Watching from the Sky:
Business Observability and Voluntary Disclosure

Clark Liu
PBC School of Finance, Tsinghua University
Email: liuyue@pbcsf.tsinghua.edu.cn

Yancheng Qiu
Hong Kong University of Science and Technology
Email: yqiuak@connect.ust.hk

Shujing Wang
School of Economics and Management, Tongji University
Email: shujingwang@connect.ust.hk

P. Eric Yeung*
Samuel Curtis Johnson Graduate School of Management, Cornell University
Email: eric.yeung@cornell.edu

Current version: April 2022

Abstract

Exploiting the staggered releases of satellite data on parking lot traffic across U.S. retailers, we study how improved business observability affects corporate voluntary disclosure. We document that following satellite traffic data release, firms significantly suppress issuing management forecasts, especially good news forecasts. This result is best explained by management's incentive to avoid missing its own earnings "guidance" after business observability improves. Consistent with this explanation, we document that good news suppression concentrates among quarterly guidance and is more pronounced when operating uncertainty, institutional ownership, or expected litigation risk are higher. When the management decides to issue good news forecasts, these forecasts also become more qualitative.

Keywords: satellite data, business observability, voluntary disclosure, management forecasts, earnings targets

JEL Classification: G30; M41

* We thank helpful comments from conference participants at the 2020 Five-Star Workshop in Finance, the 1st Conference of Financial Reporting and Disclosures at Xijing University (2020), 2021 China Fintech Research Conference, 2021 China International Risk Forum (Big Data Research Excellence Award Winning Paper), the 17th Annual Haskell & White Academic Conference (2021), and 2022 Hawaii Accounting Research Conference, and seminar participants at Beijing Institute of Technology, Modern Risk Society, Peking University, Shanghai Jiao Tong University, and Tsinghua University for their valuable comments and suggestions. All remaining errors are our own.

Watching from the Sky:
Business Observability and Voluntary Disclosure

Abstract

Exploiting the staggered releases of satellite data on parking lot traffic across U.S. retailers, we study how improved business observability affects corporate voluntary disclosure. We document that following satellite traffic data release, firms significantly suppress issuing management forecasts, especially good news forecasts. This result is best explained by management's incentive to avoid missing its own earnings "guidance" after business observability improves. Consistent with this explanation, we document that good news suppression concentrates among quarterly guidance and is more pronounced when operating uncertainty, institutional ownership, or expected litigation risk are higher. When the management decides to issue good news forecasts, these forecasts also become more qualitative.

Keywords: satellite data, business observability, voluntary disclosure, management forecasts, earnings targets

JEL Classification: G30; M41

1. Introduction

Recent rapid technological advances and vast proliferation of data are leading to revolutionary changes in the way how information is produced, disseminated, and assimilated. Firms' voluntary disclosure is an important means for managers to communicate information with the capital markets and reflects the trade-off of various costs and benefits.[1] It remains less understood how the technological changes and fast-growing big data affect firms' disclosure decisions (e.g., Miller and Skinner 2015). In particular, as part of a growing trend, corporate activities become more observable thanks to the availability of alternative data such as satellite surveillance. Both public and private sectors are gearing up investments to develop the related technologies, and certain institutional investors have already started obtaining and utilizing timely information of firm performance provided by such technologies.[2]

Our research objective is to better understand how firms may alter their voluntary disclosures when their business activities become more observable to outsiders. In this paper, we investigate specifically how retail firms change voluntary disclosure practices when their business activities become more observable due to the *release* of high-resolution satellite imagery data on their parking lot traffic (hereafter "satellite traffic data"). The release of satellite traffic data for retail firms provides a powerful setting because parking lot traffic links closely to retailers' performance, at least during our sample period.

To our knowledge, we are among the first to utilize the precise timing of the release of satellite data to investors, which allows us to empirically identify the impact of improved business observability on changes in firm voluntary disclosure.[3] We obtain the timing of the staggered releases of satellite

---

[1] Corporate disclosure has emerged from decades of research as an important way to mitigate information asymmetry and agency conflicts between managers and outside investors (e.g., Healy and Palepu 2001; Beyer, Cohen, Lys, and Walther 2010; Leuz and Wysocki 2016).

[2] For example, by analyzing the satellite imagery data provided by Orbital Insight, J.P. Morgan claimed that Whole Foods' parking lots were busier after Amazon bought Whole Foods in August 2017. For the month of September 2017, car traffic at Whole Foods' stores was up 4.6 percent, the biggest increase since June 2014. In contrast, Kroger had been adversely affected by the Amazon-Whole Foods deal and experienced a 7.3 percent decrease of car traffic in the same month. J.P. Morgan predicted that Kroger's same-store sales would be down 0.8 percent in the entire third quarter of 2017, much lower than what analysts anticipated. See, https://www.cnbc.com/2017/10/04/whole-foods-parking-lots-were-busier-post-amazon-satellites-show.html, for more details.

[3] For instance, Zhu (2019) designates June 30, 2014 as the data release date for all firms and finds that, while stock prices become more informative with the satellite traffic data, there are no significant changes in management forecasts. Zhu (2019, footnote 20) discusses the findings regarding voluntary disclosure and explicitly acknowledges her data limitations on page 2032.

imagery data for publicly listed U.S. retailers from major data vendors specializing in satellite imaging for parking lot traffic. The important function of the high-tech data vendors is that they use advanced technology and machine learning algorithms to extract meaningful and ready-to-use data from a variety of location-based satellite imagery sources.

Using the within-firm and within-year generalized difference-in-differences (DID) method based on staggered release events during the period 2011-2018, we identify the causal effect of the release of satellite traffic data on management forecasts. We document that treated firms decrease their management forecasts by about 26% relative to control firms following the release of satellite traffic data. We conduct a wide range of assessments to strengthen a causal interpretation. This result cannot be explained by the general substitution between observability and voluntary disclosure (e.g., Diamond and Verrecchia 1991), as the decrease is concentrated in good news forecasts (relative to analysts' expectation). Other determinants of voluntary disclosures commonly found in the literature also fail to explain this decrease in good news forecasts.[4]

We propose a "meeting guidance" hypothesis to account for the disappearing good news forecasts, which is motivated by the management's strong incentive not to miss its own earnings guidance (e.g., Graham, Harvey, and Rajgopal 2005). Our basic idea is that, as business observability improves price informativeness (and the earning expectations imbedded in prices), the risk of missing its own earnings guidance becomes higher if the management issues good news forecasts, leading to less propensity to issue good news forecasts. By contrast, the propensity to issue bad news forecasts remains unchanged conditional on improved price informativeness.

Consider the following illustrative example. Assume that the market expects an earnings per share (EPS) of \$2 at Time 0, and the EPS realization at Time 2 will be \$3 under "good state" or \$1 under "bad

[4] These explanations suggest that higher business observability *increases* voluntary disclosure, inconsistent with what we find. First, business observability reduces the uncertainty regarding the management's possession of information, which forces the managers to disclose (Dye 1985; Jung and Kwon 1988), leading to more management forecasts. Second, managers face strong incentives, such as career concerns, to withhold bad news and gamble that subsequent events will turn in their favor (e.g., Kothari, Shu, and Wysocki 2009). Business observability makes it more difficult for managers to hide bad news, resulting in more bad news disclosure. Third, the proprietary cost concerns regarding voluntary disclosure should decrease as competitors also better observe its business (Verrecchia 1983; Darrough and Stoughton 1990; Teoh and Hwang 1991), which in turn predicts more management forecasts.

state". The management observes a private signal about upcoming EPS at Time 1. Assume for simplicity that, without the satellite traffic data, the market remains its EPS expectation at \$2 at Time 1. To keep investors informed, the management chooses to issue an EPS guidance at Time 1. Due to the post-forecast uncertainty, the management has the incentive to be conservative when issuing the guidance. Assume that its guidance is \$2.7 under good state or \$0.7 under bad state, creating a 30-cents "buffer zone" (in either case) to meet its own guidance.

With the satellite traffic data, however, the market is more informative, and its EPS expectation will move closer towards actual EPS at Time 1. Assume that the revised market expectation is \$2.8 under good state or \$1.2 under bad state at Time 1. This incremental informativeness creates *asymmetric* incentive for the management: It becomes less willing to further "talk up" EPS expectation under the good state for fear of missing its own guidance (i.e., \$2.8 is already very close to \$3); By contrast, the incentive of "walking down" EPS expectation remains under the bad state (i.e., \$1.2 is still above \$1).

Under our meeting guidance hypothesis, improved business observability causes the management to become more averse to issuing good news forecasts relative to issuing bad news forecasts under the following three conditions: i) The market revises its expectation under the partial adjustment model (i.e., a Bayesian fashion with non-zero weight on its prior expectation);[5] ii) The management has relatively more precise information about EPS than the market does (e.g., costs-related information that remains unobservable to investors); And iii) management's payoffs are asymmetric with respect to meet and miss its own guidance.

Our hypothesis generates the following specific predictions. If the management is concerned about the risk of missing its own guidance, the suppressed good news forecasts should be i) targeting more at short-horizon guidance than at long-horizon forecasts, and ii) more pronounced when the post-forecast uncertainty is higher. Both of these predictions stem from management's risk-aversion. Third, as stock prices become more informative (i.e., market expectation gets closer to the actual EPS), the asymmetric reduction should be more pronounced (i.e., talking up is riskier but walking down remains beneficial). Fourth, the asymmetric reduction in management forecasts should be more pronounced when the costs

[5] Katona, Painter, Patatoukas, and Zeng (2021) find that stock prices do not fully incorporate the value-relevant information embedded in satellite imagery data before the public disclosure of retail firms' performance.

of missing EPS guidance are higher for the firm/management. Consistent with all four predictions, we find that the asymmetric reduction is more pronounced i) when management forecasts serve as short-horizon quarterly guidance, ii) when cash flows volatility is higher, iii) when institutional ownership is higher, and iv) when litigation risk (associated with large stock price declines) is higher.

Our hypothesis also yields the unique prediction that management suppresses bad news forecasts if the market is already overly pessimistic about firm performance. In the bad state of our illustrative example, if the market already revises its EPS expectation down to \$0.8 (<\$1), the management has no incentive to further walk down the market expectation. Consistent with this prediction, we find that when the firm has experienced extremely negative abnormal stock returns, bad news forecasts decrease significantly following the release of satellite traffic data.

Our last analysis focuses on the precision of issued management forecasts. Our hypothesis predicts that, as the management becomes more concerned with missing its own guidance, it may be more likely to issue qualitative good news forecasts to avoid providing "smoking gun" evidence against its reputation. Consistent with this prediction, we find that the management is more likely to issue qualitative forecasts (i.e., replacing more precise quantitative forecasts) after the release of satellite traffic data.

One caveat of our empirical analyses is that, in line with Bertrand, Duflo, and Mullainathan (2004), Roberts and Whited (2013), and Yagan (2015), we use the term DID simply to describe a model that compares trends in disclosure policy between different groups of treated and control observations. We do not claim that the assignments to treated and control groups are purely random. We are aware of the fact that the selection procedure of satellite data release is based on many factors, such as business model, technological feasibility, the availability of satellite images, and so on. The identifying assumption underlying our model specification is that, in the absence of satellite data release, the temporal trend in retail firms' disclosure policy would have been the same between treated and control groups. We take several steps to mitigate the endogeneity concern. First, we verify the parallel trend assumption by showing that the pre-treatment trends are indistinguishable between the treated and control firms. Second, we present evidence that it is the data release date that causes the change in disclosure policy as opposed to the data start date, which is usually earlier than the release date.[6] Third, we show that our results remain similar when we match each treated firm with a similar control firm in the same industry based on the propensity-score-matching (PSM) procedure. Fourth, we demonstrate that the initiation of satellite data release has little relation with the underlying firm's prior disclosure policy. Fifth, we implement the Callaway and Sant'Anna (2021) estimator and verify that our conclusion is robust to potential biases in staggered DID regressions due to treatment effect heterogeneity.

Our study contributes to the literature in several ways. First, our paper adds to the research on the growing importance of technology advancements and alternative data for capital markets and corporate decisions. Recent studies focus on the information content of alternative data (Zhu 2019; Kang, Stice-Lawrence, and Wong 2021; Katona, Painter, Patatoukas, and Zeng 2021). Our paper provides fresh evidence on how the availability of alternative data influences voluntary disclosures. Second, our findings enrich the literature concerning the impact of information environment on corporate voluntary disclosures by focusing on third-party big data suppliers. Prior studies focus on traditional institutions such as sell-side analysts, institutional or active investors, and banks,[7] and there has been little evidence on how firms' voluntary disclosure is exogenously affected by the growing third-party big data suppliers (Goldstein, Spatt, and Ye 2021). Our paper provides new findings on managers' response to exogenous shocks in their information environments due to the emergence of third-party big data vendors.

