StockTwits about focal firms between 21:38:52 and 21:43:52 on 27 January 2017. Negative, neutral, positive, and compound represent the assigned VADER negative, neutral, positive, and compound score accordingly.

Appendix B Variable Definition—NPI Sentiment and Management Guidance

Variable	Definition and measurement
Panel A: Variables used in the main analyses	
CTO[0,+1]	A proxy for NPI reaction to guidance, comprised of the overnight stock return (close-to-open price) measured as the natural logarithm of the ratio of closing price on guidance date and opening price one day after guidance date.
Guide	An indicator variable for 'bundled' earning guidance issuance that equals 1 if the firm provided an earnings guidance during the five-day window centred on the guidance issuance, and 0 otherwise.
GuidePos	An indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise.
NPISent	A measure of NPI sentiment equal to the difference between the number of positive and negative tweets observed in the 30-day window prior to the issuance of guidance.
NewsSent	A proxy for the sentiment of all actual and potential investors in a firm, measured as the difference between the number of positive and negative news articles concerning the firm in the 30-day window prior to the issuance of guidance.
TweetCount	The total number of tweets discussing a firm (in a 30-day window prior to the guidance issuance).
NewsCount	The total number of news articles discussing a firm (in a 30-day window prior to the guidance issuance).
MktSent	Market sentiment, proxied by the Michigan Consumer Sentiment Index.
Surprise	Earnings surprise equals the actual earnings minus the prevailing median analyst estimate, deflated by stock price three trading days prior to the guidance issuance.
Loss	An indicator of loss making that equals 1 if reported earnings is negative, and 0 otherwise.
PriorRet	The stock return over the 90-day period ending three trading days prior to the guidance issuance.
MarketCap	The market value of equity (i.e., stock price multiplied by number of stocks outstanding) measured at three trading days prior to the guidance issuance.
AnalystCov	Analyst coverage measured as the number of analysts with outstanding estimates three trading days prior to the guidance issuance.
MeetBeat	The proportion of the previous four quarters in which firms' reported earnings met or exceeded analysts' prevailing median consensus estimates.

Variable	Definition and measurement
LitiRisk	A measure of litigation risk estimated from the probit model of Kim and Skinner (2012), which regresses an indicator of a class action lawsuit filing during the fiscal year (Stanford Litigation Database) against (i) an indicator of litigious industries (biotech, computer, electronics, or retail), (ii) total assets, (iii) sales growth, (iii) annual market-adjusted return, (iv) returns skewness, (v) returns volatility, and (vi) trading volume. *LitiRisk* is the predicted value from the probit regression.
AnalystDisp	The standard deviation of prevailing analysts' estimates for the current period's earnings.
RetVol	Stock return volatility, measured as the standard deviation of daily stock returns over the 90 days prior to the guidance issuance.
MtB	Market-to-book ratio.
IdioVol	Average implied volatility (for a 30-day window, at the money option) measured at the release date of guidance issuance.
MultiGuide	An indicator variable that equals 1 if the firm previously provided earnings guidance for the current quarter's earnings, and 0 otherwise.
GuidePrior	An indicator variable that equals 1 if the firm issued an earnings guidance during the five-day window centred on the guidance issuance last quarter, and 0 otherwise.
RecentGuider	An indicator that equals 1 if the firm is a guiding firm, as measured by the presence of at least three guidance issues in the prior 12 quarters, and 0 otherwise.
InsiderTrade	The total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the current quarter.
InsiderTradePost	The total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the 15 days after the earnings announcement.
VIX	Chicago Board Option's Exchange implied volatility index (a.k.a. the 'fear index') during the three-day window centred on the guidance issuance.
Panel B: Variables used in additional analyses	
SECAlert	An indicator that equals 1 if an earnings guidance is issued after the date of the SEC investor alert, and 0 otherwise.
NPISent4Q	The average of *NPISent* for the prior four quarters.
SalesGrowth	Sales growth measured as the change in quarterly sales deflated by total assets.
TradeVolume	The annual trading volume deflated by the beginning-of-the-year outstanding shares.

Variable	Definition and measurement
LitiIndustry	An indicator that equals 1 if the firm is in an industry with high litigation risk (i.e., four-digit SIC code 2833–2836, 3570–3577, 3600–3675, 5200–5961, and 7370–7374), and 0 otherwise.
NPISentPred	The predicted NPI sentiment by the 2SLS approach used in Chapter 3.6.1.3.
Disagree	An indicator that equals 1 if NPI sentiment and market sentiment disagree (i.e., high (low) NPI sentiment and low (high) market sentiment), and 0 otherwise.
HighMktSent	An indicator that equals 1 if *MCIS* is above the median value, and 0 otherwise.
StockOptions	The accumulated value of exercisable options held by directors and officers at the end of the previous fiscal year.
StandalonePos	An indicator that equals 1 if the standalone guidance estimate is greater than the pre-forecast prevailing median analyst estimate, and 0 otherwise.
PihTra	The percentage of institutional ownership by transient institutional investors, which is defined by Bushee (1998).
PihDed	The percentage of institutional ownership by dedicated institutional investors, which is defined by Bushee (1998).
PihQix	The percentage of institutional ownership by quasi-index institutional investors, which is defined by Bushee (1998).
Panel C: Variables used in robustness checks	
IMR	The inverse Mills ratio obtained from the first-stage model of the Heckman two-stage model in Chapter 3.5.3.
ΔIdioVolPrior	The natural logarithm of the ratio of implied volatility measured at the close of day prior to the report date of quarterly earnings to implied volatility measured 15 days prior to the report date of earnings.
InstConcent	The aggregate percentage of a firm's shares held by the five largest institutional investors at the beginning of the fiscal quarter.
MktRet	Market returns measured as the CRSP Value-weighted Index return including distributions.
Def	The difference between the yields to maturity on BBB- and AAA-rated bond yields.
Yld	The yield on the three-month Treasury bill.
GDP	Gross Domestic Product (GDP) growth measured as 100 times the quarterly change in the natural logarithm of chained (1996) GDP.

Variable	Definition and measurement
Cons	Consumption growth measured as 100 times the quarterly change in the natural logarithm of personal consumption expenditures.
Labor	Labor income growth measured as 100 times the quarterly change in the natural logarithm of labor income computed as total personal income minus dividend income per capita, and deflated by the Personal Consumption Expenditure (PCE) Deflator.
URate	The average monthly unemployment rate, as reported by US Bureau of Labor Statistics.
CPIQ	The Consumer Price Index inflation rate.
CAY	Consumption-to-wealth ratio obtained from the website of Sydney Ludvigson, Professor of Economics at New York University (NYU) (https://www.sydneyludvigson.com/data-and-appendixes/).
NPISentResid1	The predicted NPI sentiment residuals from the two-stage approach in Chapter 3.7.2.
NPISentResid2	The predicted NPI sentiment residuals from the two-stage approach in Chapter 3.7.3.

Appendix C Variable Definition—Investor Earnings Expectations and Analyst Forecasts

Variable	Definition and measurement
Panel A: Variables used in the main analyses	
$Revision_{j,t}$	Analyst EPS forecast consensus [−30, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus the analyst EPS forecast consensus [−60, −31] days before the actual earnings announcement for firm *j* in quarter *t*, deflated by the stock price at the end of the previous quarter. Analyst EPS forecast consensus [−60, −31] *([−30, −1])* is measured as the average of all analysts' forecast EPS issued [−60, −31] *([−30, −1])* days before the actual earnings announcement.
$RevDown_{j,t}$	A binary variable that equals 1 if *$Revision_{j,t}$* is less than 0; 0 otherwise.
$Error_{j,t}$	The error in the analyst EPS forecast consensus, measured as the analyst EPS forecast consensus [−30, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus the actual EPS, deflated by the stock price at the end of the previous quarter.
$Pessimism_{j,t}$	A binary variable that equals 1 if *$Error_{j,t}$* is less than 0; 0 otherwise.
$PessimismMag_{j,t}$	The magnitude of analyst forecast pessimism, measured as *$Pessimism_{j,t}$* times the absolute value of *$Error_{j,t}$*.
$EstmzRev_{j,t}$	The Estimize EPS forecast consensus [−60, −31] days before the actual earnings announcement for firm *j* in quarter *t* minus the Estimize EPS forecast [−90, −61] days before the actual earnings announcement for firm *j* in quarter *t*, deflated by the stock price at the end of the previous quarter. The Estimize EPS forecast consensus [−90, −61] *([−60, −31])* is measured as the average of daily Estimize EPS consensus made [−90, −61] *([−60, −31])* days before the actual earnings announcement.
$DownFollow_{j,t}$	A binary variable that equals 1 if both *$Revision_{j,t}$* and *$EstmzRev_{j,t}$* are less than 0; 0 otherwise.
$DownAgainst_{j,t}$	A binary variable that equals 1 if *$Revision_{j,t}$* is less than 0 while *$EstmzRev_{j,t}$* is greater than (or equal to) 0; 0 otherwise.
$DownFollowMag_{j,t}$	The magnitude of the analyst downward revision when *$EstmzRev_{j,t}$* is less than 0, measured as *$DownFollow_{j,t}$* times the absolute value of *$Revision_{j,t}$*.
$DownAgainstMag_{j,t}$	The magnitude of the analyst downward revision when *$EstmzRev_{j,t}$* is greater than (or equal to) 0, measured as *$DownAgainst_{j,t}$* times the absolute value of *$Revision_{j,t}$*.
$FirmSize_{j,t}$	The natural logarithm of market capitalization for firm *j* computed as share price times total shares outstanding as of the end of the fiscal year before the earnings announcement date for firm *j*.