2. Literature Review

2.1 Technology development and big data

The recent development of technologies and the emergence of big data have significantly changed the way information is produced and disseminated. Miller and Skinner (2015) discuss the new forces

[6] When the data vendors commercially release the satellite data of a specific company to its customers, the release date is generally later than the "start date" of the satellite traffic data (i.e., satellite traffic data is already possessed by the data vendor but not yet released to investors). For example, when RS Matrix released the satellite traffic data of Best Buy Co. on March 19, 2013 to its customers, the satellite traffic data actually starts from October 30, 2010.

[7] For example, see Langberg and Sivaramakrishnan (2008) and Balakrishnan, Billings, Kelly, and Ljungqvist (2014) regarding analysts, Ajinkya, Bhojraj, and Sengupta (2005), Jung (2013), Boone and White (2015), Bourveau and Schoenfeld (2017), and Khurana, Li, and Wang (2018) regarding institutional or active investors, and Lo (2014), Vashishtha (2014) and Chen and Vashishtha (2017) regarding banks.

such as the increased penetration of mobile devices and the role of social media that are likely to change disclosure in important ways and call for future research that continue to explore these issues. For example, Brown, Stice, and White (2015) show that mobile communications significantly influence local information flow and local investor activities. Froot, Kang, Ozik, and Sadka (2017) develop a real-time corporate sales index for U.S. retail stores from mobile devices and show that it can explain future releases of firm fundamentals and future stock returns. Blankespoor, Miller, and White (2014) demonstrate that the dissemination of firm-initiated news via Twitter improves firms' market liquidity. Chen, De, Hu, and Hwang (2014) find that individual opinions of financial securities shared on one of the most popular social media platforms in the U.S. can predict future stock returns and earnings surprises. Lee, Hutton, and Shu (2015) show that corporate social media, in general, attenuates the negative price reaction to product recall announcements. Dai, Parwada, and Zhang (2015) suggest that social media plays a disciplining role in restricting insider trading profits.

Satellite imagery has recently emerged as an important data source for economic studies. For example, satellite data has been used to measure economic activities (Henderson, Storeygard, and Weil 2012), weather and climate variations (Guiteras, Jina, and Mobarak 2015), land usage and crop choice (Holmes and Lee 2012), natural resources (Foster and Rosenzweig 2003), and pollution (Foster, Gutierrez, and Kumar 2009).

However, only recently has *high-resolution* satellite imagery data been used to investigate firm performance. Limited examples include Zhu (2019), who shows that the availability of satellite imagery and consumer transactions increases stock price informativeness and disciplines managers' actions by reducing opportunistic insider trading and improving investment efficiency. Kang et al. (2021) use satellite data to examine the effects of information acquisition costs and find that investors adjust their holdings more in response to local performance. Katona et al. (2021) find that stock prices do not fully incorporate the value-relevant information embedded in satellite imagery data before the public disclosure of retail firms' performance and unequal access to alternative data among market participants does not necessarily facilitate stock price discovery.

While existing studies mainly focus on how the use of alternative data impacts financial markets and asset management, there is scarce evidence on how firms react to the big data revolution when making real corporate decisions (Goldstein et al. 2021). We contribute to this line of literature by providing new evidence on how the emergence of satellite data impacts corporate voluntary disclosure.

2.2 Information environment and firm voluntary disclosure

Existing studies provide evidence consistent with the notion that market participants' demand and supply of information exert important influence on corporate voluntary disclosure. On the demand side, a number of studies investigate the role of institutional investors. It is suggested that firms issue more voluntary disclosures when they have higher institutional ownership (Ajinkya et al. 2005; Jung 2013; Boone and White 2015) or when they are at risk of attack by activist investors (Bourveau and Schoenfeld 2017). Khurana et al. (2018) further show that target firms decrease bad news disclosure during hedge fund interventions. Regarding banks, Lo (2014) shows that borrowers whose banking relationship is threatened by declining bank health increase both the quantity and informativeness of their voluntary disclosures. Vashishtha (2014) finds that firms reduce disclosures following covenant violations when bank monitoring is enhanced. Chen and Vashishtha (2017) show that borrowers increase disclosures after their lending banks engage in mergers.

On the supply side, consistent with the substitution effect, Balakrishnan et al. (2014) show that after firms lose analyst coverage due to plausible exogenous shocks and thus face decreased supply of public information, they provide more voluntary disclosures. Our paper adds to this strand of research by investigating how managers respond to exogenous shocks of their information environment when satellite traffic data of retail firms is released by third-party data vendors.

3. Data and Sample

3.1. Data sources

We obtain satellite imagery data of parking lot traffic for U.S. retailers from two major data vendors, RS Metrics (RS) and Orbital Insight (OB).[8] RS Metrics is the first major data vendor that has released real-time parking lot traffic data of a wide range of U.S. retail firms based on satellite imagery starting from the first quarter of the year 2011. Orbital Insight is the most prominent competing data

[8] J. P. Morgan's Big Data and AI Strategies report provides a summary of major data vendors, which is available at https://www.cfasociety.org/cleveland/Lists/Events%20Calendar/Attachments/1045/BIG-Data_AI-JPMmay2017.pdf.

vendor, which started to release similar data in the second quarter of the year 2015 and covered the largest number of U.S. retail firms at the end of 2018. The data vendors obtain satellite imagery from diverse imagery providers, apply deep learning algorithms to classify information or detect objects in each image, and utilize data science to contextualize each observation. The images have a high resolution of nearly 50 square centimeters per pixel, which allows the data vendors to accurately distinguish cars from neighboring objects. Kang et al. (2021) provide a comprehensive review of the process based on which Orbital Insight use proprietary algorithms to extract car count information from satellite images. The final data consists of daily store-level and firm-level information about parking lot car counts and parking lot utilization across major U.S. retailers. We further obtain detailed information from the two data vendors on the exact time when the satellite imagery data of each retail firm started to be released to their clients. As disclosed by the data vendors, their clients include hedge funds, asset management firms, government entities, and nonprofit organizations.

We obtain management forecast and analyst forecast data from Institutional Brokers' Estimate System (I/B/E/S). Share price and trading data is obtained from the Center for Research in Security Prices (CRSP) and accounting data from the Compustat database. Information of institutional investors is obtained from the Thomson Reuter's 13F database. Equity and debt issuance data is obtained from SDC Platinum database.

3.2. Sample construction and summary statistics

We combine the satellite data from RS and OB, and generate a comprehensive dataset covering all the U.S. retail firms whose satellite data has been released by the two major data vendors from 2011 to 2018. We obtain 142 unique retail firms after merging the satellite data with the CRSP-Compustat data. RS releases data for 48 firms and OB for 139 firms, among which the data of 45 firms are released by both data vendors. RS and OB release the data of these retail firms at different times. When the satellite data of a retail firm has been released by both data vendors, we use the earlier time as the event time. Figure 1 presents the number of satellite data release events from 2011 to 2018. It is evident that the release of satellite data is staggered over time, with the highest numbers in 2016 mainly due to the reason that OB substantially expanded its data coverage in that year.

Similar to Katona et al. (2021), we identify control firms as those with the same six-digit Global

Industry Classification Standard (GICS) codes as the treated firms, which span over 13 GICS industries.[9] We select control firms from the same industry as the treated firms to difference away unobserved industry-specific trends. Our sample starts in 2008, three years before the first release event of satellite imagery data, and ends in 2018. To mitigate the influence of newly listed firms, we delete firm observations in the first two years after the initial public offering (Fama and French 1993).

We further merge the sample with management forecast data provided by I/B/E/S guidance. Our final sample with available management forecast data covers 117 treated firms in 12 GICS industries whose satellite imagery data are gathered and eventually released by either of the two commercial satellite data vendors and 526 control firms in the same industries whose satellite data has not been released by the end of 2018. Table 1 presents the industry distribution of treated firms before (Panel A) and after (Panel B) merging the sample with management forecast data.

The final sample consists of 5,515 firm-year observations, including 1,393 observations of treated firms and 4,122 observations of control firms. We winsorize all continuous variables at the 1% and 99% levels to minimize the influence of outliers. Table 2 presents the summary statistics of variables used in the paper for the full sample, the treated sample, and the control sample, respectively. Following previous literature (e.g., Baginski and Rakow 2012; Bourveau, Lou, and Wang 2018; Houston, Lin, Liu, and Wei 2019), our main measure of voluntary disclosure is the natural logarithm of the total number of management forecasts issued over a fiscal year plus one, denoted as *Ln*(*#MgmtFcst+1*). We also construct similar measures based specifically on sales forecasts, *Ln*(*#MgmtSalesFcst+1*), as satellite data of parking lot traffic is closely related to sales numbers. All our measures are based on both annual and quarterly forecasts aggregated at the firm-year level. The detailed construction of empirical variables is described in Appendix A.

4. Research Design

We exploit the staggered release of satellite data on parking lot traffic as an exogenous shock to the information asymmetry between corporate insiders and outside investors for publicly listed U.S. retail firms. The availability of satellite data provides outside investors with near real-time information

[9] Our analyses remain qualitatively similar when control firms are selected based on the two-digit Standard Industrial Classification (SIC).

on parking lot traffic, which is an important proxy for sales growth and operating performance and significantly improves the business observability of U.S. retail firms. The staggered nature of satellite data release provides a set of counterfactuals for how corporate voluntary disclosure would have evolved in the absence of satellite data and so allows us to disentangle the effect of satellite data from other forces shaping the decisions of corporate disclosure policy. In addition, satellite data is provided by third-party data vendors and out of managers' control, which is more likely to be exogenous from the viewpoint of individual firms. We use within-firm and within-year generalized DID regression models to control for unobserved firm attributes and temporal trends in corporate disclosure policies that are unrelated to satellite data coverage. The specification is as follows:

$$Disclosure_{i,t} = \alpha_i + \alpha_t + \beta Released_{i,t} + \gamma \boldsymbol{z_{i,t}} + \epsilon_{i,t}, \tag{1}$$

where $Disclosure_{i,t}$ is a measure of voluntary disclosure of firm *i* in year *t*; $Released_{i,t}$ is a dummy variable that equals one if the satellite data has been released for firm *i* by time *t*, and zero otherwise; α_i and α_t represent firm and year fixed effects, respectively; $\boldsymbol{z_{i,t}}$ is a vector of control variables.

The DID coefficient estimate β on $Released_{i,t}$ captures the effect of the information shock caused by satellite data release on retail firms' corporate voluntary disclosure. Including firm fixed effects (α_i) ensures that the DID estimate reflects average within-firm changes in voluntary disclosure when satellite data becomes available to outside investors. Year fixed effects (α_t) control for concomitant aggregate economic conditions and management forecast trends. To account for the potential cross correlations within firms and over time, standard errors are clustered at both the firm and year levels.

Following previous literature (e.g., Ali, Klasa, and Yeung 2014), our control variables ($\boldsymbol{z_{i,t}}$**)** include stock return volatility (*RetVol*), absolute change in annual earnings per share scaled by stock price (*AbsChEPS*), market-adjusted stock return (*MktAdjRet*), research and development expense scaled by book assets (*R&D*), the natural logarithm of market value of equity (*Size*), analyst coverage (*Coverage*), institutional fractional ownership (*InstOwn*), equity or debt issuance dummy (*D_Issuance*), market-to-book equity ratio (*MB*), leverage (*Lev*), standard deviation of earnings (*StdEarn*), positive earnings change dummy (*D_PosChEarn*), and analyst optimism (*Optimism*). Table 2 shows the descriptive statistics of these variables for the full sample in Panel A, treated firms in Panel B, and control firms in Panel C.

The key identifying assumption that guarantees the consistency of the DID estimates is that conditional on all covariates and fixed effects, treated and control firms have parallel trends in the absence of satellite data release. In this case, the coefficient estimate on $Released_{i,t}$ gives the causal treatment effects of satellite data release on voluntary disclosure of retail firms. We perform extensive tests to validate this identifying assumption in the following section of empirical results.

5. Empirical Results

5.1 Baseline difference-in-differences results

We apply the generalized DID approach to identify the effect of staggered release of satellite data on corporate voluntary disclosure. We follow the specification in Equation (1) and present the baseline results on the frequency of management forecasts in Table 3. We present the results for management forecasts in columns 1 and 2 and sales forecasts in columns 3 and 4.

The DID coefficients on $Released_{i,t}$ are significantly negative across all specifications. Because we use natural logarithm to measure dependent variables, the coefficients can be approximately interpreted as percentage changes. For example, the estimated coefficient in column 1 is -0.286 (t-statistic = -4.24), indicating that following the release of satellite data, treated firms decrease their management forecast frequency by 28.6% relative to control firms. The DID estimate does not change much after firm-level controls are added in column 2, which shows a 26.1% decrease (t-statistic = -4.23). This indicates that the identified effect does not coincide systematically with firm-specific characteristics and the unobserved omitted variable bias is possibly limited (e.g., Altonji, Elder, and Taber 2005).

The coefficients on control variables are also consistent with prior literature (e.g., Ali et al. 2014). For instance, firms issue fewer management forecasts when they face greater uncertainty and make more forecasts when they are larger.[10] Next, we specifically investigate how the release of satellite data affects management sales forecasts. The DID coefficients on $Released_{i,t}$ show that following the release of satellite data, treated firms decrease their sales forecasts by more than 28% (statistically significant at the 1% level) as reported in columns 3 and 4.

[10] Because we add firm fixed effects and our sample is confined to retail firms, control variables become less statistically significant compared with prior work.

5.2 Major identification assessment

The causal interpretation of the DID estimate depends crucially on the parallel trend assumption. To validate that the baseline results are not driven by the pre-existing trend difference between treated and control firms, we examine the dynamic effect of satellite data release as suggested by previous studies (e.g., Bertrand and Mullainathan 2003; Roberts and Whited 2013). We replace $Released_{i,t}$ in Equation (1) with three dummy variables: i) $Released_{\{i,-2\leq t\leq -1\}}$ is a dummy variable that equals one if year t is within two years before the satellite data of firm i is released, and zero otherwise; ii) $Released_{\{i,0\leq t\leq 1\}}$ is a dummy variable that equals one if year t is within the year or one year after the satellite data of firm i is released, and zero otherwise; and iii) $Released_{\{i,t\geq 2\}}$ is a dummy variable that equals one if year t is two years or more after the satellite data of firm i is released, and zero otherwise.

Results are reported in Panel A of Table 4. The dummy variable $Released_{\{i,-2\leq t\leq -1\}}$ allows us to assess whether there is any effect on management forecasts prior to the release of satellite data. Finding such an "effect" before satellite data release could be a symptom of different pre-trends between treated and control firms. The results show that the coefficients on $Released_{\{i,-2\leq t\leq -1\}}$ are economically small and statistically insignificant for all specifications, suggesting that the pre-treatment trends are indistinguishable between the treated and control firms and thus validating the parallel trend assumption of our identification.

The coefficients on $Released_{\{i,0\leq t\leq 1\}}$ and $Released_{\{i,t\geq 2\}}$ allow us to track the effect on management forecasts after the release of satellite data over time. Consistent with a causal interpretation of our basic results, we find that these coefficients are significantly negative at the 1% level. And the estimated coefficients on $Released_{\{i,0\leq t\leq 1\}}$ are economically smaller than those on $Released_{\{i,t\geq 2\}}$, suggesting that it takes time for firms to adjust their voluntary disclosure policy in response to the information shock of satellite data release.

To further visualize the pre-trend test, we estimate the DID coefficient β_j prior to and after the event year of satellite data release by performing the following regression:

$$Disclosure_{i,t} = \alpha_i + \alpha_t + \sum_j \beta_j Released_{\{i,j\}} + \gamma X_{i,t} + \epsilon_{i,t}, \quad (2)$$

where j represents the event years from t = -3 to t >= 3, where t=0 is the event year of satellite data release. Figure 2 plots the estimated coefficients and 95% confidence intervals of β_j. Event years of $t <$ -3 are used as the benchmark. It is evident that before satellite data release (when $t <= -1$), the estimated coefficient is small and statistically insignificant, confirming that there is no significant difference in the pre-treatment trend between the treated and control groups. After satellite data release (when $t >= 0$), however, the estimated coefficient becomes significantly negative, indicating that there is a significant decrease in management forecasts of treated firms compared with control firms. Moreover, the magnitude of the estimated coefficient at $t = 0$ is economically smaller than those at $t > 0$, confirming that it takes some time for firms to change their voluntary disclosure practices.

The second identification check is the importance of the timing of data releases as opposed to the start of the data. As noted earlier, satellite traffic data may have been possessed by the data vendor but not yet released to the market participants, and we expect that it is the release date that causes the change in disclosure policy as opposed to the data start date. We thus perform the following DID regression:

$$Disclosure_{i,t} = \alpha_i + \alpha_t + \beta_1 Released_{i,t} + \beta_2 StartNotRelease_{i,t} + \gamma \boldsymbol{z}_{i,t} + \varepsilon_{i,t}. \tag{3}$$

We add $StartNotRelease_{i,t}$ to our main regression model in Equation (1), which is a dummy variable that equals one if the satellite data on parking lot traffic has started but not yet been released for firm i by year t, and zero otherwise.

The results reported in Panel B of Table 4 show that while the coefficients on $Released_{i,t}$ remain significantly negative across various specifications, the coefficients on $StartNotRelease_{i,t}$ are statistically and economically insignificant. These results confirm that when satellite data is possessed by the data vendor but not released to the market participants, the data has no impact on the frequency of management forecasts. In other words, only after the satellite data is released to the market participants, is there a significant decrease of management forecasts for treated firms relative to control firms.

As a third identification check, we match each treated firm with a similar control firm in the same industry based on the propensity-score-matching (PSM) procedure. We implement the PSM procedure by first estimating a logit regression to model the probability that the satellite data of a retailer is released based on all control variables in the baseline regression. We then match each treated firm to a control firm in the same industry using the nearest neighbor matching technique with no replacement. In unreported tables, we show that the treated and matched control firms have similar firm characteristics. The results based on the PSM sample are reported in Panel C of Table 4. The estimated coefficients on $Released_{i,t}$ remain economically large and statistically significant.

5.3 Management EPS forecast frequency of good and bad news

In this section, we investigate the effect of satellite data release on the disclosure of good news and bad news earnings forecasts separately, which are bifurcated based on analyst consensus forecasts. In all the good news vs. bad news analyses, we focus specifically on management EPS forecasts because analysts' non-EPS forecasts (e.g., net income, return on assets, etc.) are scarce.

To test the effect of satellite data release on the two types of forecasts, we employ the same DID regression model in Equation (1) for good news and bad news EPS forecast frequency, respectively. Following previous literature (e.g., Kothari et al. 2009; Ali, Li, and Zhang 2019), good news and bad news EPS forecasts are defined as management forecasts of quarterly or annual EPS that are above or below the analysts' most recent consensus forecasts.[11] We follow Rogers and Van Buskirk (2013) and construct updated analyst forecast consensus by incorporating the latest information in earnings announcements.[12] First, we use the parameters of the analyst revision model in Rogers and Van Buskirk (2013) to estimate a revision that the analysts should make given the new information in the earnings announcement.[13] Second, we update analyst forecast consensus by adjusting for the analyst revision. Finally, we classify bundled management forecasts into good or bad news forecasts based on the updated analyst forecast consensus.

Table 5 reports the estimation for good news and bad news EPS forecast frequency measured by

[11] Specifically, a management forecast is classified as good (bad) news if the point estimate, or the midpoint of the range forecast, is above (below) the analyst consensus forecast before the management forecast. For open-ended management forecasts, the forecast is classified as good (bad) when its bottom (upper) bound is higher (lower) than the analyst consensus forecast.

[12] In recent years, especially after the passage of Regulation Fair Disclosure (Reg FD) in 2000, an increasing number of management forecasts are issued in conjunction with earnings announcements, which are referred to as bundled forecasts (Anilowski, Feng, and Skinner 2007). The conventional classification of forecast news may be biased for bundled forecasts due to measurement errors.

[13] We also estimate the analyst revision model as in Rogers and Van Buskirk (2013) during our sample period from 2008 to 2018, and use the estimated model parameters to calculate analyst revision given the new information in the earnings announcement. The results are qualitatively similar and available upon request.

the natural logarithm of the number of good news or bad news management EPS forecasts plus one. It is evident that the DID coefficients on $Released_{i,t}$ are negative for good news forecasts (e.g., -0.187 with *t*-statistic of -3.26 in column 2). This indicates that compared to the control firms, treated firms experience a significant decrease in the frequency of good news EPS forecasts after their satellite data is released. In contrast, the DID coefficients on $Released_{i,t}$ are close to zero and statistically insignificant for bad news forecasts (e.g., -0.059 with *t*-statistic of -1.32 in column 4), indicating that treated firms do not experience a significant decrease in bad news forecasts, measured relative to control firms.

Taken together, our findings suggest that the decrease in management forecasts after satellite data release is mainly driven by suppressed good new forecasts. This one-sided decrease cannot be explained by the general substitution relation between observability and voluntary disclosure suggested by, for example, Diamond and Verrecchia (1991).

5.4 Predictions from the meeting guidance hypothesis

We propose a "meeting guidance" hypothesis to account for the asymmetric decrease in management forecasts. Our basic idea is that, conditional on investors already observing good traffic data, the risk of missing its own earnings guidance becomes higher if the management further talks up earnings expectations. On the other hand, conditional on investors observing bad traffic data, the management's incentive to walk down earnings expectation remains given that the market revises its expectation under the partial adjustment model. In this section, we test the predictions from this meeting guidance hypothesis.

5.4.1 Quarterly Guidance vs. Long-term Forecasts

The meeting guidance hypothesis predicts that the management is more averse to missing its quarterly guidance issued near fiscal end than long-term forecasts, which are less viewed by the market as earnings target since the businesses may change substantially between the forecast date and fiscal end. To test this prediction, we define "quarterly guidance" as those forecasts with horizon less than or equal to 90 days (from the fiscal end) and "long-term forecasts" as those forecasts with horizon greater than 90 days. We test how much the management reduces its quarterly guidance or long-term forecasts after satellite traffic date release by estimating DID regression model in Equation (1).

Table 6 reports the results. The dependent variables are the natural logarithm of the number of quarterly guidance plus one in columns 1 and 2 and the natural logarithm of the number of long-term forecasts in columns 3 and 4. We document that the negative coefficients on $Released_{i,t}$ are larger and carry higher statistical significance for quarterly guidance. For instance, quarterly guidance decreases by nearly 26% (column 2), but long-term forecasts decrease by about 17% (column 4).

We further test whether the decrease in good news EPS forecasts is more pronounced for quarterly guidance. In columns 5 and 6, the dependent variables are the natural logarithm of the number of good news EPS quarterly guidance plus one, while the dependent variables in columns 7 and 8 are the natural logarithm of the number of good news EPS long-term forecasts. As expected, the estimated coefficients on $Released_{i,t}$ are more negative and carry higher statistical significance for good news EPS quarterly guidance than for good news EPS long-term forecasts. The coefficients suggest that the decrease in quarterly guidance is twice as large as that for long-term forecasts.[14]

5.4.2 Forecast Uncertainty

The management should be more averse to issuing good news forecasts if the post-forecast uncertainty is higher such that it is more likely to miss its own good news guidance. Note that uncertainty should not reduce the management's incentive to issue bad news forecasts (and perhaps walking down market expectation becomes even more important for risk-averse management).

We thus examine how the negative relationship between satellite data release and management forecast frequency varies with the uncertainty of operating environment. We use cash flow volatility as the proxy for firm-level operating uncertainty, which is defined as the coefficient of variation in a firm's quarterly operating cash flow over the past six-year period following Minton and Schrand (1999). We estimate the following DID regression of management forecast frequency by adding an interaction term between satellite data release and operating uncertainty to the baseline regression model in Equation (1):

$$Disclosure_{i,t} = \alpha_i + \alpha_t + \beta_1 Released_{i,t} + \beta_2 Released_{i,t} \times X_{i,t} + \beta_3 X_{i,t} + \gamma \boldsymbol{z_{i,t}} + \epsilon_{i,t}, \qquad (4)$$

where $X_{i,t}$ represents a variable that interacts with $Released_{i,t}$ and captures high operating uncertainty in this test ($HighCFVol_{i,t}$), which is a dummy variable that equals one if firm *i* has high cash flow volatility

[14] Untabulated results verify that estimated coefficients on $Released_{i,t}$ are insignificant when we use bad news EPS quarterly guidance or bad news EPS long-term forecasts as the dependent variables.