Variable	Definition and measurement
$BtM_{j,t}$	The book value of equity for the most recent fiscal year before the earnings announcement date, scaled by market capitalization as of the end of the same fiscal year for firm *j*.
$Turnover_{j,t}$	Average daily turnover is defined as share volume scaled by shares outstanding in the calendar year before the earnings announcement date for firm *j*.
$Loss_{j,t}$	A loss dummy for firm *j*, set to 1 if firm *j* generated a net loss in the most recent fiscal year before the earnings announcement date; 0 otherwise.
$Leverage_{j,t}$	Leverage for firm *j*, defined as the book value of long-term debt at the most recent fiscal year before the earnings announcement date, deflated by total assets.
$RetVol_{j,t}$	The standard deviation of daily returns over the calendar year prior to the earnings announcement date for firm *j*.
$Guidance_{j,t}$	A binary variable equal to 1 if firm *j* issues earnings guidance for quarter *t*.
$Follow_{j,t}$	The number of analysts following firm *j* in quarter *t* [−30,−1] days before the earnings announcement date.
$Dispersion_{j,t}$	The standard deviation in the analyst forecasts for firm *j* in quarter *t* [−30,−1] days before the earnings announcement date.
Panel B: Variables used in additional analyses	
$Revision_{i,j,t}$	Analyst *i*'s EPS forecast [−30, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus analyst *i*'s EPS forecast [−60, −31] days before the actual earnings announcement for firm *j* in quarter *t*, deflated by the stock price at the end of the previous quarter.
$RevDown_{i,j,t}$	A binary variable that equals 1 if $Revision_{i,j,t\ t}$ is less than 0; 0 otherwise.
$Error_{i,j,t}$	The error of analyst *i*'s EPS forecast, measured as analyst *i*'s EPS forecast [−30, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus the actual EPS, deflated by the stock price at the end of the previous quarter.
$Pessimism_{i,j,t}$	A binary variable that equals 1 if $Error_{i,j,t}$ is less than 0; 0 otherwise.
$PessimismMag_{i,j,t}$	The magnitude of analyst *i*'s forecast pessimism, measured as $Pessimism_{i,j,t}$ times the absolute value of $Error_{i,j,t}$.
$Revision_{i,j,t}$	Analyst *i*'s EPS forecast [−30, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus analyst *i*'s EPS forecast [−60, −31] days before the actual earnings announcement for firm *j* in quarter *t*, deflated by the stock price at the end of the previous quarter.
$DownFollow_{i,j,t}$	A binary variable that equals 1 if both $Revision_{i,j,t}$ and $EstmzRev_{j,t}$ are less than 0; 0 otherwise.
$DownAgainst_{i,j,t}$	A binary variable that equals 1 if $Revision_{i,j,t}$ is less than 0 while $EstmzRev_{j,t}$ is greater than (or equal to) 0; 0 otherwise.

Variable	Definition and measurement
DownFollowMag$_{i,j,t}$	The magnitude of the analyst downward revision when *EstmzRev*$_{j,t}$ is less than 0, measured as *DownFollow*$_{i,j,t}$ times the absolute value of *Revision*$_{i,j,t}$.
DownAgainstMag$_{i,j,t}$	The magnitude of the analyst downward revision when *EstmzRev*$_{j,t}$ is greater than (or equal to) 0, measured as *DownAgainst*$_{i,j,}$ times the absolute value of *Revision*$_{i,j,t}$.
AccScore$_{i,j,t}$	The measure of relative forecast accuracy for analyst *i*'s EPS forecast [−60, −31] days before the actual earnings announcement for firm ***j*** in quarter ***t***. Analysts are ordered based on their absolute forecast errors (AFE). The analyst with the lowest AFE receives the first rank (*AccuRank*$_{i,j,t}$); the analyst with the second-lowest AFE receives the second rank; and so on. Analysts with the same AFE are assigned the same rank (the midpoint value of their ranks). The ranks are then transformed into scores to account for differences in the number of analysts covering the different firms. The scores are obtained by applying the following formula: $AccScore_{i,j,t} = 100 - [(AccRank_{i,j,t} - 1)/(M_{j,t} - 1)]*100$ where $M_{j,t}$ is the number of analysts who issue a forecast [−60, −31] days before the actual earnings announcement for firm ***j*** in quarter ***t***.
BoldScore$_{i,j,t}$	The measure of relative forecast boldness for analyst *i*'s EPS forecast [−60, −31] days before actual earnings announcement for firm ***j*** in quarter ***t***: $BoldScore_{i,j,t} = 100 - [(BoldRank_{i,j,t} - 1)/(M_{j,t} - 1)]*100$ where *BoldRank*$_{i,j}$ is the rank of analyst ***i***'s forecast deviation from the average of EPS forecasts made by all other analysts, $\lvert ForecastEPS_{i,j,t} - AvgForecastEPS_{j,t} \rvert$. The process to rank analysts according to their forecast boldness is similar to the one used for the measure of relative forecast accuracy.
BiasScore$_{i,j,t}$	The measure of relative forecast bias for analyst *i*'s EPS forecast [−60, −31] days before the actual earnings announcement for firm ***j*** in quarter ***t***: $BiasScore_{i,j,t} = 100 - [(BiasRank_{i,j,t} - 1)/(M_{j,t} - 1)]*100$ where *BiasRank*$_{i,j,t}$ is the rank of analyst *i*'s forecast error.. The process to rank analysts according to their forecast bias is similar to the one used for the measure of relative forecast accuracy.
FirmExp$_{i,j,t}$	Analyst *i*'s firm-specific experience, measured by counting the number of quarters to date in which the analyst has followed the firm in I/B/E/S.
IndusExp$_{i,j,t}$	Analyst *i*'s industry-specific experience, measured by counting the number of quarters to date in which the analyst has followed the industry in I/B/E/S.
GenExp$_{i,j,t}$	Analyst *i*'s general experience, measured by counting the number of quarters to date during which the analyst has issued an earnings forecast (for this or any other firm) in I/B/E/S.

Variable	Definition and measurement
BrokerSize$_{i,t}$	Analyst *i*'s brokerage house size, measured as the number of analysts at the brokerage house that employs analyst *i* in the calendar year during which the analyst make an EPS forecast.
Error90d$_{j,t}$	The error in the analyst EPS forecast consensus, measured as the analyst EPS forecast consensus [−90, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus the actual EPS, deflated by the stock price at the end of the previous quarter.
Return$_{j,t}$	Annual stock return at year prior to the earnings announcement date, adjusted for contemporaneous annual market return.
Post$_t$	A binary variable that equals 1 for the period from 2012 to 2018 and 0 for the periods between 2009 to 2011.
Treat$_j$	A binary variable that equals 1 for treated firms and 0 for control firms.
GdpChg$_{j,t}$	The absolute value of quarterly change in the seasonal growth rate in GDP. GDP data are obtained from the Federal Reserve Bank of St. Louis web site: https://fred.stlouisfed.org/searchresults/?st=gdp&isTst=1
Miss$_{j,t}$	A binary variable set to 1 if a firm fails to meet the consensus forecast in a given quarter; 0 otherwise.
ROA$_{j,t}$	Earnings before extraordinary items scaled by lagged total assets.
PublicPrc$_{j,t}$	The proportion of public information to total information in analyst forecasts.
TotalInfo$_{j,t}$	The precision of total information: $= \frac{(D_{j,t})}{(SE_{j,t}-\frac{D_{j,t}}{N_{j,t}}+D_{j,t})^2} + \frac{(SE_{j,t}-\frac{D_{j,t}}{N_{j,t}})}{(SE_{j,t}-\frac{D_{j,t}}{N_{j,t}}+D_{j,t})^2}$ $D_{j,t}$ is the realised forecast dispersion $= \frac{1}{N-1}\sum_{i=1}^{N}(F_{i,j,t}-\bar{F}_{j,t})^2$ where $F_{i,j,t}$ refer to analyst *i*'s forecast for firm *j*; $\bar{F}_{j,t}$ refers to the mean of all analyst forecasts for firm *j*; and *N* is the number of analyst following. $SE_{j,t}$ is the square error in the mean forecast $= (Actual_{j,t}-\bar{F}_{j,t})^2$ where $Actual_{j,t}$ is the actual EPS for firm *j*.
PrivateInfo$_{j,t}$	The precision of private information $= \frac{(D_{j,t})}{(SE_{j,t}-\frac{D_{j,t}}{N_{j,t}}+D_{j,t})^2}$
PublicInfo$_{j,t}$	The precision of public information $= \frac{(SE_{j,t}-\frac{D_{j,t}}{N_{j,t}})}{(SE_{j,t}-\frac{D_{j,t}}{N_{j,t}}+D_{j,t})^2}$

Panel C: Variables used in robustness analyses

Variable	Definition and measurement
MeetBeat $_{,j,t}$	A binary variable that equals 1 if the raw forecast error (i.e., analyst EPS forecast consensus [−30, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus the actual EPS) is [−0.02, 0]; 0 otherwise.
MeetBeat $_{i,j,t}$	A binary variable that equals 1 if the raw forecast error (i.e., analyst *i*'s EPS forecast [−30, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus the actual EPS) is [−0.02, 0]; 0 otherwise.

Appendix D Additional Analysis and Robustness Check—NPI Sentiment and Management Guidance

Table A1 NPI reaction to positive guidance including *GuideSurp*

	Dependent variable = *CTO[0,+1]*			
	Column I		Column II	
Variable	**coef.**	**t-stat.**	**coef.**	**t-stat.**
GuidePos	0.055	1.144	0.055	1.137
GuideSurp	0.024	2.462**	0.028	3.027**
GuidePos× GuideSurp	−0.006	−0.636	−0.004	−0.516
NPISent			−0.041	−0.823
GuidePos×NPISent			0.030	3.054***
GuideSurp×NPISent			−0.025	−2.059**
GuidePos× GuideSurp×NPISent			−0.011	−0.976
NewsSent	0.006	0.684	0.006	0.710
GuidePos×NewsSent	−0.005	−0.614	−0.006	−0.718
GuideSurp×NewsSent	−0.010	−1.163	−0.011	−1.287
GuidePos× GuideSurp×NewsSent	−0.001	−0.117	0.005	0.602
TweetCount			0.015	0.343
NewsCount	0.005	0.623	0.008	0.896
GuidePos× NewsCount	−0.022	−3.005***	−0.025	−3.417***
MktSent	0.172	1.026	0.170	1.015
GuidePos× MktSent	0.003	0.071	0.002	0.035
Surprise	0.143	12.418***	0.143	12.420***
GuidePos× Surprise	0.034	2.780***	0.034	2.786***
Loss	−0.002	−0.200	−0.002	−0.221
PriorRet	0.014	1.911*	0.015	1.938*
MarketCap	0.004	0.545	0.005	0.774
AnalystCov	−0.006	−0.638	−0.005	−0.577
MeetBeat	0.089	13.169***	0.089	13.144***
LitRisk	−0.035	−2.601***	−0.035	−2.581**
AnalystDisp	−0.009	−1.420	−0.010	−1.458
RetVol	0.016	1.313	0.018	1.426
MtB	0.005	0.696	0.004	0.665
Industry FE	Yes		Yes	
Yearr-Quarter FE	Yes		Yes	
Observations	23,061		23,061	
F statistics	8,295		8.090	
Prob > F	0.000		0.000	
Adjusted R^2	0.048		0.049	