(above the median) at the end of year t-1, and zero otherwise.

Results are reported in Panel A of Table 7. In columns 1 and 2, the coefficients on both $Released_{i,t}$ and $Released_{i,t} \times HighCFVol_{i,t}$ are significantly negative for management forecasts. The results suggest that after satellite data release, treated firms significantly decrease their management forecast frequency and such decrease is more pronounced for high-uncertainty firms. More importantly, this negative effect stems mainly from good news forecasts. The estimated coefficients on $Released_{i,t} \times HighCFVol_{i,t}$ are -0.176 (t-statistic = -2.24) and -0.172 (t-statistic = -2.16) for good news forecasts in columns 3 and 4, but are much smaller in magnitude (and statistically insignificant) for bad new forecasts in columns 5 and 6. The insignificant results for bad news forecasts suggest that the management's incentive to walk down market expectation remains strong under high operating uncertainty.

5.4.3 Institutional Ownership

As stock prices become more informative (i.e., market expectation gets closer to the actual EPS), the asymmetric reduction in management forecasts should be more pronounced (i.e., talking up becomes riskier but walking down remains beneficial). Because the high prices of satellite traffic data can likely be afforded only by large institutional investors, we expect that stock prices are more informative with respect to this data when institutional ownership is higher. This reasoning predicts that the asymmetric reduction is more pronounced when institutional ownership is higher.

To test this prediction, we estimate the DID regression of management forecast frequency by adding an interaction term between satellite data release and institutional ownership. We set $X_{i,t}$ in Equation (4) as high institutional ownership ($HighInstOwn_{i,t}$), which is a dummy variable that equals one if firm i has high institutional ownership (above the median) at the end of year t-1, and zero otherwise.

Results are reported in Panel B of Table 7. In columns 1 and 2, the dependent variable is the frequency of all management forecasts. The coefficients on $Released_{i,t}$ remain significantly negative, while the coefficients on $Released_{i,t} \times HighInstOwn_{i,t}$ are small and statistically insignificant. However, when we focus on good news EPS forecasts frequency in columns 3 and 4, the estimated coefficients on $Released_{i,t} \times HighInstOwn_{i,t}$ are significantly negative, suggesting that after satellite data release, treated firms decrease good news EPS forecasts significantly more when their institutional ownership is high. For instance, the estimated coefficient in column 4 suggests that the decrease in good news EPS forecasts for firms with high institutional ownership is 15.2% more than that for firms with low institutional ownership. The coefficients on $Released_{i,t}$ become small and statistically insignificant, suggesting that the release has attenuated effects on good news EPS forecasts for firms with low institutional ownership. Moreover, as expected, institutional ownership has no impact on bad news EPS forecasts. Specifically, the estimated coefficients on $Released_{i,t} \times HighInstOwn_{i,t}$ are close to zero in columns 5 and 6.

5.4.4 Reputational Costs of Missing Guidance

The meeting guidance hypothesis also predicts that the asymmetric reduction in management forecasts should be more pronounced when the reputational costs of missing guidance to the firm or management are higher. We use expected litigation risk as the proxy for the risk of losing reputation with the markets (Kim and Skinner 2012), which captures the management's incentive to avoid missing its own guidance. By contrast, firms with higher ex-ante litigation risk remain incentivized to issue bad news forecasts to avoid any potential litigation associated with large stock price decline following missing earnings targets.

To estimate the expected litigation risk faced by individual firms, we calculate the probability of litigation based on the coefficient estimates reported in model 3 of Table 7 by Kim and Skinner (2012). We classify our sample into firms with high (above median) and low (below median) expected litigation risk. We then set $X_{i,t}$ in the DID regression model of Equation (4) as $HighLitRisk_{i,t}$, which is a dummy variable that equals one if firm *i* has high expected litigation risk (above the median) at the end of year *t*-1, and zero otherwise.

Panel C of Table 7 reports the results for management forecasts (columns 1 and 2), good news EPS forecasts (columns 3 and 4), and bad news EPS forecasts (columns 5 and 6), respectively. In columns 1 and 2, the estimated coefficients on $HighLitRisk_{i,t}$ and $Released_{i,t} \times HighLitRisk_{i,t}$ are significantly negative. The results suggest that after satellite data release, treated firms significantly decrease their management forecasts and such decrease is more pronounced for firms with high litigation risk (measured relative to control firms). For instance, the estimated coefficient in column 2 suggests that the decrease for high-litigation risk firms is 14.6% more than that for low-litigation risk firms. Consistent with the prediction of our hypothesis, results in columns 3 to 6 show that the effect of litigation risk is mainly concentrated in good news forecasts. The estimated coefficients on $Released_{i,t} \times HighLitRisk_{i,t}$ are significantly negative for good news EPS forecasts in columns 3-4 but insignificant for bad news EPS forecasts in columns 5-6.

5.4.5 Poor Stock Returns and Bad News Forecasts

While our above analyses focus on several predictions that the reduction in good news forecasts is asymmetrically more pronounced under certain circumstances, we attend to the frequency of bad news and test a unique prediction from the meeting guidance hypothesis that management suppresses bad news forecasts only if the market already has very low earnings expectations. The intuition is that, if the market has already fully adjusted to, or overreacted to bad fundamental signals (i.e., the market is overly pessimistic about the firm's performance), the management is confident to meet the low bar and thus has no incentive to further walk down market expectations.

To test this prediction, we use extremely negative abnormal stock returns as a proxy for market's pessimism about firm performance. Specifically, we create a dummy variable $LowPastRet_{i,t}$, which equals one if firm *i* has extremely low abnormal past return (below the bottom quintile) over year *t*-1, and zero otherwise. Annual abnormal stock return is measured as the difference between cumulative monthly raw returns minus cumulative monthly DGTW (1997) benchmark portfolio returns over a year.[15] We choose the bottom quintile as a rough-cut, which is -43% in our sample. We believe that annual abnormal returns below this figure reasonably capture the market's pessimism.

We then set $X_{i,t}$ in the DID regression model of Equation (4) as $LowPastRet_{i,t}$, and report in Panel D of Table 7 the results for management forecasts (columns 1 and 2), good news EPS forecasts (columns 3 and 4), and bad news EPS forecasts (columns 5 and 6), respectively. In columns 1 and 2, the coefficients on $Released_{i,t} \times LowPastRet_{i,t}$ are significantly negative for management forecasts, indicating that firms with pessimistic performance expectation incrementally suppress their management forecasts. Notably, the results in columns 3 to 6 indicate that such incremental decrease

[15] Monthly DGTW (1997) benchmark portfolio consists of stocks in the same market capitalization, book-to-market, and prior-year return quintile following the methodology as in Daniel, Grinblatt, Titman, and Wermers (1997).

stems from bad news forecasts. The estimated coefficients on $Released_{i,t} \times LowPastRet_{i,t}$ for good news forecasts (in columns 3 and 4) are small and statistically insignificant, while the coefficients for bad news forecasts are large and significantly negative (in columns 5 and 6).

5.4.6 Precision of Management Forecasts

Lastly, we test the prediction that, as the management becomes more concerned with missing its own guidance, it may be more likely to issue less precise good news forecasts to avoid providing "smoking gun" evidence against its reputation. To test this prediction, we use *Precision* as the dependent variable in the DID regression model of Equation (1). Following previous literature (e.g., Baginski and Hassell 1997; Baginski, Hassell, and Kimbrough 2002), we assign each forecast a score ranging from 0 to 3, which gives the highest value to the most precise forecasts. We designate a score of 0 if the forecast is a qualitative forecast, 1 if the forecast is an open-ended range, 2 if the forecast is a closed range, and 3 if the forecast is a point estimate. Precision is defined as the average score of management forecasts made by a firm during a fiscal year (e.g., Houston et al. 2019).

Panel A of Table 8 reports the DID regression of management forecast precision on satellite data release. We report the results for management forecasts in columns 1 and 2 and for sales forecasts in columns 3 and 4, respectively. The DID coefficients on $Released_{i,t}$ are significantly negative for sales forecasts and are negative but insignificant for management forecasts. It is not surprising that the results are stronger for sales forecasts as satellite data on parking lot traffic provide more direct and relevant information for sales numbers of retail firms. In Panel B, when we further bifurcate management forecasts into good news EPS forecasts and bad news EPS forecasts, we find significant negative coefficients on $Released_{i,t}$ for good news forecasts in columns 1 and 2. The coefficients on $Released_{i,t}$ in columns 3 and 4 for bad news forecasts are close to zero and statistically insignificant.

Taken together, our results show that after satellite data is released and business activities become more observable, retail firms not only reduce the frequency of good news forecasts but also reduce the precision associated with good news forecasts. These results are consistent with our meeting guidance hypothesis.

6. Additional Tests

6.1 Does satellite data contain valuable information about retail firms' performance?

In this section, we provide evidence that high-resolution satellite imagery data of parking lot traffic provides timely information for retail firm's operating performance (Zhu 2019; Kang et al. 2021; Katona et al. 2021). Since satellite imagery data is released in near real time, it can be used to predict firm operating performance in the same quarter, which is usually announced to the public with a delay of a few weeks up to a couple of months after the quarter end. Specifically, we show that retail firms' quarterly traffic growth rate can predict firm-level year-over-year quarterly sales growth, net income growth, and cumulative abnormal return (*CAR*) around earnings announcements of the same quarter.

Following Kang et al. (2021), we first calculate store-level quarterly traffic growth as the average daily car count of a store in the current quarter relative to that in the same quarter of the previous year. We then calculate firm-level quarterly traffic growth as the weighted average of store-level quarterly traffic growth. The weight is the store's relative size within the firm, which is defined as the quarterly average car count of a store divided by the sum of the quarterly average car count of all stores within the firm. The mean and standard deviation of *traffic growth* for our sample firms are 30.5% and 47.0%, respectively.

Table 9 reports the regressions of a retail firm's performance measures on its traffic growth of the same quarter. We control for lagged sales growth, quarterly stock market returns, and other firm characteristics used in our baseline regression. In column 1, the dependent variable is quarterly sales growth (in %), which is defined as year-over-year percentage change of quarterly sales. The coefficient on *traffic growth* is 0.015, which is significant at the 5% level (*t*-statistic = 2.42). The estimate suggests that a one-standard-deviation increase in *traffic growth* is associated with a 0.7 percentage-point increase in its *sales growth*.[16] In column 2, the dependent variable is *net income growth* (in %), which is defined as year-over-year percentage change of quarterly net income. The coefficient on *traffic growth* is 0.169 (*t*-statistic = 2.16), indicating that a one-standard-deviation increase in *traffic growth* is associated with an 8.0 percentage-point increase in *net income growth*. In column 3, the dependent

[16] This figure is calculated as the coefficient 0.015×47% (the standard deviation of traffic growth) = 0.7%. Similarly, the economic significance in column 2 is calculated as the coefficient 0.169×47% = 8.0%, and in column 3 as 0.013×47% = 0.6%.

variable is *CAR[0,2],* which is defined as the cumulative abnormal return calculated using daily market-adjusted stock excess returns around the three-day [0,2] event window of quarterly earnings announcement. The coefficient on *traffic growth* is 0.013, significant at the 5% level (*t*-statistic = 2.54). The estimate indicates that a one-standard-deviation increase in *traffic growth* is associated with a 0.6 percentage-point increase in earnings announcement *CAR* over the three-day [0,2] event window. Overall, our results are consistent with recent findings that satellite imagery data of parking lot traffic provides valuable and timely information about firm performance.

6.2 Test for reverse causality: Can change in management forecasts predict the initiation of satellite data release?

To further address the endogeneity concern, in particular the reverse causality from voluntary disclosure to the release of satellite data (i.e., fewer forecasts induce the release of satellite traffic data), we examine whether the changes in management forecasts lead to the initiation of satellite data release. Following previous studies (e.g., Kim, Shroff, Vyas, and Wittenberg-Moerman 2018), we estimate the model below for the five-year period before and after the initiation year of satellite data release for the treated sample of retail firms:

$$ReleaseInitiation_{i,t} = \alpha_t + \beta \Delta Ln(\#MgmtFcst + 1)_{i,t-1} + \gamma \Delta \boldsymbol{z_{i,t}} + \epsilon_{it}, \quad (5)$$

where $ReleaseInitiation_{i,t}$ takes a value of one if year *t* is the initiation year of satellite data release for firm i, and zero otherwise. $\Delta Ln(\#MgmtFcst + 1)_{i,t-1}$ is the lagged change in the number of management forecasts relative to the previous year. We include changes of all control variables used in our previous main regression, where the change is measured relative to the previous year. Since the specification of first difference removes unobserved firm-specific fixed effects, we only include time fixed effects in the regression.