Note: Table A1 presents standardized coefficients and t-statistics from regressions of NPI reaction on *GuideSurp*, its two-way and three-way interactions with *GuidePos* and *NPISent*, and relevant controls. *GuideSurp* is measured as the guidance estimate minus the prevailing median analyst estimate, deflated by stock price three trading days prior to the guidance issuance. *CTO[0,+1]*, an NPI reaction proxy, is the overnight stock return (close-to-open price) measured as the natural logarithm of the ratio of closing price on the guidance date and opening price one day after the guidance date; *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *NPISent* is the number of net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the number of net positive news articles, measured as the number of positive news articles minus the number of negative news articles (in a 30−day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing a firm (in a 30−day window prior to the guidance issuance); *MktSent* is the Michigan Consumer Sentiment Index; *Surprise* measures earnings surprise that equals the actual earnings minus the prevailing median analyst estimate, deflated by stock price three trading days prior to the guidance issuance; *Loss* is an indicator of loss making that equals 1 if reported earnings is negative, and 0 otherwise; *PriorRet* measures the stock return over the 90-day period ending three trading days prior to the guidance issuance; *MarketCap* is the market value of equity (i.e., stock price multiplied by number of stocks outstanding) measured at three trading days prior to the guidance issuance; *AnalystCov* is measured as the number of analysts with outstanding estimates three trading days prior to the guidance issuance; *MeetBeat* is the proportion of the previous four quarters in which firms' reported earnings met or exceeded analysts' prevailing median consensus estimates; *LitiRisk* is the measure of litigation risk estimated from the probit model by Kim and Skinner (2012); *AnalystDisp* measures the standard deviation of prevailing analyst estimates for the current period's earnings; *RetVol* measures the standard deviation of daily stock returns over the 90 days prior to the guidance issuance; *MtB* is the market-to-book ratio. Industry-fixed effects based on the Fama–French 48 classification and fiscal quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

Table A2 Additional analyses: Regressions on sub-samples defined by insider trading and institutional shareholding percentage

	Dependent variable = ***GuidePos***							
	Distinguishing the opportunistic and informative views				**Do StockTwits capture NPI-specific sentiment?**			
	Column I: High insider trading		**Column II: Low insider trading**		**Column III: High institutional percentage**		**Column IV: Low institutional percentage**	
Variable	**coef.**	**t-stat.**	**coef.**	**t-stat**	**coef.**	**t-stat.**	**coef.**	**t-stat.**
NPISent	0.059	4.228***	0.019	1.489	0.025	1.687*	0.044	3.356***
NewsSent	0.015	5.700***	0.011	4.614***	0.016	5.977***	0.010	4.155***
TweetCount	−0.010	−3.342***	−0.003	−1.007	0.003	0.769	−0.009	−3.279**
NewsCount	0.005	4.213***	0.010	9.213***	0.005	4.009***	0.009	8.209***
MktSent	0.104	1.103	−0.040	−0.269	0.020	1.418	0.030	0.374
IdioVol	−0.446	−2.751***	−0.724	−4.966***	−0.656	−3.560***	−0.493	−3.568***
MultiGuide	−0.475	−10.335***	−0.413	−9.075***	−0.472	−9.954***	−0.422	−9.408***
GuidePrior	0.049	0.732	0.013	0.215	0.068	0.993	0.005	0.079
InsideTrade	0.096	5.033***	0.027	0.973	0.100	4.320***	0.061	2.918***
InsideTradePost	−0.026	−0.911	2.333	1.447	0.008	0.205	0.024	0.658
VIX	0.013	0.279	0.077	1.161	0.081	0.729	0.038	1.021
Surprise	27.478	4.880***	22.758	5.602***	35.562	6.001***	18.830	4.713***
Loss	−0.529	−5.113***	−0.371	−3.982***	−0.416	−3.451***	−0.473	−5.492***
PriorRet	0.161	1.820	−0.077	−0.863	0.048	0.506	0.018	0.211
MarketCap	0.000	−0.022	−0.005	−3.160***	0.005	2.240**	−0.005	−3.201***
AnalystCov	0.004	0.984	−0.001	−0.224	−0.003	−0.748	0.005	1.227
MeetBeat	0.766	9.944***	0.635	8.574***	0.604	7.729***	0.778	10.391***
LitiRisk	15.907	6.268***	11.354	5.000***	15.705	5.592***	11.651	5.350***
AnalystDisp	−2.002	−2.654***	−0.645	−1.016	−1.442	−2.030**	−1.122	−1.655*
Industry FE	Yes		Yes		Yes		Yes	
Quarter FE	Yes		Yes		Yes		Yes	
Observations	11,920		12,737		11,424		12,868	
LR chi^2	1017.005		1016.529		747.188		1276.769	
Prob > chi^2	0.000		0.000		0.000		0.000	
Pseudo R^2	0.064		0.061		0.049		0.076	
Chi^2 for difference in coefficient of ***NPISent*** between subsamples				4.312**				0.880

Note: Columns I and II of Table A2 present results from regressing the incidence of positive guidance on NPI sentiment and controls for sub-samples of high and low insider trading among senior management, respectively. Sub-samples of high and low senior management option holdings are determined by the industry-median values of the level of insider trading among senior management during the 15 days after the guidance issuance date (*InsiderTradePost*). Columns III and IV of Table A2 present results from regressing the incidence of positive guidance on NPI sentiment and controls for sub-samples of high and low institutional ownership percentage, respectively. The proportionate institutional shareholdings (*InstHold*) is measured by the number of common shares held by institutional investors divided by the total common shares outstanding. *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the net positive news articles, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing a firm (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan Consumer Sentiment Index. Other variables are defined in Appendix B. Industry-fixed effects based on Fama–French 48 classification and fiscal-quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

Table A3 Descriptive statistics for standalone guidance—partitioned based on *GuidePos*

	Full sample (n = 17,736)			***GuidePos* = 1 (n = 3,979)**			***GuidePos* = 0 (n = 13,757)**			**Difference**	
	Mean	Median	Std.dev	Mean	Median	Std.dev.	Mean	Median	Std.dev.	Mean	Median
NPISent	0.722	0.000	1.434	0.513	0.000	1.271	0.783	0.000	1.472	***	***
NewsSent	2.177	1.000	8.351	1.225	0.000	7.996	2.453	1.000	8.431	***	***
TweetCount	4.389	3.000	4.903	3.426	2.000	4.516	4.668	3.000	4.974	***	***
NewsCount	13.249	6.000	21.154	11.570	4.000	19.442	13.734	6.000	21.600	***	***
MktSent	76.545	75.300	11.093	74.134	73.600	11.894	77.243	76.200	10.750	***	***
IdioVol	0.432	0.397	0.187	0.477	0.441	0.199	0.419	0.384	0.181	***	***
MultiGuide	0.966	1.000	0.180	0.955	1.000	0.208	0.970	1.000	0.171	***	***
InsideTrade	0.449	0.000	1.137	0.398	0.000	1.144	0.464	0.000	1.135	***	***
InsideTradePost	0.185	0.000	0.673	0.164	0.000	0.658	0.191	0.000	0.678	**	***
VIX	21.048	18.050	10.236	24.215	19.340	13.695	20.132	17.590	8.781	***	***
PriorRet	0.037	0.033	0.241	−0.020	−0.024	0.263	0.054	0.046	0.231	***	***
MarketCap	0.713	0.186	1.644	0.500	0.120	1.350	0.775	0.214	1.715	***	***
AnalystCov	11.941	11.000	7.090	10.042	8.000	6.397	12.490	11.000	7.186	***	***
MeetBeat	0.413	0.500	0.333	0.452	0.500	0.340	0.402	0.250	0.330	***	***
LitiRisk	0.031	0.023	0.039	0.028	0.021	0.035	0.031	0.024	0.039	***	***
AnalystDisp	0.020	0.010	0.023	0.025	0.020	0.029	0.018	0.010	0.021	***	***

Note: All variables are defined in Appendix B; *, **, *** denotes instances where the samples differ significantly at the 10%, 5%, 1% level for two−tailed tests.

Table A4 Correlation matrix for standalone guidance

Variable		**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
GuidePos	(1)	1.00	**−0.09**	**−0.07**	**−0.13**	**−0.06**	**−0.11**	**0.13**	**−0.04**	**−0.06**
NPISent	(2)	**−0.08**	1.00	**0.15**	**0.55**	**0.20**	**0.27**	**−0.15**	0.00	**0.07**
NewsSent	(3)	**−0.06**	**0.14**	1.00	**0.19**	**0.29**	**0.16**	**−0.21**	**0.02**	**0.08**
TweetCount	(4)	**−0.11**	**0.64**	**0.14**	1.00	**0.37**	**0.60**	**−0.32**	0.01	**0.09**
NewsCount	(5)	**−0.04**	**0.19**	**0.09**	**0.29**	1.00	**0.12**	**−0.30**	**0.04**	0.02
MktSent	(6)	**−0.12**	**0.24**	**0.11**	**0.49**	0.01	1.00	**−0.35**	0.01	**0.06**
IdioVol	(7)	**0.13**	**−0.11**	**−0.17**	**−0.22**	**−0.22**	**−0.33**	1.00	**−0.02**	**−0.04**
MultiGuide	(8)	**−0.04**	0.00	0.01	0.01	0.02	0.02	**−0.02**	1.00	**0.03**
InsideTrade	(9)	**−0.02**	**0.03**	**0.03**	0.01	**−0.07**	−0.01	**0.02**	0.01	1.00
InsideTradePost	(10)	−0.02	0.00	0.00	0.00	**−0.06**	−0.01	0.02	0.01	**0.55**
VIX	(11)	**0.17**	**−0.19**	**−0.17**	**−0.34**	0.01	**−0.63**	**0.37**	−0.01	**−0.03**
PriorRet	(12)	**−0.13**	**0.12**	**0.20**	**0.08**	−0.01	**0.11**	**−0.13**	**0.04**	**0.12**
MarketCap	(13)	**−0.07**	**0.23**	0.13	**0.31**	**0.65**	**0.08**	**−0.32**	0.01	**−0.08**
AnalystCov	(14)	**−0.14**	**0.26**	**0.10**	**0.39**	**0.43**	**0.07**	**−0.33**	0.04	**−0.03**
MeetBeat	(15)	**0.06**	**−0.03**	**0.03**	**−0.04**	**0.08**	**0.05**	**−0.11**	−0.01	**−0.02**
LitiRisk	(16)	**−0.03**	**0.22**	0.01	**0.33**	**0.17**	**0.08**	**0.03**	0.00	0.01
AnalystDisp	(17)	**0.13**	0.02	**−0.03**	**0.04**	**0.09**	**−0.02**	0.00	**−0.13**	**−0.07**

Variable		**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**	**(16)**	**(17)**
GuidePos	(1)	**−0.03**	**0.11**	**−0.13**	**−0.14**	**−0.15**	**0.06**	**−0.06**	**0.14**
NPISent	(2)	**0.03**	**−0.25**	**0.14**	**0.23**	**0.22**	0.01	**0.25**	**0.01**
NewsSent	(3)	**0.04**	**−0.20**	**0.21**	**0.17**	**0.09**	**0.02**	**0.10**	**−0.04**
TweetCount	(4)	**0.05**	**−0.51**	**0.12**	**0.40**	**0.36**	**0.04**	**0.46**	**0.04**
NewsCount	(5)	**−0.02**	**−0.10**	**0.02**	**0.54**	**0.44**	**0.10**	**0.40**	**0.07**
MktSent	(6)	**0.03**	**−0.68**	**0.10**	**0.14**	**0.05**	**0.05**	**0.21**	−0.01
IdioVol	(7)	0.01	**0.37**	**−0.18**	**−0.65**	**−0.36**	**−0.14**	**−0.16**	**−0.08**
MultiGuide	(8)	0.02	0.00	**0.04**	**0.03**	**0.04**	0.00	**0.03**	**−0.09**
InsideTrade	(9)	**0.56**	**−0.08**	**0.14**	**0.05**	**0.05**	**−0.03**	**0.05**	**−0.10**
InsideTradePost	(10)	1.00	**0.04**	**−0.20**	0.21	0.17	**0.09**	**0.02**	**0.10**
VIX	(11)	**−0.02**	1.00	**−0.19**	**−0.13**	**−0.05**	**−0.03**	**−0.15**	**0.03**
PriorRet	(12)	**0.07**	**−0.31**	1.00	**0.13**	**0.03**	**0.04**	−0.01	**−0.05**
MarketCap	(13)	**−0.06**	**−0.06**	**0.03**	1.00	**0.71**	**0.10**	**0.44**	**0.14**
AnalystCov	(14)	**−0.05**	**−0.07**	0.02	**0.45**	1.00	**0.04**	**0.47**	**0.06**
MeetBeat	(15)	−0.02	0.00	**0.03**	**0.06**	**0.03**	1.00	0.00	**0.10**
LitiRisk	(16)	−0.01	**−0.03**	−0.01	**0.05**	**0.25**	**−0.04**	1.00	**0.11**
AnalystDisp	(17)	**−0.05**	**0.07**	**−0.06**	**0.04**	**0.00**	**0.07**	**0.10**	1.00

Note: All variables are defined in Appendix B; bold typeface indicates significance at the 1% level. Pearson's correlation coefficients are shown in the lower triangle (shaded), including the diagonal, and Spearman's rank correlations appear above the diagonal.

Table A5 NPI sentiment and the direction of guidance for standalone guidance

	Dependent variable = *GuidePos*	
Variable	**coef.**	**t-stat.**
NPISent	−0.035	−1.804
NewsSent	0.000	−0.160
TweetCount	0.014	1.913
NewsCount	0.000	0.229
MktSent	0.005	0.552
IdioVol	0.774	5.113***
MultiGuide	−0.146	−1.213
InsideTrade	−0.009	−0.330
InsideTradePost	−0.007	−0.158
VIX	−0.002	−0.314
PriorRet	−0.542	−4.537***
MarketCap	−0.017	−0.610
AnalystCov	−0.054	−11.240***
MeetBeat	0.844	11.744***
LitiRisk	−0.752	−1.033
AnalystDisp	11.769	12.358***
Industry FE	Yes	
Quarter FE	Yes	
Observations	17,736	
LR chi^2	905.319	
Prob > chi^2	0.000	
Pseudo R^2	0.072	

Note: Table A5 presents results of a logistic regression of the incidence of positive standalone guidance on NPI sentiment and relevant controls. *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *IMR* is the inverse Mills ratio obtained from the first-stage model of the Heckman two-stage model; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the net positive news articles, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing a firm (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan Consumer Sentiment Index. *IdioVol* is the average implied volatility (for a 30-day window, at the money option) measured at the release date of guidance issuance; *MultiGuide* is an indicator variable that equals 1 if the firm previously provided earnings guidance for the current quarter's earnings, and 0 otherwise; *InsiderTrade* measures the total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the current quarter; *InsiderTradePost* measures the total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the 15 days after the earnings announcement; *VIX* is the Chicago Board Option's Exchange Implied Volatility Index (a.k.a. the 'fear index') during the three-day window centred on the guidance issuance; *PriorRet* measures the stock return over the 90-day period ending three trading days prior to the guidance issuance; *MarketCap* is the market value of equity (i.e., stock price multiplied by number of stocks outstanding) measured at three trading days prior to the guidance issuance; *AnalystCov* is measured as the number of analysts with outstanding estimates three trading days prior to the guidance issuance; *MeetBeat* is the proportion of the previous four quarters in which firms' reported earnings met or exceeded analysts' prevailing median consensus estimates; *LitiRisk* is the measure of litigation risk that is estimated from the probit model by Kim and Skinner (2012); *AnalystDisp* measures the standard deviation of prevailing analyst estimates for the current period's earnings. Industry-fixed effects based on the Fama–French 48 classification and fiscal-quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

Table A6 The space of NPI sentiment in the stock market

Panel A: NPI reaction to positive guidance								
	Dependent variable = ***CTO[0,+1]***							
	Column I: High institutional ownership		**Column II: Low institutional ownership**		**Column III: High transient institutional ownership**		**Column IV: Low transient institutional ownership**	
Variable	**coef.**	**t-stat.**	**coef.**	**t-stat**	**coef.**	**t-stat.**	**coef.**	**t-stat.**
GuidePos	0.142	1.723*	-0.033	-0.441	0.096	0.766	0.101	0.995
NPISent	-0.042	-1.799*	0.000	-0.005	0.063	0.551	-0.002	-0.043
GuidePos×NPISent	0.016	0.800	0.024	2.873***	0.068	1.968**	-0.009	-0.318
NewsSent	0.001	0.043	0.028	2.210**	0.029	1.052	0.02	1.153
GuidePos×NewsSent	-0.001	-0.077	-0.016	-1.406	-0.013	-0.535	-0.015	-0.977
TweetCount	-0.006	-0.216	-0.031	-0.482	-0.035	-0.172	-0.109	-2.349**
NewsCount	0.008	0.462	-0.043	-2.274**	-0.040	-1.274	-0.000	-0.013
GuidePos×NewsCount	-0.028	-2.047**	-0.009	-0.743	-0.003	-0.099	-0.031	-1.846*
MktSent	0.441	0.923	0.270	1.228	0.606	1.600	0.498	1.089
GuidePos×MktSent	-0.099	-1.207	0.075	1.013	-0.009	-0.068	-0.047	-0.451
Surprise	0.169	6.180***	0.185	9.584***	0.217	6.862***	0.179	6.255***
GuidePos×Surprise	0.003	0.114	0.047	2.135**	0.006	0.167	0.039	1.477
Loss	0.033	2.137**	0.034	1.852	0.062	2.409**	0.03	1.046
PriorRet	-0.015	-1.386	0.012	1.023	0.014	0.809	-0.007	-0.444
MarketCap	-0.049	-2.092**	-0.045	-1.526	-0.103	-2.451**	-0.037	-0.901
AnalystCov	-0.043	-1.371	-0.039	-1.050	-0.087	-1.603	-0.108	-2.103**
MeetBeat	0.073	5.796***	0.082	7.035***	0.108	5.255***	0.082	4.840***
LitRisk	-0.030	-1.112	-0.059	-1.924*	-0.004	-0.087	-0.054	-1.831*
AnalystDisp	-0.036	-2.712***	0.036	2.608**	0.006	0.266	0.013	0.589
RetVol	0.069	2.676***	0.032	1.251	0.034	0.894	0.061	1.461
MtB	-0.017	-1.193	-0.031	-2.544**	-0.038	-1.712*	-0.007	-0.458
PihTra	0.069	2.086**	0.097	2.788***	0.135	1.940*	0.07	2.234**
PihDed	-0.017	-0.784	-0.014	-0.621	0.006	0.150	-0.037	-1.079
PihQix	-0.097	-2.576**	-0.036	-0.992	-0.013	-0.170	-0.064	-1.593
Industry FE	Yes		Yes		Yes		Yes	
Quarter FE	Yes		Yes		Yes		Yes	
Observations	4,927		6,156		4,927		6,156	
F statistics	1.932		2.052		1.591		2.034	
Prob > F	0.000		0.000		0.000		0.000	
Adjusted R^2	0.075		0.099		0.090		0.126	
Chi^2 for difference in coefficient of ***GuidePos×NPISent*** between subsamples				3.639*				5.716**

Notes: Panel A of Table A6 presents standardized coefficients and t-statistics from regressions of NPI reaction on the interaction between positive guidance and NPI sentiment and controls. Column I and II represents the regression results for subsamples of upper and lower quartiles (determined at the industry-quarter level) of percentage of institutional holdings respectively. Column III and Iv represents the regression results for subsamples of upper and lower quartiles (determined at the industry-quarter level) of percentage of transient institutional holdings respectively.*CTO[0,+1]*, NPI reaction proxy, is the overnight stock return (Close-To-Open price) measured as the natural logarithm of the ratio of closing price on guidance date and opening price one day after guidance date; *GuidePos* is an indicator variable for positive guidance that equals one if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and zero otherwise; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the net positive news, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing about a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing about a firm (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan consumer sentiment index. Industry-fixed effects based on the Fama–French 48 classification and fiscal quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

Panel B: NPI sentiment and the direction of guidance

	Dependent variable = *GuidePos*							
	Column I: High institutional ownership		**Column II: Low institutional ownership**		**Column III: High transient institutional ownership**		**Column IV: Low transient institutional ownership**	
Variable	**coef.**	**t-stat.**	**coef.**	**t-stat**	**coef.**	**t-stat.**	**coef.**	**t-stat.**
NPISent	0.017	0.935	0.071	4.621***	0.095	3.233***	0.074	2.850***
NewsSent	0.016	4.678***	0.006	2.099**	0.023	3.015***	0.000	-0.009
TweetCount	0.006	1.030	-0.010	-2.647***	-0.004	-0.570	-0.019	-2.556**
NewsCount	0.007	3.261***	0.004	2.309**	0.013	2.910***	0.003	1.353
MktSent	-0.435	-0.659	-0.035	-0.263	0.092	0.288	0.221	0.742
IdioVol	0.525	1.839	-0.189	-0.879	0.278	0.596	0.033	0.087
MultiGuide	-0.344	-4.893***	-0.205	-3.180***	-0.329	-2.449**	-0.406	-3.766***
GuidePrior	0.146	1.727	-0.111	-1.464	-0.216	-1.374	-0.040	-0.327
InsideTrade	0.086	3.099**	0.055	1.996**	0.067	1.479	0.062	1.126
InsideTradePost	0.038	0.868	0.057	1.266	0.013	0.167	0.115	1.336
VIX	-0.144	-0.765	0.044	0.840	0.073	0.636	0.149	1.254
Surprise	82.169	9.296***	26.924	5.409***	43.349	4.154***	73.140	5.464***
Loss	-0.611	-3.359***	-0.517	-3.801***	-1.015	-4.148***	-0.208	-0.781
PriorRet	0.006	0.056	-0.032	-0.314	0.007	0.040	-0.259	-1.330
MarketCap	-0.007	-0.862	0.007	1.663*	0.051	1.912*	0.007	1.096
AnalystCov	-0.010	-0.900	-0.023	-2.216**	-0.069	-3.189***	-0.033	-1.780
MeetBeat	0.836	7.795***	0.970	9.670***	1.565	7.255***	0.595	3.582***
LitiRisk	-2.067	-0.396	2.943	0.651	10.889	1.323	-18.637	-1.896*
AnalystDisp	-0.011	-0.010	-1.287	-1.249	-0.377	-0.213	0.306	0.156
PihTra	2.250	2.784***	2.462	3.981***	2.723	1.636	3.267	2.218**
PihDed	-1.014	-1.067	1.460	1.846*	2.422	1.194	0.759	0.614
PihQix	-0.464	-0.622	0.972	1.802*	2.109	1.470	0.886	1.071
Firm FE	Yes		Yes		Yes		Yes	
Year-Quarter FE	Yes		Yes		Yes		Yes	
Observations	4,927		6,156		4,927		6,156	
LR chi^2	650.822		958.104		850.680		910.077	
Prob > chi^2	0.000		0.000		0.000		0.000	
Pseudo R^2	0.154		0.135		0.185		0.151	
Chi^2 for difference in coefficient of ***NPISent*** between subsamples				5.570**				4.079**