The results are reported in Table 10.[17] We present the results based on the Ordinary Least Squares (OLS) regression in columns 1 and 2 and results based on the Logit regression in columns 3 and 4. We

[17] The number of observations in the OLS regression (926) is smaller than that of the full sample (5,515) because in the test for reverse causality, we restrict the sample to the five-year period before and after the initiation year of satellite data release for treated firms following previous studies such as Kim et al. (2018). The number of observations in the logit regression further decreases to 733 because in the presence of year fixed effects, observations in years with no variations of the dependent variable are dropped out.

find that in all model specifications, the estimated coefficients on $\Delta Ln(\#MgmtFcst+1)_{i,t-1}$ are statistically and economically insignificant, confirming that changes in management disclosure frequency do not lead to the initiation of satellite data release. Our results thus alleviate the endogeneity concern of reverse causality.

6.3 Address potential biases in staggered DID regressions due to treatment effect heterogeneity

Recent studies suggest that in the presence of treatment effect heterogeneity, the two-way fixed effect (TWFE) DID estimators may be biased under staggered treatment timing (e.g., Callaway and Sant'Anna 2021; Baker, Larcker, and Wang 2022). It is shown that the estimates obtained through TWFE DID estimation are variance-weighted averages of many different "2×2" DIDs, each involving the comparison between a treated and an effective control group before and after the treated group receives the treatment. In some of the "2×2" DIDs, already-treated group can act as the effective control group, which may be problematic when treatment effects vary over time.

Callaway and Sant'Anna (2021) propose an alternative estimator, which allows for treatment effect heterogeneity and consistently estimates the average treatment effects of the treated (ATT). Baker et al. (2022) provide further evidence that the Callaway and Sant'Anna (CS) estimator is unbiased based on simulations using Compustat data. The CS approach first estimates the individual cohort-time-specific treatment effects through simple "2×2" DIDs with clean controls. The treatment effect of a particular treatment group treated at time *g* can be estimated in the following regression:

$$y_{it}=\alpha_1^{g,\tau}+\alpha_2^{g,\tau}\cdot\mathbb{I}\{E_i=g\}+\alpha_3^{g,\tau}\cdot\mathbb{I}\{t=\tau\}+\beta^{g,\tau}\cdot(\mathbb{I}\{E_i=g\}\times\mathbb{I}\{t=\tau\}),\qquad(6)$$

using observations at time τ and g-1 from treated units *i* with $\mathbb{I}\{E_i=g\}$ =1 and a set of clean controls, which are allowed to be not-yet-treated or never-treated units. $\mathbb{I}\{\cdot\}$ is an indicator variable that equals one when the condition in the braces is satisfied. E_i denotes the time when unit *i* receives treatment and $E_i = g$ for all firms that receive treatment at time *g*. The CS approach then aggregates the group-time ATT estimators ($\beta^{g,\tau}$) to produce measures of overall treatment effects. It is recommended to compute the average ATTs for each treatment cohort (across all post-treatment periods) and then report the weighted average ATTs across cohorts (e.g., weighted by each cohort's sample share).

We implement the CS estimator in our sample and report the results in Table 11. The dependent variable is *Ln(#MgmtFcst+1)* in columns 1-4. We estimate the average treatment effect using never-treated or not-yet-treated firms as clean controls in columns 1 and 2, respectively. It is evident that the CS regressions provide a negative and significant estimate of the average treatment effect. For example, the ATT is estimated to be -0.263 with a *t*-statistic of -4.27, the magnitude of which is very close to our DID coefficient reported in Table 3 (e.g., -0.286 in column 1). We further estimate the dynamic treatment effects using never-treated or not-yet-treated firms in columns 3 and 4, respectively. For brevity, we only report coefficients from *g*-3 to *g*+3. The CS estimates show that the treatment effect is insignificant before the treatment year (e.g., from *g*-3 to *g*-1), but become significantly negative after the treatment year (e.g. from *g*+0 to *g*+3). The results confirm our pre-trend tests reported in Panel A of Table 4 and plotted in Figure 2. We perform the same analysis for *Ln(#MgmtSalesFcst+1)* in columns 5-8 and the results are similar. Taken together, our analyses based on the CS estimator confirm that our results are robust for correcting the potential bias from treatment effect heterogeneity in the TWFE DID regressions with staggered treatment timing.

7. Conclusion

In this paper, we provide empirical evidence consistent with the interpretation that improved business observability (due to the release of high-resolution satellite data on parking lot traffic) causes retail firms to reduce the frequency of management forecasts, especially for those conveying good news. This asymmetric reduction cannot be explained by the general substitution between price informativeness and firm voluntary disclosure. Our evidence is more consistent with the management's incentive of meeting its own earnings guidance. In particular, the asymmetric reduction is stronger among short-term quarterly guidance and is more pronounced for firms with higher operating uncertainty, higher institutional ownership, and higher expected litigation risk. By contrast, the frequency of bad news forecasts decreases more when the market is overly pessimistic about firm performance.

To the extent that our results from retail firms can be generalized to other industries and other types of alternative data, the major implication of our results is that the increasing availability of alternative data mitigates managerial opportunism and agency problems associated with corporate disclosure decisions. Higher business observability squeezes management's discretion, and thus curbs the opportunistic earnings guidance game in particular. As product market competitors can also obtain alternative data, future research may offer additional insights regarding the implications of lower managerial optimism for the product markets.

References

Ajinkya, Bipin, Sanjeev Bhojraj, and Partha Sengupta, 2005, The association between outside directors, institutional investors and the properties of management earnings forecasts, *Journal of Accounting Research* 43, 343-376.

Ali, Ashiq, Sandy Klasa, and Eric Yeung, 2014, Industry concentration and corporate disclosure policy, *Journal of Accounting and Economics* 58, 240-264.

Ali, Ashiq, Ningzhong Li, and Weining Zhang, 2019, Restrictions on managers' outside employment opportunities and asymmetric disclosure of bad versus good news, *The Accounting Review* 94, 1-25.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, 2005, Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools, *Journal of Political Economy* 113, 151-184.

Anilowski, Carol, Mei Feng, and Douglas J. Skinner, 2007, Does earnings guidance affect market returns? The nature and information content of aggregate earnings guidance, *Journal of Accounting and Economics* 44, 36-63.

Baginski, Stephen P., and John M. Hassell, 1997, Determinants of management forecast precision, *The Accounting Review* 72, 303-312.

Baginski, Stephen P., John M. Hassell, and Michael D. Kimbrough, 2002, The effect of legal environment on voluntary disclosure: Evidence from management earnings forecasts issued in U.S. and Canadian markets, *The Accounting Review* 77, 25-50.

Baginski, Stephen P., and Kenneth C. Rakow, 2012, Management earnings forecast disclosure policy and the cost of equity capital, *Review of Accounting Studies* 17, 279-321.

Baker, Andrew C., David F. Larcker, and Charles CY Wang, 2022, How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics* 144, 370-395.

Balakrishnan, Karthik, Mary B. Billings, Bryan Kelly, and Alexander Ljungqvist, 2014, Shaping liquidity: On the causal effects of voluntary disclosure, *Journal of Finance* 69, 2237-2278.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, 2004, How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119, 249-275.

Bertrand, Marianne, and Sendhil Mullainathan, 2003, Enjoying the quiet life? Corporate governance and managerial preferences, *Journal of Political Economy* 111, 1043-1075.

Beyer, Anne, Daniel A. Cohen, Thomas Z. Lys, and Beverly R. Walther, 2010, The financial reporting environment: Review of the recent literature, *Journal of Accounting and Economics* 50, 296-343.

Blankespoor, Elizabeth, Gregory S. Miller, and Hal D. White, 2014, The role of dissemination in market liquidity: Evidence from firms' use of Twitter™, *The Accounting Review* 89, 79-112.

Boone, Audra L, and Joshua T. White, 2015, The effect of institutional ownership on firm transparency and information production, *Journal of Financial Economics* 117, 508-533.

Bourveau, Thomas, Yun Lou, and Rencheng Wang, 2018, Shareholder litigation and corporate disclosure: Evidence from derivative lawsuits, *Journal of Accounting Research* 56, 797-842.

Bourveau, Thomas, and Jordan Schoenfeld, 2017, Shareholder activism and voluntary disclosure, *Review of Accounting Studies* 22, 1307-1339.

Brown, Nerissa C., Han Stice, and Roger M. White, 2015, Mobile communication and local information flow: Evidence from distracted driving laws, *Journal of Accounting Research* 53, 275-329.

Callaway, Brantly, and Pedro HC Sant'Anna, 2021, Difference-in-differences with multiple time periods, *Journal of Econometrics* 225, 200-230.

Chen, Hailiang, Prabuddha De, Yu (Jeffrey) Hu, and Byoung-Hyoun Hwang, 2014, Wisdom of crowds: The value of stock opinions transmitted through social media, *Review of Financial Studies* 27, 1367-1403.

Chen, Qi, and Rahul Vashishtha, 2017, The effects of bank mergers on corporate information disclosure, *Journal of Accounting and Economics* 64, 56-77.

Dai, Lili, Jerry T. Parwada, and Bohui Zhang, 2015, The governance effect of the media's news dissemination role: Evidence from insider trading, *Journal of Accounting Research* 53, 331-366.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035-1058.

Darrough, Masako N., and Neal M. Stoughton, 1990, Financial disclosure policy in an entry game, *Journal of Accounting and Economics* 12, 219-243.

Diamond, Douglas W., and Robert E. Verrecchia, 1991, Disclosure, liquidity, and the cost of capital, *Journal of Finance* 46, 1325-1359.

Dye, Ronald A., 1985, Disclosure of nonproprietary information, *Journal of Accounting Research* 23, 123-145.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Foster, Andrew, Emilio Gutierrez, and Naresh Kumar, 2009, Voluntary compliance, pollution levels, and infant mortality in Mexico, *American Economic Review* 99, 191-197.

Foster, Andrew D., and Mark R. Rosenzweig, 2003, Economic growth and the rise of forests, *Quarterly Journal of Economics* 118, 601-637.

Froot, Kenneth, Namho Kang, Gideon Ozik, and Ronnie Sadka, 2017, What do measures of real-time corporate sales say about earnings surprises and post-announcement returns? *Journal of Financial Economics* 125, 143-162.

Goldstein, Itay, Chester S. Spatt, and Mao Ye, 2021, Big data in finance, *Review of Financial Studies* 34, 3213-3225.

Graham, John R., Campbell R. Harvey, and Shiva Rajgopal, 2005, The economic implications of corporate financial reporting, *Journal of Accounting and Economics* 40, 3-73.

Guiteras, Raymond, Amir Jina, and A. Mushflq Mobarak, 2015, Satellites, self-reports, and submersion: exposure to floods in Bangladesh, *American Economic Review* 105, 232-236.

Healy, Paul M., and Krishna G. Palepu, 2001, Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature, *Journal of Accounting and Economics* 31, 405-440.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil, 2012, Measuring economic growth from outer space, *American Economic Review* 102, 994-1028.

Holmes, Thomas J., and Sanghoon Lee, 2012, Economies of density versus natural advantage: Crop choice on the back forty, *Review of Economics and Statistics* 94, 1-19.

Houston, Joel F., Chen Lin, Sibo Liu, and Lai Wei, 2019, Litigation risk and voluntary disclosure: Evidence from legal changes, *The Accounting Review* 94, 247-272.

Jung, Michael J., 2013, Investor overlap and diffusion of disclosure practices, *Review of Accounting Studies* 18, 167-206.

Jung, Woon-Oh, and Young K. Kwon, 1988, Disclosure when the market is unsure of information endowment of managers, *Journal of Accounting Research* 26, 146-153.

Kang, Jung Koo, Lorien Stice-Lawrence, and Yu Ting Forester Wong, 2021, The firm next door: Using satellite images to study local information advantage, *Journal of Accounting Research* 59, 713-750.

Katona, Zsolt, Marcus Painter, Panos N. Patatoukas, and Jean Zeng, 2021, On the capital market consequences of alternative data: Evidence from outer space, Working paper.

Khurana, Inder K., Yinghua Li, and Wei Wang, 2018, The effects of hedge fund interventions on strategic firm behavior, *Management Science* 64, 4094-4117.