Notes: Panel B of Table A6 presents results of a logistic regressions of the incidence of positive guidance on NPI sentiment and relevant controls. Column I and II represents the regression results for subsamples of upper and lower quartiles (determined at the industry-quarter level) of percentage of institutional holdings respectively. Column III and Iv represents the regression results for subsamples of upper and lower quartiles (determined at the industry-quarter level) of percentage of transient institutional holdings respectively. *GuidePos* is an indicator variable for positive guidance that equals one if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and zero otherwise; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the net positive news, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing about a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing about a firm (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan consumer sentiment index. Other variables are defined in Appendix B Industry-fixed effects based on the Fama–French 48 classification and fiscal quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

Table A7 Macroeconomic factors

Panel A: Descriptive statistics – Partitioned based on *GuidePos*

	***GuidePos* = 1 (n = 10,925)**			***GuidePos* = 0 (n = 14,675)**			**Difference**	
	Mean	Median	Std.dev.	Mean	Median	St.Dev	Mean	Median
MktRet	3.799	4.555	7.413	3.450	4.030	7.281	***	***
MktRetLead	3.871	4.555	7.070	4.096	4.555	6.877	**	
MktRetLag	3.240	4.555	8.680	2.604	4.030	8.657	***	***
Def	1.117	1.000	0.432	1.120	1.000	0.465		***
DefLead	1.053	0.993	0.323	1.058	0.993	0.348		
DefLag	1.187	1.057	0.534	1.175	1.033	0.551		***
Yld	0.111	0.087	0.097	0.106	0.087	0.095	***	***
YldLead	0.123	0.087	0.128	0.118	0.087	0.125	***	***
YldLag	0.106	0.087	0.083	0.101	0.067	0.084	***	***
GDP	0.476	0.477	0.477	0.467	0.471	0.499		
GDPLead	0.516	0.477	0.401	0.513	0.471	0.407		
GDPLag	0.412	0.477	0.636	0.396	0.477	0.674	*	
Cons	0.884	0.905	0.466	0.866	0.875	0.497	***	
ConsLead	0.929	0.956	0.381	0.920	0.956	0.392	*	
ConsLag	0.790	0.875	0.706	0.765	0.873	0.751	**	*
Labor	0.308	0.485	1.126	0.335	0.508	1.139	**	***
LaborLead	0.401	0.510	1.051	0.387	0.510	1.125		
LaborLag	0.342	0.508	1.076	0.340	0.510	1.173		***
URate	7.476	7.800	1.771	7.296	7.533	1.689	***	***
URateLead	7.356	7.733	1.822	7.174	7.233	1.748	***	***
URateLag	7.555	7.800	1.712	7.380	7.533	1.627	***	***
CPIQ	0.444	0.339	0.751	0.423	0.314	0.764	**	
CPIQLead	0.438	0.339	0.746	0.408	0.314	0.760	***	**
CpidLag	0.301	0.314	1.018	0.306	0.339	1.073		**
CAY	-0.014	-0.013	0.013	-0.014	-0.017	0.013	***	***
CAYLead	-0.015	-0.013	0.013	-0.016	-0.017	0.013	***	***
CAYLag	-0.013	-0.012	0.014	-0.013	-0.012	0.014		

Panel B:	**Dependent Variable = *NPISent***	
Variable	**coef.**	**t-stat**
MktRet	0.061	2.090**
MktRetLead	0.075	3.006***
MktRetLag	0.046	3.183***
Def	−0.115	−0.127
DefLead	0.229	0.314
DefLag	−0.455	−0.773
Yld	−5.793	−2.459**
YldLead	−0.851	−0.481
YldLag	10.107	3.220***
GDP	−0.188	−0.742
GDPLead	−0.745	−2.852***
GDPLag	−0.001	−0.004
Cons	−1.367	−2.664***
ConsLead	−0.435	−0.875
ConsLag	−0.729	−3.841***
Labor	0.416	2.076**
LaborLead	0.353	2.138**
LaborLag	0.240	4.499***
URate	−1.761	−3.406***
URateLead	−0.507	−0.753
URateLag	1.659	2.491**
CPIQ	−0.251	−1.001
CPIQLead	−0.641	−2.522**
CPIQLag	0.374	1.750
CAY	−1.934	−0.087
CAYLead	78.852	2.818***
CAYLag	−80.884	−2.890***
Observations	64,190	
Quarter FE	Yes	
F statistic	128.793	
Prob > F	0.000	
Adjust R2	0.029	

Note: Table A6 Panel A describes the macroeconomic factors for the sample of 25,600 firm–quarter observations used in the main analysis. Table A6 Panel B presents the results of the first-stage regression of *NPISent* on a broad set of macroeconomic variables, as well as their lagged and lead measures. *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *MktRet* is the market returns measured as the CRSP Value-weighted Index return including distributions; *Def* is the difference between the yields to maturity on BBB- and AAA-rated bond yields; *Yld* is the yield on the three-month Treasury bill; *GDP* is measured as 100 times the quarterly change in the natural logarithm of chained (1996) GDP; *Cons* is the consumption growth, measured as 100 times the quarterly change in the natural logarithm of personal consumption expenditures; *Labor*, the labor income growth, is measured as 100 times the quarterly change in the natural logarithm of labor income computed as total personal income minus dividend income per capita, and deflated by the PCE Deflator; *Urate* is the average monthly unemployment rate, as reported by US Bureau of Labor Statistics; *CPIQ* is the Consumer Price Index inflation rate; *CAY* is the consumption-to-wealth ratio obtained from the website of Sydney Ludvigson, Professor of Economics at NYU (https://www.sydneyludvigson.com/data-and-appendixes/). Fiscal-quarter-fixed effects are included; *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two- tailed tests.

Panel C	Dependent variable = *GuidePos*	
Variable	**coef.**	**t-stat.**
NPISentResid1	0.026	2.627**
NewsSent	0.004	1.941
TweetCount	−0.004	−1.752
NewsCount	−0.001	−1.084
MktSent	−0.109	−1.337
IdioVol	0.128	1.067
MultiGuide	−0.399	−10.078***
GuidePrior	0.781	15.324***
InsideTrade	0.066	4.306***
InsideTradePost	0.015	0.583
VIX	0.007	0.201
Surprise	29.983	6.868***
Loss	−0.472	−6.097***
PriorRet	0.081	1.406
MarketCap	0.002	1.464
AnalystCov	−0.023	−6.149***
MeetBeat	0.936	13.805***
LitiRisk	5.538	2.764**
AnalystDisp	1.038	1.806
Industry FE	Yes	
Quarter FE	Yes	
Observations	25,600	
LR chi^2	1472.490	
Prob > chi^2	0.000	
Pseudo R^2	0.055	

Note: Table A6 Panel C presents results of a logistic regression of the incidence of positive guidance on residual NPI sentiment (*NPISentResid1*) from the first-stage model and relevant controls. *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *IMR* is the inverse Mills ratio obtained from the first-stage model of the Heckman two-stage model; *NewsSent* is the net positive news, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing a firm (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan Consumer Sentiment Index. *IdioVol* is the average implied volatility (for a 30-day window, at the money option) measured at the release date of guidance issuance; *MultiGuide* is an indicator variable that equals 1 if the firm previously provided earnings guidance for the current quarter's earnings, and 0 otherwise; *InsiderTrade* measures the total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the current quarter; *InsiderTradePost* measures the total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the 15 days after the earnings announcement; *VIX* is the Chicago Board Option's Exchange Implied Volatility Index (a.k.a. the 'fear index') during the three-day window centred on the guidance issuance; *PriorRet* measures the stock return over the 90-day period ending three trading days prior to the guidance issuance; *MarketCap* is the market value of equity (i.e., stock price multiplied by number of stocks outstanding) measured at three trading days prior to the guidance issuance; *AnalystCov* is measured as the number of analysts with outstanding estimates three trading days prior to the guidance issuance; *MeetBeat* is the proportion of the previous four quarters in which firms' reported earnings met or exceeded analysts' prevailing median consensus estimates; *LitiRisk* is the measure of litigation risk estimated from the probit model by Kim and Skinner (2012); *AnalystDisp* measures the standard deviation of prevailing analysts' estimates for the current period's earnings. Industry-fixed effects based on the Fama–French 48 classification and fiscal quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

Table A8 Distinguishing investor attention and investor sentiment

	Column I: First stage		**Column II: Second stage**	
	Dependent variable = *NPISent*		**Dependent variable = *GuidePos***	
Variable	**coef.**	**t-stat.**	**coef.**	**t-stat.**
NPISentResid2			0.026	2.671***
NewsSent	0.036	10.001***	0.006	2.984***
TweetCount	0.177	85.876***	0.000	−0.159
NewsCount	−0.005	−3.407***	−0.001	−1.350
MktSent	0.017	7.258***	0.021	2.683***
IdioVol			−0.003	−0.019
MultiGuide			−0.270	−6.620***
GuidePrior			0.837	15.739***
InsideTrade			0.074	4.444***
InsideTradePost			0.019	0.719
VIX			0.028	2.745***
Surprise			32.985	7.067***
Loss			−0.496	−5.870***
PriorRet			0.083	1.331
MarketCap			0.002	1.565
AnalystCov			−0.025	−6.207***
MeetBeat			1.006	14.100***
LitiRisk			5.549	2.769***
AnalystDisp			0.904	1.540
Industry FE	Yes		Yes	
Quarter FE	Yes		Yes	
Observations	64,190		25,600	
LR chi^2 or F statistics	238.735		1019.766	
Prob > chi^2 or Prob>F	0.000		0.000	
Pseudo R^2 or Adjusted R^2	0.802		0.055	

Note: Table A7 Panel A presents results of the first-stage regression of *NPISent* on *TweetCount*, *NewsCount*, *NewsSent*, and *MktSent*; Table A7 Panel B presents results of a logistic regression of the incidence of positive guidance on residual NPI sentiment (*NPISentResid2*) from the first-stage model and relevant controls; *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *IMR* is the inverse Mills ratio obtained from the first-stage model of the Heckman two-stage model; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the net positive news, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing a firm (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan Consumer Sentiment Index. *IdioVol* is the average implied volatility (for a 30-day window, at the money option) measured at the release date of guidance issuance; *MultiGuide* is an indicator variable that equals 1 if the firm previously provided earnings guidance for the current quarter's earnings, and 0 otherwise; *InsiderTrade* measures the total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the current quarter; *InsiderTradePost* measures the total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the 15 days after the earnings announcement; *VIX* is the Chicago Board Option's Exchange Implied Volatility Index (a.k.a. the 'fear index') during the three-day window centred on the guidance issuance; *PriorRet* measures the stock return over the 90-day period ending three trading days prior to the guidance issuance; *MarketCap* is the market value of equity (i.e., stock price multiplied by number of stocks outstanding) measured at three trading days prior to the guidance issuance; *AnalystCov* is measured as the number of analysts with outstanding estimates three trading days prior to the guidance issuance; *MeetBeat* is the proportion of the previous four quarters in which firms' reported earnings met or exceeded analysts' prevailing median consensus estimates; *LitiRisk* is a measure of litigation risk estimated from the probit model by Kim and Skinner (2012); *AnalystDisp* measures the standard deviation of prevailing analysts' estimates for the current period's earnings. Industry-fixed effects based on the Fama–French 48 classification and fiscal-quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

Table A9 Alternative measures of NPI sentiment

	Dependent Variable = *GuidePos*							
	Panel A: Harvard dictionary		Panel B: LM dictionary		Panel C: L1 lexicon		Panel D: L2 lexicon	
Variable	coef.	t-stat	coef.	t-stat	coef.	t-stat	coef.	t-stat
NPISent	0.003	3.265***	0.008	2.424**	0.003	2.622***	0.003	3.378***
NewsSent	0.005	2.226**	0.005	2.159**	0.005	2.174**	0.005	2.156**
TweetCount	−0.001	−0.953	0.000	−0.004	0.000	0.262	−0.001	−0.973
NewsCount	−0.001	−1.465	−0.001	−0.888	−0.001	−1.480	−0.001	−1.473
MktSent	0.019	2.461**	0.019	2.450**	0.019	2.509**	0.019	2.451**
IdioVol	−0.049	−0.376	−0.003	−0.020	−0.034	−0.258	−0.051	−0.386
MultiGuide	−0.272	−6.674***	−0.271	−6.661***	−0.271	−6.648***	−0.272	−6.674***
GuidePrior	0.840	15.798***	0.837	15.746***	0.838	15.776***	0.840	15.797***
InsideTrade	0.073	4.401***	0.073	4.405***	0.073	4.387***	0.073	4.392***
InsideTradePost	0.019	0.715	0.019	0.720	0.019	0.706	0.019	0.715
VIX	0.027	2.657***	0.026	2.556**	0.028	2.745***	0.027	2.666***
Surprise	32.899	7.057***	32.961	7.060***	32.901	7.058***	32.906	7.057***
Loss	−0.501	−5.966***	−0.495	−5.866***	−0.498	−5.917***	−0.502	−5.973***
PriorRet	0.086	1.386	0.084	1.354	0.084	1.351	0.086	1.387
MarketCap	0.001	0.769	0.002	1.334	0.001	0.886	0.001	0.723
AnalystCov	−0.026	−6.365***	−0.025	−6.178***	−0.026	−6.401***	−0.026	−6.374***
MeetBeat	1.005	14.062***	1.002	14.047***	1.003	14.042***	1.004	14.046***
LitiRisk	4.464	2.252**	5.588	2.801***	4.927	2.485**	4.436	2.240**
AnalystDisp	0.842	1.431	0.896	1.522	0.853	1.450	−0.101	−3.278***
Industry FE	Yes		Yes		Yes		Yes	
Quarter FE	Yes		Yes		Yes		Yes	
Observations	25,600		25,600		25,600		25,600	
LR chi2	1032.744		1024.431		1035.381		1037.279	
Prob > chi2	0.000		0.000		0.000		0.000	
Pseudo R2	0.055		0.055		0.055		0.055	

Note: Table A8 presents results of a logistic regression of the incidence of positive guidance on alternative measures of NPI sentiment and relevant controls. *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *IMR* is the inverse Mills ratio obtained from the first-stage model of the Heckman two-stage model; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the net positive news, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing a firm (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan Consumer Sentiment Index. *IdioVol* is the average implied volatility (for a 30-day window, at the money option) measured at the release date of guidance issuance; *MultiGuide* is an indicator variable that equals 1 if the firm previously provided earnings guidance for the current quarter's earnings, and 0 otherwise; *InsiderTrade* measures the total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the current quarter; *InsiderTradePost* measures the total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the 15 days after the earnings announcement; *VIX* is the Chicago Board Option's Exchange Implied Volatility Index (a.k.a. the 'fear index') during the three-day window centred on the guidance issuance; *PriorRet* measures the stock return over the 90-day period ending three trading days prior to the guidance issuance; *MarketCap* is the market value of equity (i.e., stock price multiplied by number of stocks outstanding) measured at three trading days prior to the guidance issuance; *AnalystCov* is measured as the number of analysts with outstanding estimates three trading days prior to the guidance issuance; *MeetBeat* is the proportion of the previous four quarters in which firms' reported earnings met or exceeded analysts' prevailing median consensus estimates; *LitiRisk* is the measure of litigation risk estimated from the probit model by Kim and Skinner (2012); *AnalystDisp* measures the standard deviation of prevailing analysts' estimates for the current period's earnings. Industry-fixed effects based on the Fama–French 48 classification and fiscal-quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

Table A10 Alternative time periods for NPI sentiment

Variable	Dependent Variable = *GuidePos*			
	Panel A: [−91,−61] period		Panel B: [−61,−31] period	
	coef.	t-stat.	coef.	t-stat.
NPISent	0.025	2.489**	0.029	3.244***
Observations	25,600		25,600	
LR chi2	1506.530		1505.467	
Prob > chi2	0.000		0.000	
Pseudo R2	0.054		0.054	
	Panel C: [−16,−1] period		Panel D: [−8,−1] period	
	coef.	t-stat.	coef.	t-stat.
NPISent	0.027	1.889*	0.034	1.387
Observations	25,600		25,600	
LR chi2	1498.118		1119.213	
Prob > chi2	0.000		0.000	
Pseudo R2	0.054		0.054	

Other controls included: *Industry effects, time effects, NewsSent, MktSent, TweetCount, NewsCount, IdioVol, MultiGuide, GuidePrior, InsideTrade, InsideTradePost, VIX, Surprise, Loss, PriorRet, MarketCap, AnalystCov, MeetBeat, LitRisk, AnalystDisp.*

Note: Table A9 presents results of a logistic regression of the incidence of positive guidance on NPI sentiment measured using alternative time periods, and relevant controls. *GuidePos* is an indicator variable for positive guidance that equals 1 if the guidance estimate is greater than the pre-guidance prevailing median analyst forecast estimate, and 0 otherwise; *IMR* is the inverse Mills ratio obtained from the first-stage model of the Heckman two-stage model; *NPISent* is the net positive tweets, measured as the number of positive tweets minus the number of negative tweets (in a 30-day window prior to the guidance issuance); *NewsSent* is the net positive news, measured as the number of positive news articles minus the number of negative news articles (in a 30-day window prior to the guidance issuance); *TweetCount* is the total number of tweets discussing a firm (in a 30-day window prior to the guidance issuance). *NewsCount* is the total number of news articles discussing a firm (in a 30-day window prior to the guidance issuance); *MktSent* is the Michigan Consumer Sentiment Index. *IdioVol* is the average implied volatility (for a 30-day window, at the money option) measured at the release date of guidance issuance; *MultiGuide* is an indicator variable that equals 1 if the firm previously provided earnings guidance for the current quarter's earnings, and 0 otherwise; *InsiderTrade* measures the total insider trades (i.e., sales + purchases) of the senior management (scaled by shares outstanding at the beginning of the quarter) during the current quarter; *InsiderTradePost* measures the total insider trades (i.e., sales + purchases) by senior management (scaled by shares outstanding at the beginning of the quarter) during the 15 days after the earnings announcement; *VIX* is the Chicago Board Option's Exchange Implied Volatility Index (a.k.a. the 'fear index') during the three-day window centred on the guidance issuance; *PriorRet* measures the stock return over the 90-day period ending three trading days prior to the guidance issuance; *MarketCap* is the market value of equity (i.e., stock price multiplied by number of stocks outstanding) measured at three trading days prior to the guidance issuance; *AnalystCov* is measured as the number of analysts with outstanding estimates three trading days prior to the guidance issuance; *MeetBeat* is the proportion of the previous four quarters in which firms' reported earnings met or exceeded analysts' prevailing median consensus estimates; *LitiRisk* is a measure of litigation risk estimated from the probit model by Kim and Skinner (2012); *AnalystDisp* measures the standard deviation of prevailing analysts' estimates for the current period's earnings. Industry-fixed effects based on the Fama–French 48 classification and fiscal-quarter-fixed effects are included. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.

Appendix E Additional Analysis and Robustness Check—Investor Earnings Expectations and Analyst Forecasts

Table A11 Sample selection at individual analyst level

Sample Selection	No. of observations	No. of firm-quarters	No. of distinct firms
Estimize Daily Consensus (1 Jan 2012 to 31 Dec 2018)	3,896,909	40,466	2729
Less: consensus made outside [−90, 0] period prior to earnings announcement	(2,311,463)		(0)
Number of observations remaining	1,585,446	38,240	2729
[−90,−61]	414,321	15,120	1900
[−60, −31]	478,631	18,139	2102
[−30, −1]	655,533	34,406	2669
[0]	36,961	36,961	2729
Keep one obs. per firm quarter	(1,548,485)	(0)	(0)
Number of observations remaining	38,240	38,240	2729
I/B/E/S (FPI=6) Individual Forecasts (1 Jan 2012 to 31 Dec 2018)	1,210,847	118,602	7531
Less: excluded forecasts	(62,974)	(1765)	(115)
Less: forecast made outside [−90, 0] period prior to earnings announcement	(123,328)	(4227)	(113)
Number of observations remaining	1,024,545	112,610	7303
Keep the most recent forecast by analyst per window	(73,477)	(0)	(0)
Number of observations remaining	951,068	112,610	7303
[−90,−61]	269,545	79304	6679
[−60, −31]	210,079	74,849	6730
[−30, −1]	450,868	90,859	6802
[0]	20,576	15,245	3981
Estimize and I/B/E/S merged	358,001	26,419	2065
Less: Estimize and I/B/E/S report actual EPS does not match to two decimal places	(51,338)	(4258)	(113)
Number of observations remaining	306,663	22,161	1952
Less: observations with missing control variables			
Compustat – fundamentals	(31,467)	(2275)	(208)
CRSP – daily stock	(171)	(33)	(4)
Number of observations remaining	275,016	19,853	1740
[−90,−61]	132103	7613	1175

[−60, −31]	154,003	9213	1317
[−30, −1]	258,973	18,232	1709
[0]	15,434	1526	865
Total useable observations	10,361	3226	686

Note: Estimize and I/B/E/S data are merged using Cusip, Year-quarter, and Actual announcement date.