Kim, Irene, and Douglas J. Skinner, 2012, Measuring securities litigation risk, *Journal of Accounting and Economics* 53, 290-310.

Kim, Jae B., Pervin Shroff, Dushyantkumar Vyas, and Regina Wittenberg-Moerman, 2018, Credit default swaps and managers' voluntary disclosure, *Journal of Accounting Research*, 56, 953-988.

Kothari, Sabino P., Susan Shu, and Peter D. Wysocki, 2009, Do managers withhold bad news? *Journal of Accounting Research* 47, 241-276.

Langberg, Nisan, and K. Sivaramakrishnan, 2008, Voluntary disclosures and information production by analysts, *Journal of Accounting and Economics* 46, 78-100.

Lee, Lian Fen, Amy P. Hutton, and Susan Shu, 2015, The role of social media in the capital market: Evidence from consumer product recalls, *Journal of Accounting Research* 53, 367-404.

Leuz, Christian, and Peter D. Wysocki, 2016, The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research, *Journal of Accounting Research* 54, 525-622.

Lo, Alvis K., 2014, Do declines in bank health affect borrowers' voluntary disclosures? Evidence from international propagation of banking shocks, *Journal of Accounting Research* 52, 541-581.

Miller, Gregory S., and Douglas J. Skinner, 2015, The evolving disclosure landscape: How changes in technology, the media, and capital markets are affecting disclosure, *Journal of Accounting Research* 53, 221-239.

Minton, Bernadette A., and Catherine Schrand, 1999, The impact of cash flow volatility on discretionary investment and the costs of debt and equity financing, *Journal of Financial Economics* 54, 423-460.

Roberts, Michael R., and Toni M. Whited, 2013, Endogeneity in empirical corporate finance, *Handbook of the Economics of Finance* 2, 493-572.

Rogers, Jonathan L., and Andrew Van Buskirk, 2013, Bundled forecasts in empirical accounting research, *Journal of Accounting and Economics* 55, 43-65.

Teoh, Siew Hong, and Chuan Yang Hwang, 1991, Nondisclosure and adverse disclosure as signals of firm value, *Review of Financial Studies* 4, 283-313.

Vashishtha, Rahul, 2014, The role of bank monitoring in borrowers' discretionary disclosure: Evidence from covenant violations, *Journal of Accounting and Economics* 57, 176-195.

Verrecchia, Robert E., 1983, Discretionary disclosure, *Journal of Accounting and Economics* 5, 179-194.

Yagan, Danny, 2015, Capital tax reform and the real economy: The effects of the 2003 dividend tax cut, *American Economic Review* 105, 3531-3563.

Zhu, Christina, 2019, Big data as a governance mechanism, *Review of Financial Studies* 32, 2021-2061.

Appendix A. Variable Definitions

Variable	Definition
Disclosure measures	
#MgmtFcst	The total number of management forecasts issued over a fiscal year.
Ln(#MgmtFcst+1)	The natural logarithm of the total number of management forecasts issued over a fiscal year plus one.
Ln(#MgmtSalesFcst+1)	The natural logarithm of the total number of management sales forecasts issued over a fiscal year plus one.
Ln(#GoodNewsFcst+1)	The natural logarithm of the number of good news EPS forecasts plus one. A management EPS forecast is classified as good news if the point estimate, or the midpoint of the range forecast, is above the analyst consensus forecast before the management forecast. For open-ended management forecasts, the forecast is classified as good when its lower bound is higher than the analyst consensus forecast. Following Rogers and Van Buskirk (2013), we correct the potential measurement errors of forecast news when management forecasts are bundled with earnings announcements by updating analyst expectations with the latest information in earnings announcements. We then classify management forecast news based on the updated analyst forecast consensus.
Ln(#BadNewstFcst+1)	The natural logarithm of the number of bad news EPS forecasts plus one. A management EPS forecast is classified as bad news if the point estimate, or the midpoint of the range forecast, is below the analyst consensus forecast before the management forecast. For open-ended management forecasts, the forecast is classified as bad when its upper bound is lower than the analyst consensus forecast.
Precision	Each forecast is assigned a score ranging from 0 to 3, which gives the highest value to the most precise forecasts. We designate a score of 0 if the forecast is a qualitative forecast, 1 if the forecast is an open-ended range, 2 if the forecast is a closed range, and 3 if the forecast is a point estimate. Precision is defined as the average score of management forecasts made by a firm during a fiscal year.
Firm-level control variables	
RetVol	Stock return volatility, calculated as the standard deviation of monthly stock returns over a firm's fiscal year, with a requirement of at least six months of observations.
AbsChEPS	The absolute value of the annual change in earnings per share divided by stock price at the beginning of the fiscal year.
MktAdjRet	Market-adjusted stock return, defined as the firm's buy-and-hold 12-month fiscal year stock return minus the CRSP value-weighted stock return over the same period.
R&D	R&D expense divided by the book value of assets at the beginning of the fiscal year.
Size	Firm size, defined as the natural logarithm of the market value of equity.
Coverage	Analyst coverage, defined as the average number of analysts making earnings forecasts for a firm over a fiscal year.
InstOwn	Institutional fractional ownership, collected from the Thomson Reuters CDA/Spectrum Institutional(13f) Holdings database and represents institutional holdings at the beginning of the fiscal year.

Continued next page

D_Issuance	Equity or debt issuance dummy, which takes a value of one if a firm issues public equity or public debt in a subsequent two-year period, and zero otherwise.
MB	Market-to-book equity ratio, defined as the market value of equity divided by the book value of equity.
Lev	Leverage, defined as total liabilities minus deferred taxes scaled by total book assets.
StdEarn	Standard deviation of earnings, defined as the standard deviation of earnings before extraordinary items over the prior five years, with a requirement of at least three years of observations.
D_PosChEarn	Positive earnings change dummy, which takes the value of one if the annual earnings per share divided by stock price of the current year is higher than that of the previous year, and zero otherwise.
Optimism	Analyst optimism, defined as the difference between analyst forecast consensus at the beginning of the fiscal year and the actual earnings per share, scaled by the absolute value of actual earnings per share.
Conditioning variables	
HighCFVol	A dummy variable that equals one if firm *i* has high cash flow volatility (above the median) at the end of year *t*-1, and zero otherwise. Cash flow volatility is the coefficient of variation in a firm's quarterly operating cash flow over the past six-year period following Minton and Schrand (1999).
HighInstOwn	A dummy variable that equals one if firm *i* has high institutional ownership (above the median) at the end of year *t*-1, and zero otherwise.
HighLitRisk	A dummy variable that equals one if firm *i* has high expected litigation risk (above the median) at the end of year *t*-1, and zero otherwise. Firm-level expected litigation risk is calculated as the probability of litigation based on the coefficient estimates reported in model 3 of Table 7 by Kim and Skinner (2012).
LowPastRet	A dummy variable that equals one if firm *i* has extremely low abnormal past return (below the bottom quintile) over year *t*-1, and zero otherwise. Annual abnormal stock return is measured as the difference between cumulative monthly raw returns minus cumulative monthly DGTW (1997) benchmark portfolio returns over a year.

Figure 1. Number of Satellite Data Release Events Across Years

This figure reports the number of release events across years. The sample includes U.S. retail firms whose satellite data of parking lot traffic are released from 2011 to 2018.

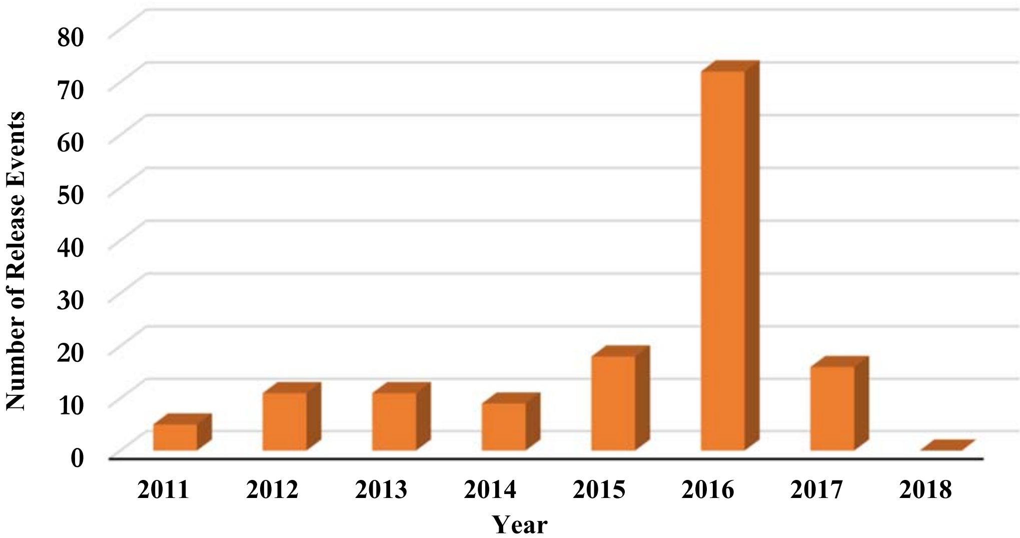

Figure 2. Parallel Trend Analysis: Dynamic Treatment Effect on Management Forecasts
This figure presents the difference-in-differences coefficients prior to and after the event year of satellite data release (t=0) in the regression of Equation (2) and their 95% confidence intervals. Event years of t < -3 are used as the benchmark.

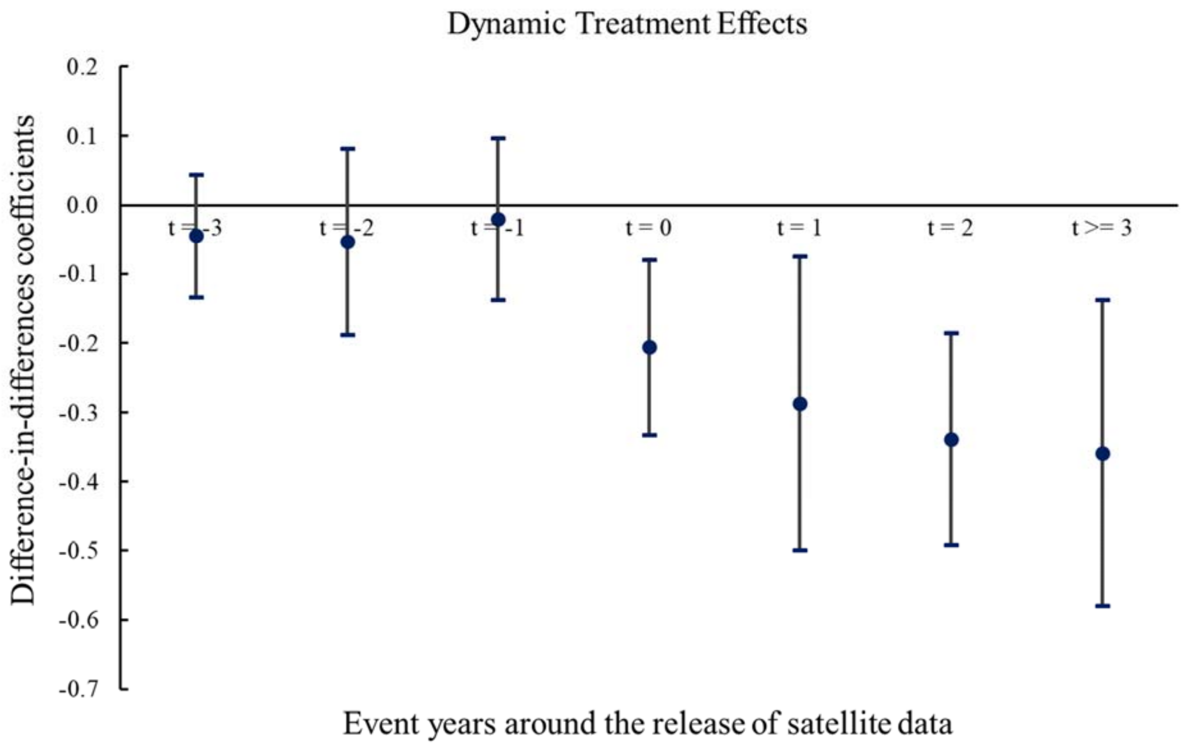

Table 1. Industry Distribution of the Treated Sample

This table reports the GICS industry distribution of treated firms before and after merging the sample with management forecast data. The treated sample includes U.S. retail firms with parking lot traffic data eventually released by either of the two commercial satellite data vendors: RS metrics and Orbital Insight. Panel A presents the sample of 142 treated firms before merging with management forecasts data. Panel B presents the sample of 117 treated firms after merging with management forecasts data. Both the number and percentage are reported.