Table A12 Descriptive statistics at individual analyst level

Variable	N	Mean	SD	Min	P25	P50	P75	Max
$\boldsymbol{Revision_{i,j,t}}$	10361	−0.0003	0.0027	−0.0121	−0.0009	−0.0001	0.0005	0.0092
$\boldsymbol{Error_{i,j,t}}$	10361	−0.0009	0.0132	−0.1154	−0.0020	−0.0006	0.0000	0.6970
$\boldsymbol{Pessimism_{i,j,t}}$	10361	0.7275	0.4452	0.0000	0.0000	1.0000	1.0000	1.0000
$\boldsymbol{PessimismMag_{i,j,t}}$	10361	0.0017	0.0030	0.0000	0.0000	0.0006	0.0020	0.0182
$\boldsymbol{RevDown_{i,j,t}}$	10361	0.5516	0.4974	0.0000	0.0000	1.0000	1.0000	1.0000
$\boldsymbol{DownFollow_{i,j,t}}$	10361	0.3156	0.4648	0.0000	0.0000	0.0000	1.0000	1.0000
$\boldsymbol{DownAganist_{i,j,t}}$	10361	0.2360	0.4246	0.0000	0.0000	0.0000	0.0000	1.0000
$\boldsymbol{DownFollowMag_{i,j,t}}$	10361	0.0540	0.1453	0.0000	0.0000	0.0000	0.0270	0.8955
$\boldsymbol{DownAgainstMag_{i,j,t}}$	10361	0.0345	0.1078	0.0000	0.0000	0.0000	0.0000	0.7333
$\boldsymbol{EstmzRev_{j,t}}$	10361	−0.0002	0.0017	−0.0076	−0.0005	0.0000	0.0003	0.0062
$\boldsymbol{Follow_{j,t}}$	10361	12.6535	7.1099	2.0000	7.0000	11.0000	17.0000	31.0000
$\boldsymbol{Dispersion_{j,t}}$	10361	0.0759	0.0768	0.0050	0.0250	0.0502	0.0993	0.4445
$\boldsymbol{AccScore_{i,j,t}}$	10361	19.4928	54.2971	−121.4286	−20.0000	25.0000	66.6667	100.0000
$\boldsymbol{BoldScore_{i,j,t}}$	10361	25.5411	49.4534	−100.0000	−12.5000	27.7778	66.6667	100.0000
$\boldsymbol{BiasScore_{i,j,t}}$	10361	18.4967	54.1490	−116.6667	−21.4286	20.8333	64.2857	100.0000
$\boldsymbol{FirmExp_{i,j,t}}$	10361	17.8915	16.5879	0.0000	6.0000	13.0000	25.0000	76.0000
$\boldsymbol{IndusExp_{i,j,t}}$	10361	30.4650	23.4795	0.0000	12.0000	25.0000	45.0000	101.0000
$\boldsymbol{GenExp_{i,j,t}}$	10361	39.0132	28.6772	1.0000	17.0000	32.0000	58.0000	123.0000
$\boldsymbol{BrokerSize_{i,t}}$	10361	67.5591	37.7331	4.0000	35.0000	69.0000	103.0000	130.0000
$\boldsymbol{FirmSize_{j,t}}$	10361	9.7053	1.4744	6.1912	8.5994	9.7298	10.7986	12.8331
$\boldsymbol{BtM_{j,t}}$	10361	0.4253	0.4428	−0.2598	0.1802	0.3301	0.5294	2.7232
$\boldsymbol{Loss_{j,t}}$	10361	0.2749	0.4465	0.0000	0.0000	0.0000	1.0000	1.0000
$\boldsymbol{Leverage_{j,t}}$	10361	0.3009	0.1782	0.0000	0.1918	0.2833	0.3842	0.8976
$\boldsymbol{Turnover_{j,t}}$	10361	13.1618	10.2305	2.8047	6.5026	9.9099	15.8580	57.4500
$\boldsymbol{RetVol_{j,t}}$	10361	0.0205	0.0089	0.0084	0.0141	0.0184	0.0247	0.0531
$\boldsymbol{Guidance_{j,t}}$	10361	0.1132	0.3169	0.0000	0.0000	0.0000	0.0000	1.0000

Note: for presentation, $DownFollowMag_{i,j,t}$ and $DownAgainstMag_{i,j,t}$ are multiplied by 100.

Table A13 Correlation matrix at individual analyst level

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1)	***Revision***$_{i,j,t}$	1.00	0.00	0.01	0.00	**−0.86**	**−0.54**	**−0.42**	**−0.60**	**−0.45**	**0.14**	−0.01	0.00	**−0.09**
(2)	***Error***$_{i,j,t}$	0.02	1.00	**−0.77**	**−0.99**	**0.03**	**0.05**	−0.02	**0.03**	**−0.03**	**−0.03**	**−0.07**	**−0.12**	**0.05**
(3)	***Pessimism***$_{i,j,t}$	0.01	−0.21	1.00	**0.78**	−0.01	**−0.03**	0.02	**−0.04**	0.02	**0.03**	**0.04**	**−0.07**	0.02
(4)	***PessimismMag***$_{i,j,t}$	**−0.06**	−0.33	0.36	1.00	−0.03	**−0.05**	0.02	**−0.03**	**0.04**	**0.03**	**0.07**	**0.15**	−0.06
(5)	***RevDown***$_{i,j,t}$	**−0.57**	0.00	−0.01	−0.02	1.00	**0.61**	**0.50**	**0.60**	**0.50**	**−0.12**	0.00	−0.02	0.12
(6)	***DownFollow***$_{i,j,t}$	**−0.37**	0.02	**−0.03**	**−0.04**	**0.61**	1.00	**−0.38**	**0.98**	**−0.37**	**−0.56**	0.02	−0.02	**0.06**
(7)	***DownAganist***$_{i,j,t}$	**−0.26**	−0.02	0.02	0.02	**0.50**	**−0.38**	1.00	**−0.37**	**0.99**	**0.48**	−0.02	0.00	**0.08**
(8)	***DownFollowMag***$_{i,j,t}$	**−0.62**	**0.03**	**−0.04**	**0.12**	**0.33**	**0.55**	**−0.21**	1.00	**−0.36**	**−0.57**	**0.03**	0.01	**0.04**
(9)	***DownAgainstMag***$_{i,j,t}$	**−0.48**	**−0.04**	−0.01	**0.14**	**0.29**	**−0.22**	**0.58**	**−0.12**	1.00	**0.48**	−0.01	0.02	**0.07**
(10)	***EstmzRev***$_{j,t}$	**0.13**	**−0.10**	0.02	0.01	**−0.08**	**−0.35**	**0.29**	**−0.36**	**0.25**	1.00	0.01	−0.02	0.00
(11)	***Follow***$_{j,t}$	0.01	**−0.03**	**0.03**	**0.05**	0.00	0.01	−0.02	0.01	0.01	**0.04**	1.00	**0.16**	**−0.08**
(12)	***Dispersion***$_{j,t}$	**−0.03**	**0.09**	**−0.06**	**0.15**	−0.01	−0.02	0.01	**0.10**	**0.08**	**−0.03**	**0.06**	1.00	**−0.04**
(13)	***AccScore***$_{i,j,t}$	**−0.03**	0.00	0.02	**−0.08**	**0.12**	**0.06**	**0.07**	**−0.04**	**−0.03**	−0.01	**−0.07**	−0.01	1.00
(14)	***BoldScore***$_{i,j,t}$	**0.17**	−0.02	**0.10**	**0.06**	**−0.25**	**−0.13**	**−0.15**	**−0.08**	**−0.12**	0.00	**−0.05**	0.00	**−0.12**
(15)	***BiasScore***$_{i,j,t}$	**0.32**	**−0.06**	**0.15**	**0.13**	**−0.37**	**−0.22**	**−0.19**	**−0.19**	**−0.19**	0.01	**−0.07**	−0.02	**−0.15**
(16)	***FirmExp***$_{i,j,t}$	0.00	−0.02	**0.04**	**−0.04**	0.01	0.02	0.00	−0.01	−0.02	−0.01	**0.04**	**−0.07**	−0.01
(17)	***IndusExp***$_{i,j,t}$	0.01	0.00	**0.03**	−0.01	−0.02	0.00	−0.01	0.00	−0.01	0.00	**0.05**	−0.02	−0.02
(18)	***GenExp***$_{i,j,t}$	−0.01	0.00	**0.03**	0.00	0.01	0.02	0.00	0.02	0.00	−0.02	0.02	**−0.04**	−0.01
(19)	***BrokerSize***$_{i,t}$	0.00	−0.02	0.02	**−0.04**	0.02	0.00	0.01	**−0.04**	−0.01	0.01	**−0.10**	**−0.03**	0.03
(20)	***FirmSize***$_{j,t}$	0.01	**−0.03**	**0.10**	**−0.20**	**0.03**	0.02	0.01	**−0.12**	**−0.11**	**0.05**	**0.17**	**0.04**	0.00
(21)	***BtM***$_{j,t}$	**−0.03**	**−0.10**	0.00	**0.36**	0.00	−0.02	**0.03**	**0.12**	**0.16**	0.00	**0.14**	**0.07**	−0.01
(22)	***Loss***$_{j,t}$	**0.04**	0.01	**−0.04**	**0.17**	**−0.06**	**−0.03**	**−0.04**	**0.06**	**0.04**	0.02	**0.22**	**0.12**	0.00
(23)	***Leverage***$_{j,t}$	0.01	**0.06**	**−0.04**	**0.09**	−0.01	0.01	−0.02	**0.04**	**0.03**	**−0.03**	**−0.05**	0.02	0.02
(24)	***Turnover***$_{j,t}$	**−0.05**	**0.03**	**−0.07**	**0.29**	−0.01	0.02	**−0.04**	**0.18**	**0.12**	**−0.09**	**0.04**	**0.15**	0.01
(25)	***RetVol***$_{j,t}$	**−0.04**	**0.04**	**−0.09**	**0.31**	**−0.03**	0.00	**−0.03**	**0.19**	**0.14**	**−0.08**	**0.08**	**0.14**	0.01
(26)	***Guidance***$_{j,t}$	0.00	0.01	**0.05**	**−0.11**	−0.02	0.01	**−0.03**	**−0.03**	**−0.06**	0.01	**−0.14**	**−0.20**	**0.03**