Panel A. Before merging with management forecasts

GICS Industries	Number	Percentage
Chemicals	1	0.7%
Household Durables	1	0.7%
Textiles, Apparel, and luxury goods	7	4.9%
Hotels, Restaurants, and Leisure	28	19.7%
Media	1	0.7%
Distributors	3	2.1%
Multiline Retail	14	9.9%
Specialty Retail	67	47.2%
Food and Staples Retailing	14	9.9%
Food Products	2	1.4%
Health Care Providers and Services	1	0.7%
Consumer Finance	1	0.7%
Entertainment	2	1.4%
Total	142	100.0%

Panel B. After merging with management forecasts

GICS Industries	Number	Percentage
Chemicals	1	0.9%
Household Durables	1	0.9%
Textiles, Apparel, and luxury goods	7	6.0%
Hotels, Restaurants, and Leisure	23	19.7%
Media	1	0.9%
Distributors	3	2.6%
Multiline Retail	13	11.1%
Specialty Retail	53	45.3%
Food and Staples Retailing	11	9.4%
Food Products	2	1.7%
Health Care Providers and Services	1	0.9%
Consumer Finance	1	0.9%
Total	117	100.0%

Table 2. Summary Statistics

This table reports the mean, standard deviation, 25th percentile, median, and 75th percentile of firm characteristics for the full sample (Panel A), treated sample (Panel B), and control sample (Panel C), respectively. The treated sample includes U.S. retail firms with parking lot traffic data eventually released by either of the two commercial satellite data vendors: RS metrics and Orbital Insight. The control sample includes other U.S. retail firms in the same GICS industry of treated firms and their satellite data has not been released by the end of 2018. The full sample combines treated and control samples. Firm characteristics include the total number of management forecasts (*#MgmtFcst*), the natural logarithm of the total number of management forecasts plus one (*Ln*(*#MgmtFcst+1*)), the natural logarithm of the number of management sales forecasts plus one (*Ln*(*#MgmtSalesFcst+1*)), the natural logarithm of the number of good news EPS forecasts plus one (*Ln*(*#GoodNewsFcst+1*)), the natural logarithm of the number of bad news EPS forecasts plus one (*Ln*(*#BadNewsFcst+1*)), stock return volatility (*RetVol*), absolute change in annual earnings per share scaled by stock price (*AbsChEPS*), market-adjusted stock return (*MktAdjRet*), research and development expense scaled by book assets (*R&D*), the natural logarithm of the market value of equity (*Size*), analyst coverage (*Coverage*), institutional ownership (*InstOwn*), equity or debt issuance dummy (*D_Issuance*), market-to-book equity ratio (*MB*), leverage (*Lev*), standard deviation of earnings (*StdEarn*), positive earnings change dummy (*D_PosChEarn*), and analyst optimism (*Optimism*). Detailed variable definitions are reported in Appendix A. The sample period is from 2008 to 2018. All variables are winsorized at the 1% and 99% levels.

Panel A. Full sample (N = 5,515)

Variable	Mean	Std.	P25	Median	P75
#MgmtFcst	9.746	10.199	1.000	7.000	16.000
Ln(#MgmtFcst+1)	1.776	1.227	0.693	2.079	2.833
Ln(#MgmtSalesFcst+1)	0.717	0.882	0.000	0.000	1.609
Ln(#GoodNewsFcst+1)	0.524	0.711	0.000	0.000	1.099
Ln(#BadNewsFcst+1)	0.285	0.487	0.000	0.000	0.693
RetVol	0.122	0.076	0.071	0.100	0.150
AbsChEPS	0.131	0.318	0.012	0.030	0.090
MktAdjRet	0.130	0.520	-0.173	0.076	0.343
R&D	0.005	0.016	0.000	0.000	0.000
Size	6.812	2.121	5.335	6.925	8.281
Coverage	9.347	7.033	3.917	7.500	13.583
InstOwn	0.673	0.286	0.501	0.761	0.900
D_Issuance	0.253	0.435	0.000	0.000	1.000
MB	2.299	4.154	1.032	1.792	3.062
Lev	0.554	0.253	0.370	0.533	0.702
StdEarn	0.059	0.071	0.016	0.031	0.072
D_PosChEarn	0.462	0.499	0.000	0.000	1.000
Optimism	0.045	0.356	-0.027	-0.005	0.020

Continued next page

Table 2 Continued

Panel B. Treated sample (N = 1,393)

Variable	Mean	Std.	P25	Median	P75
#MgmtFcst	13.200	10.313	5.000	12.000	20.000
Ln(#MgmtFcst+1)	2.281	1.007	1.792	2.565	3.045
Ln(#MgmtSalesFcst+1)	0.935	0.917	0.000	0.693	1.792
Ln(#GoodNewsFcst+1)	0.705	0.755	0.000	0.693	1.386
Ln(#BadNewsFcst+1)	0.435	0.589	0.000	0.000	0.693
RetVol	0.112	0.067	0.067	0.094	0.133
AbsChEPS	0.078	0.209	0.009	0.021	0.052
MktAdjRet	0.134	0.477	-0.142	0.076	0.340
R&D	0.000	0.004	0.000	0.000	0.000
Size	7.555	1.804	6.318	7.541	8.923
Coverage	13.306	8.506	6.167	11.917	19.917
InstOwn	0.754	0.241	0.661	0.830	0.927
D_Issuance	0.210	0.408	0.000	0.000	0.000
MB	2.534	5.056	1.239	2.102	3.475
Lev	0.553	0.256	0.364	0.513	0.689
StdEarn	0.043	0.052	0.014	0.026	0.051
D_PosChEarn	0.472	0.499	0.000	0.000	1.000
Optimism	0.012	0.280	-0.031	-0.008	0.007

Panel C. Control sample (N = 4,122)

Variable	Mean	Std.	P25	Median	P75
#MgmtFcst	8.579	9.892	0.000	5.000	14.000
Ln(#MgmtFcst+1)	1.605	1.247	0.000	1.792	2.708
Ln(#MgmtSalesFcst+1)	0.643	0.858	0.000	0.000	1.609
Ln(#GoodNewsFcst+1)	0.463	0.684	0.000	0.000	1.099
Ln(#BadNewsFcst+1)	0.234	0.436	0.000	0.000	0.693
RetVol	0.125	0.078	0.073	0.103	0.155
AbsChEPS	0.149	0.345	0.014	0.037	0.109
MktAdjRet	0.128	0.533	-0.185	0.076	0.343
R&D	0.007	0.018	0.000	0.000	0.004
Size	6.561	2.161	4.980	6.653	8.073
Coverage	8.010	5.887	3.417	6.917	10.917
InstOwn	0.646	0.294	0.442	0.728	0.887
D_Issuance	0.268	0.443	0.000	0.000	1.000
MB	2.219	3.799	0.975	1.687	2.902
Lev	0.554	0.252	0.373	0.539	0.705
StdEarn	0.064	0.076	0.017	0.034	0.082
D_PosChEarn	0.458	0.498	0.000	0.000	1.000
Optimism	0.056	0.377	-0.026	-0.005	0.028

Table 3. Frequency of Management Forecasts and Satellite Data Release

This table reports the estimation results of the following difference-in-differences regression:

$$Disclosure_{i,t} = \alpha_i + \alpha_t + \beta_1 Released_{i,t} + \gamma \boldsymbol{z_{i,t}} + \varepsilon_{i,t}.$$

The dependent variable is *Ln(#MgmtFcst+1)* in columns 1-2 and *Ln(#MgmtSalesFcst+1)* in columns 3-4. $Released_{i,t}$ is a dummy variable that equals one if the satellite data on parking lot traffic has been released for firm i by year *t* and zero otherwise. Control variables $\boldsymbol{z_{i,t}}$ include: stock return volatility, absolute change in annual earnings per share scaled by stock price, market-adjusted stock return, research and development expense scaled by book assets, natural logarithm of market value of equity, analyst coverage, institutional ownership, equity or debt issuance dummy, market-to-book equity ratio, leverage, standard deviation of earnings, positive earnings change dummy, and analyst optimism. Detailed variable definitions are reported in Appendix A. All regressions include firm and year fixed effects. The sample period is from 2008 to 2018. The *t*-statistics based on two-way clustered standard errors at firm and year levels are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively. Coefficients on main variables of interest are in bold.

	Ln(#MgmtFcst+1)		*Ln(#MgmtSalesFcst+1)*	
	(1)	(2)	(3)	(4)
Released	**-0.286*****	**-0.261*****	**-0.293*****	**-0.283*****
	(-4.24)	**(-4.23)**	**(-5.02)**	**(-4.79)**
RetVol		-0.537**		-0.306**
		(-2.44)		(-2.00)
AbsChEPS		-0.052		-0.053
		(-0.97)		(-1.32)
MktAdjRet		-0.009		-0.000
		(-0.37)		(-0.00)
R&D		0.090		0.736
		(0.02)		(0.21)
Size		0.148***		0.099***
		(4.81)		(3.96)
Coverage		-0.000		0.004
		(-0.09)		(0.88)
InstOwn		0.091		0.070
		(1.38)		(1.19)
D_Issuance		0.039		0.043**
		(1.24)		(2.23)
MB		-0.003		-0.003
		(-0.89)		(-0.84)
Lev		0.146		0.064
		(1.17)		(0.59)
StdEarn		-0.378		0.143
		(-1.08)		(0.59)
D_PosChEarn		-0.016		0.007
		(-0.81)		(0.40)
Optimism		-0.096***		-0.032*
		(-4.38)		(-1.68)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
# Obs	5,515	5,515	5,515	5,515
Adj. R^2	0.784	0.790	0.707	0.712

Table 4. Assessing Identification

This table reports tests of the identification strategy. Panel A presents the dynamic effects of satellite data release on management forecast frequency and tests for pre-trend differences between the treated and control firms. We replace $Released_{i,t}$ in the baseline regression with three dummy variables: $Released_{i,-2\le t\le -1}$ is a dummy variable that equals one if year *t* is within two years before the satellite data of firm *i* is released, and zero otherwise; $Released_{i,0\le t\le 1}$ is a dummy variable that equals one if year *t* is in the year or one year after the satellite data of firm *i* is released, and zero otherwise; $Released_{i,2\le t}$ is a dummy variable that equals one if year *t* is two years or more after the satellite data of firm *i* is released, and zero otherwise. In Panel B, we add $StartNotRelease_{i,t}$, which is a dummy variable that equals one if the satellite data on parking lot traffic has started but has not been released for firm *i* by year *t*, and zero otherwise. Control variables are the same as those in the baseline regression. Panel C presents the baseline regressions in the PSM sample. We implement the PSM procedure by first estimating a logit regression to model the probability that the satellite data of a retail firm is released based on all control variables in the baseline regression. We then match each treated firm to a control firm in the same industry using the nearest neighbor matching technique with no replacement. Detailed variable definitions are reported in Appendix A. All regressions include firm and year fixed effects. The sample period is from 2008 to 2018. The *t*-statistics based on two-way clustered standard errors at firm and year levels are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively. Coefficients on main variables of interest are in bold.