		(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)
(1)	***Revision***$_{i,j,t}$	**0.26**	**0.42**	−0.02	0.02	−0.01	0.00	0.00	**−0.05**	**0.05**	**0.03**	−0.02	0.00	0.01
(2)	***Error***$_{i,j,t}$	**−0.11**	**−0.19**	−0.01	−0.01	−0.01	0.01	0.02	**−0.17**	**−0.07**	−0.02	**−0.08**	**−0.08**	**0.07**
(3)	***Pessimism***$_{i,j,t}$	**0.10**	**0.16**	**0.04**	**0.03**	**0.03**	0.02	**0.09**	0.01	**−0.04**	**−0.03**	**−0.07**	**−0.09**	**0.05**
(4)	***PessimismMag***$_{i,j,t}$	**0.11**	**0.18**	0.00	0.01	0.01	−0.01	**−0.05**	**0.18**	**0.09**	0.02	**0.11**	**0.11**	**−0.09**
(5)	***RevDown***$_{i,j,t}$	**−0.25**	**−0.38**	0.02	−0.02	0.01	0.02	**0.03**	0.01	**−0.06**	−0.02	−0.02	**−0.04**	−0.02
(6)	***DownFollow***$_{i,j,t}$	**−0.13**	**−0.23**	0.01	−0.01	0.01	0.00	0.02	−0.02	**−0.03**	0.00	0.01	0.00	0.01
(7)	***DownAganist***$_{i,j,t}$	**−0.14**	**−0.19**	0.01	−0.01	0.00	0.01	0.01	**0.03**	**−0.04**	−0.02	**−0.03**	**−0.04**	**−0.03**
(8)	***DownFollowMag***$_{i,j,t}$	**−0.14**	**−0.25**	0.01	−0.01	0.01	−0.01	−0.01	0.02	0.00	0.00	**0.06**	**0.04**	0.00
(9)	***DownAgainstMag***$_{i,j,t}$	**−0.15**	**−0.21**	0.00	−0.01	0.00	0.01	−0.01	**0.06**	**−0.03**	−0.02	−0.01	−0.02	**−0.04**
(10)	***EstmzRev***$_{j,t}$	0.01	0.02	−0.01	0.01	−0.01	0.01	0.02	0.02	0.02	**−0.04**	**−0.06**	**−0.04**	−0.01
(11)	***Follow***$_{j,t}$	**−0.07**	**−0.08**	**0.05**	**0.07**	**0.03**	**−0.09**	**0.19**	**0.18**	**0.18**	**−0.04**	**0.07**	**0.09**	**−0.14**
(12)	***Dispersion***$_{j,t}$	−0.01	**−0.03**	**−0.05**	0.00	−0.02	**−0.07**	0.00	**0.14**	**0.18**	**0.05**	**0.22**	**0.23**	**−0.29**
(13)	***AccScore***$_{i,j,t}$	**−0.15**	**−0.19**	−0.01	−0.01	−0.01	**0.03**	−0.01	**−0.04**	0.00	0.01	0.00	0.00	**0.03**
(14)	***BoldScore***$_{i,j,t}$	1.00	**0.71**	−0.02	−0.01	−0.01	0.01	−0.02	−0.01	0.01	0.00	0.01	0.01	0.01
(15)	***BiasScore***$_{i,j,t}$	0.72	1.00	−0.02	−0.01	−0.01	**0.03**	−0.02	−0.02	0.00	0.00	0.00	0.00	**0.03**
(16)	***FirmExp***$_{i,j,t}$	−0.01	−0.01	1.00	**0.65**	**0.55**	0.01	**0.17**	0.02	**−0.08**	0.01	**−0.15**	**−0.18**	0.00
(17)	***IndusExp***$_{i,j,t}$	−0.01	−0.02	**0.63**	1.00	**0.82**	**−0.05**	**0.07**	**0.03**	0.01	−0.01	**−0.05**	**−0.03**	−0.02
(18)	***GenExp***$_{i,j,t}$	−0.01	−0.01	**0.52**	**0.79**	1.00	0.01	**0.04**	**0.03**	**−0.04**	**−0.03**	**−0.03**	**−0.05**	−0.01
(19)	***BrokerSize***$_{i,t}$	0.01	**0.03**	0.01	**−0.04**	−0.01	1.00	**0.14**	**−0.04**	**−0.13**	**−0.03**	**−0.14**	**−0.15**	**0.04**
(20)	***FirmSize***$_{j,t}$	−0.02	−0.01	**0.18**	**0.07**	**0.03**	**0.15**	1.00	**−0.14**	**−0.29**	**−0.16**	**−0.70**	**−0.63**	−0.02
(21)	***BtM***$_{j,t}$	0.00	0.00	**−0.03**	0.02	**0.04**	**−0.04**	**−0.22**	1.00	**0.15**	**−0.19**	**0.17**	**0.16**	**−0.14**
(22)	***Loss***$_{j,t}$	0.01	0.01	**−0.08**	−0.01	**−0.05**	**−0.13**	**−0.30**	**0.14**	1.00	**0.16**	**0.39**	**0.50**	**−0.13**
(23)	***Leverage***$_{j,t}$	0.00	0.00	−0.02	−0.02	−0.02	**−0.03**	**−0.19**	**−0.14**	**0.13**	1.00	**0.16**	**0.17**	**−0.03**
(24)	***Turnover***$_{j,t}$	0.01	0.00	**−0.15**	**−0.05**	−0.02	**−0.12**	**−0.59**	**0.31**	**0.35**	**0.18**	1.00	**0.81**	**−0.08**
(25)	***RetVol***$_{j,t}$	0.01	0.00	**−0.17**	**−0.04**	**−0.04**	**−0.15**	**−0.64**	**0.32**	**0.50**	**0.20**	**0.79**	1.00	**−0.09**
(26)	***Guidance***$_{j,t}$	0.01	0.03	0.00	−0.01	−0.01	**0.03**	**−0.03**	**−0.12**	**−0.13**	**−0.03**	**−0.09**	**−0.09**	1.00

Note: All variables are defined in Appendix C; bold typeface indicates significance at the 1% level. Pearson's correlation coefficients are shown in the lower triangle (shaded), including the diagonal, and Spearman's rank correlations appear above the diagonal.

Table A14 Sample selection—Estimize coverage

Sample Selection	Firm-Quarter-Estimator	Firm
I/B/E/S Detail 2009-2018	**5,590,036**	**9,272**
Less: N-quarter ahead forecast (N>1)	(3,964,309)	(9,083)
No. of observations remained	1,625,727	8,848
Less: Forecasts made outside (−120,−1) window prior to earnings announcement	(55,891)	(6,833)
No. of observations remained	1,569,836	8,737
Less: multiple forecasts by an estimator	(456,647)	(6,958)
No. of observation remained	1,113,189	8,737
Less: firms with non-consecutive observations (7 years: 3 years before and after add to Estimize)	(276,294)	(5,915)
No. of observations remained	836,895	2,822
Less: firms with missing book value of equity or firms with stock price less than $5 in the year prior to the Estimize's introduction	(132,086)	(596)
No. of observations remained	**704,809**	**2,226**
No. of observations without Estimate coverage	82,513	539
No. of observations with Estimate coverage	622,296	1,687
2012 Additions	*401,155*	*825*
2013 Additions	*154,470*	*519*
2014 Additions	*23,440*	*108*
2015 Additions	*10,975*	*61*
Note: The period of interest in this study. Because I require (−3,3) year data around addition year.		
2016 Additions	*29,304*	*161*
2017 Additions	*2,717*	*12*
2018 Additions	*235*	*1*
Compustat 2009-2018		
Less: without annual fundamentals	(239,930)	(727)
No. of observations remained	**464,879**	**1,499**
CRSP 2009-2018		
Less: without beta or volatility or trading volume or return	(287)	(0)
No. of observations remained	**464,601**	**1,499**
TR Institutional Holding 2009-2018		
Less: without institutional holding	(59,026)	(99)
No. of observations remained	**405,575**	**1,400**
No. of Firm-Quarter observations	**39,814**	**1,274**
Less: missing controls	(0)	(0)
No. of useable observations	**39,814**	**1,274**
PSM (without replacement)		
Treat=1	6,890	519
Treat=0	6,890	594
Less: absolute difference in pscore >0.5%	(404)	(2)
Treat=1	6,486	517
Treat=0	6,486	562
No. of PSM observations	**12,968**	**1,161**

Table A15 The probability of analysts' forecast errors (at the individual analyst level) being slightly pessimistic is greater when analysts' prior forecast revisions follow downward revisions in investor earnings expectations (Full sample)

	DV = $MeetBeat_{i,j,t}$			
	Column I: Full sample		**Column II: Exclude standalone guidance**	
Variable	**Coef.**	**t-statistics**	**Coef.**	**t-statistics**
$RevDown_{i,j,t}$	−0.160	−1.217	−0.215	−1.549
$EstmzRev_{j,t}$	−0.147	−1.387	−0.134	−1.211
$Follow_{j,t}$	0.191	1.235	0.184	1.084
$Dispersion_{j,t}$	−0.777	−1.937*	−0.865	−1.938*
$AccScore_{i,j,t}$	0.704	6.252***	0.763	6.123***
$BoldScore_{i,j,t}$	0.098	0.699	0.118	0.811
$BiasScore_{i,j,t}$	−0.465	−2.952***	−0.510	−3.080***
$FirmExp_{i,j,t}$	0.034	0.261	0.057	0.416
$IndusExp_{i,j,t}$	−0.001	−0.007	0.057	0.273
$GenExp_{i,j,t}$	0.108	0.649	0.048	0.264
$BrokerSize_{i,t}$	−0.239	−2.272**	−0.230	−2.023**
$FirmSize_{j,t}$	1.892	1.947*	1.682	1.654*
$BtM_{j,t}$	0.347	1.404	0.320	1.190
$Loss_{j,t}$	0.539	1.948*	0.513	1.862
$Leverage_{j,t}$	0.291	0.597	0.399	0.826
$Turnover_{j,t}$	−0.389	−0.712	−0.418	−0.744
$RetVol_{j,t}$	0.462	1.063	0.390	0.867
$Guidance_{j,t}$	0.416	1.727*	−0.136	−0.612
Regression Type	Logistic		Logistic	
No. of obs.	8416		7918	
Year-Quarter FE	Yes		Yes	
Firm FE	Yes		Yes	
Pseudo R Square	0.185		0.186	

Note: Column I reports the regression results using all 8,416 observations. Column II reports the regression results using observations without management guidance for year quarter *t* announced during the [−90, −1]-day period for actual earnings announcement date. $MeetBeat_{j,t}$ is a binary variable that equals 1 if the raw forecast error (i.e., analyst *i*'s EPS forecast [−30, −1] days before the actual earnings announcement for firm *j* in quarter *t* minus the actual EPS) is [−0.02, 0]; 0 otherwise; $RevDown_{j,t}$ is a binary variable that equals 1 if $Revision_{i,j,t}$ is less than 0; 0 otherwise. All variables are defined in Appendix C; *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively, for two-tailed tests.