Panel A. Pre-trend test

	Ln(#MgmtFcst+1)		*Ln(#MgmtSalesFcst+1)*	
	(1)	(2)	(3)	(4)
$\boldsymbol{Released_{i,-2\le t\le -1}}$	**-0.017**	**-0.025**	**-0.048**	**-0.055**
	(-0.31)	**(-0.50)**	**(-0.98)**	**(-1.10)**
$\boldsymbol{Released_{i,0\le t\le 1}}$	**-0.238*****	**-0.233*****	**-0.267*****	**-0.269*****
	(-3.87)	**(-3.74)**	**(-3.83)**	**(-3.72)**
$\boldsymbol{Released_{i,t\ge 2}}$	**-0.379*****	**-0.333*****	**-0.384*****	**-0.363*****
	(-5.02)	**(-4.56)**	**(-4.71)**	**(-4.20)**
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
# Obs	5,515	5,515	5,515	5,515
Adj. R^2	0.784	0.790	0.708	0.712

Panel B: Start or release of the data

	Ln(#MgmtFcst+1)		*Ln(#MgmtSalesFcst+1)*	
	(1)	(2)	(3)	(4)
Released	**-0.255*****	**-0.252*****	**-0.278*****	**-0.287*****
	(-4.07)	**(-4.33)**	**(-4.19)**	**(-4.29)**
StartNotRelease	**0.039**	**0.011**	**0.020**	**-0.005**
	(0.94)	**(0.28)**	**(0.38)**	**(-0.09)**
Controls	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
# Obs	5,515	5,515	5,515	5,515
Adj. R^2	0.784	0.790	0.707	0.712

Continued next page

Table 4 Continued

Panel C. PSM sample

	Ln(#MgmtFcst+1)		***Ln(#MgmtSalesFcst+1)***	
	(1)	(2)	(3)	(4)
Released	**-0.296*****	**-0.256*****	**-0.311*****	**-0.284*****
	(-4.17)	**(-4.02)**	**(-5.04)**	**(-4.60)**
Controls	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
# Obs	2,409	2,409	2,409	2,409
Adj. R^2	0.777	0.788	0.726	0.733

Table 5. Management EPS Forecast Frequency of Good News and Bad News
This table presents difference-in-differences regression of management forecasts conveying different news. The dependent variables are *Ln(#GoodNewsFcst+1),* the natural logarithm of the number of good news EPS forecasts plus one, in columns 1-2 and *Ln(#BadNewsFcst+1),* the natural logarithm of the number of bad news EPS forecasts plus one, in columns 3-4. A management EPS forecast is classified as good (bad) news if the point estimate, or the midpoint of the range forecast, is above (below) the analyst consensus forecast before the management forecast. For open-ended management forecasts, the forecast is classified as good (bad) news when its bottom (upper) bound is higher (lower) than the analyst consensus forecast. Following Rogers and Van Buskirk (2013), we correct the potential measurement errors of forecast news when management forecasts are bundled with earnings announcements by updating analyst expectations with the latest information in earnings announcements. We then classify management forecast news based on the updated analyst forecast consensus. Detailed variable definitions are reported in Appendix A. All regressions include firm and year fixed effects. The sample period is from 2008 to 2018. The *t*-statistics based on two-way clustered standard errors at firm and year levels are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively. Coefficients on main variables of interest are in bold.

	Good News Forecasts		**Bad News Forecasts**	
	(1)	(2)	(3)	(4)
Released	**-0.208*****	**-0.187*****	**-0.066**	**-0.059**
	(-3.42)	**(-3.26)**	**(-1.29)**	**(-1.32)**
RetVol		-0.408***		-0.247*
		(-2.71)		(-1.68)
AbsChEPS		-0.034		0.047**
		(-1.30)		(1.98)
MktAdjRet		0.041**		-0.049***
		(2.57)		(-4.32)
R&D		1.302		-0.469
		(0.54)		(-0.24)
Size		0.035		0.132***
		(1.42)		(6.07)
Coverage		-0.005		0.009***
		(-1.08)		(2.60)
InstOwn		0.091		0.025
		(1.32)		(0.52)
D_Issuance		-0.003		0.030
		(-0.15)		(1.46)
MB		-0.002		-0.002
		(-0.72)		(-1.10)
Lev		-0.083		0.022
		(-1.22)		(0.24)
StdEarn		0.074		-0.078
		(0.33)		(-0.30)
D_PosChEarn		0.068***		-0.063***
		(4.96)		(-11.04)
Optimism		-0.039*		-0.010
		(-1.93)		(-0.70)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
# Obs	5,515	5,515	5,515	5,515
Adj. R^2	0.547	0.553	0.686	0.695

Table 6. Management Forecast Frequency: Quarterly Guidance vs. Long-term Forecasts
This table presents difference-in-differences regression of quarterly guidance vs. long-term management forecasts. Quarterly guidance refers to forecasts with horizon less than or equal to 90 days. Long-term forecasts are defined as forecasts with horizon greater than 90 days. The dependent variables are the natural logarithm of the number of quarterly guidance plus one in columns 1-2, the natural logarithm of the number of long-term forecasts in columns 3-4, the natural logarithm of the number of good news EPS quarterly guidance plus one in columns 5-6, and the natural logarithm of the number of good news EPS long-term forecasts in columns 7-8. All regressions include firm and year fixed effects. The sample period is from 2008 to 2018. The *t*-statistics based on two-way clustered standard errors at firm and year levels are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively. Coefficients on main variables of interest are in bold.

	Quarterly Guidance		**Long-term Forecast**		**Good News Quarterly Guidance**		**Good News Long-term Forecast**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Released	**-0.275*****	**-0.256*****	**-0.194*****	**-0.174*****	**-0.176*****	**-0.160*****	**-0.096****	**-0.084***
	(-4.56)	**(-4.32)**	**(-3.51)**	**(-3.44)**	**(-3.19)**	**(-3.11)**	**(-2.06)**	**(-1.79)**
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# Obs	5,515	5,515	5,515	5,515	5,515	5,515	5,515	5,515
Adj. R^2	0.737	0.742	0.758	0.765	0.430	0.438	0.444	0.447

Table 7. Cross Sectional Analysis: The Effect of Operating Uncertainty, Institutional Ownership, Expected Litigation Risk and Extremely Negative Abnormal Stock Returns

This table reports the following difference-in-differences regressions of management forecast frequency:

$$Ln(\#MgmtFcst + 1)_{i,t} = \alpha_i + \alpha_t + \beta_1 Released_{i,t} + \beta_2 Released_{i,t} \times X_{it} + \beta_3 X_{it} + \gamma \boldsymbol{z_{i,t}} + \varepsilon_{i,t},$$

where X is a dummy variable that equals *HighCFVol*, *HighInstOwn*, *HighLitRisk*, and *LowPastRet* for Panels A-D, respectively. *HighCFVol* is a dummy variable that equals one if firm *i* has high cash flow volatility (above the median) at the end of year *t*-1, and zero otherwise. Cash flow volatility is defined as the coefficient of variation in a firm's quarterly operating cash flow over the past six-year period. *HighInstOwn* is a dummy variable that equals one if firm *i* has high institutional ownership (above the median) at the end of year *t*-1, and zero otherwise. *HighLitRisk* is a dummy variable that equals one if firm *i* has high expected litigation risk (above the median) at the end of year *t*-1, and zero otherwise. Firm-level expected litigation risk is calculated based on the coefficient estimates reported in Kim and Skinner (2012, Model 3 of Table 7), which predicts the probability of litigation based on a list of firm characteristics. *LowPastRet* is a dummy variable that equals one if firm *i* has extremely low abnormal stock return (below the bottom quintile) over year *t*-1, and zero otherwise. The dependent variable is the natural logarithm of the number of management forecasts plus one for all forecasts (columns 1-2), good news EPS forecasts (columns 3-4), and bad news EPS forecasts (columns 5-6), respectively. All regressions include firm and year fixed effects. The sample period is from 2008 to 2018. The *t*-statistics based on two-way clustered standard errors at firm and year levels are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively. Coefficients on main variables of interest are in bold.

Panel A. Operating uncertainty

	Management Forecasts		Good News Forecasts		Bad News Forecasts	
	(1)	(2)	(3)	(4)	(5)	(6)
Released	-0.218***	-0.205***	-0.139**	-0.119**	-0.045	-0.054
	(-3.41)	(-3.43)	(-2.20)	(-2.04)	(-0.86)	(-1.09)
Released × HighCFVol	**-0.169****	**-0.139***	**-0.176****	**-0.172****	**-0.050**	**-0.012**
	(-2.12)	**(-1.91)**	**(-2.24)**	**(-2.16)**	**(-0.64)**	**(-0.15)**
HighCFVol	-0.063	-0.028	-0.005	0.004	-0.065**	-0.034
	(-1.37)	(-0.62)	(-0.23)	(0.19)	(-2.09)	(-1.13)
Controls	No	Yes	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
# Obs	5,515	5,515	5,515	5,515	5,515	5,515
Adj. R^2	0.784	0.790	0.548	0.554	0.687	0.696

Continued next page

Table 7 Continued

Panel B. Institutional ownership

	Management Forecasts		Good News Forecasts		Bad News Forecasts	
	(1)	(2)	(3)	(4)	(5)	(6)
Released	-0.307**	-0.269**	-0.106	-0.079	-0.035	-0.008
	(-2.49)	(-2.36)	(-1.62)	(-1.28)	(-0.63)	(-0.16)
Released × HighInstOwn	**0.029**	**0.011**	**-0.144****	**-0.152****	**-0.044**	**-0.072**
	(0.27)	**(0.10)**	**(-2.14)**	**(-2.32)**	**(-0.61)**	**(-1.02)**
HighInstOwn	0.095***	0.054	0.053	0.030	0.096***	0.083**
	(3.22)	(1.11)	(1.54)	(0.75)	(3.03)	(2.45)
Controls	No	Yes	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
# Obs	5,515	5,515	5,515	5,515	5,515	5,515
Adj. R^2	0.784	0.790	0.548	0.553	0.687	0.696

Panel C. Expected litigation risk

	Management Forecasts		Good News Forecasts		Bad News Forecasts	
	(1)	(2)	(3)	(4)	(5)	(6)
Released	-0.182**	-0.181**	-0.115	-0.100	-0.015	-0.027
	(-2.21)	(-2.38)	(-1.54)	(-1.38)	(-0.25)	(-0.50)
Released ×HighLitRisk	**-0.190****	**-0.146****	**-0.171****	**-0.161****	**-0.091**	**-0.056**
	(-2.38)	**(-2.00)**	**(-2.38)**	**(-2.15)**	**(-1.15)**	**(-0.74)**
HighLitRisk	-0.005	-0.004	0.016	0.013	-0.035	-0.030
	(-0.21)	(-0.18)	(1.38)	(1.40)	(-1.55)	(-1.29)
Controls	No	Yes	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
# Obs	5,515	5,515	5,515	5,515	5,515	5,515
Adj. R^2	0.784	0.790	0.548	0.554	0.687	0.696

Panel D. Extremely negative abnormal stock returns

	Management Forecasts		Good News Forecasts		Bad News Forecasts	
	(1)	(2)	(3)	(4)	(5)	(6)
Released	-0.225***	-0.213***	-0.199***	-0.181***	-0.030	-0.036
	(-3.46)	(-3.47)	(-3.44)	(-3.30)	(-0.68)	(-0.89)
Released ×LowPastRet	**-0.255*****	**-0.215*****	**-0.034**	**-0.029**	**-0.151*****	**-0.104****
	(-3.75)	**(-2.83)**	**(-0.45)**	**(-0.41)**	**(-3.05)**	**(-2.24)**
LowPastRet	-0.044**	0.014	-0.030*	0.013	-0.025	-0.013
	(-2.42)	(0.70)	(-1.94)	(0.88)	(-1.36)	(-0.69)
Controls	No	Yes	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
# Obs	5,515	5,515	5,515	5,515	5,515	5,515
Adj. R^2	0.785	0.790	0.547	0.553	0.687	0.696

Table 8. Precision of Management Forecasts

This table reports the difference-in-differences regression of management forecast precision. The dependent variable is the precision of management forecasts. Each forecast is assigned a score ranging from 0 to 3, which gives the highest value to the most precise forecasts. We designate a score of 0 if the forecast is a qualitative forecast, 1 if the forecast is an open-ended range, 2 if the forecast is a closed range, and 3 if the forecast is a point estimate. Precision is defined as the average score of management forecasts made by a firm during a fiscal year. Results for management forecasts and management sales forecasts are reported in columns 1-2 and 3-4 of Panel A, respectively. Results for good news and bad news EPS forecasts are reported in columns 1-2 and 3-4 of Panel B, respectively. Detailed variable definitions are reported in Appendix A. All regressions include firm and year fixed effects. The sample period is from 2008 to 2018. The *t*-statistics based on two-way clustered standard errors at firm and year levels are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively. Coefficients on main variables of interest are in bold.

Panel A. Management forecasts and sales forecasts

	Management Forecasts		Sales Forecasts	
	(1)	(2)	(3)	(4)
Released	**-0.023**	**-0.020**	**-0.131*****	**-0.118****
	(-0.97)	**(-0.81)**	**(-2.63)**	**(-2.10)**
RetVol		0.003		-0.224
		(0.04)		(-1.09)
AbsChEPS		-0.006		0.109***
		(-0.18)		(2.92)
MktAdjRet		-0.016**		-0.009
		(-2.21)		(-0.39)
R&D		0.762		2.693
		(0.45)		(1.06)
Size		0.019		0.054*
		(1.10)		(1.82)
Coverage		0.000		-0.001
		(0.14)		(-0.25)
InstOwn		0.007		-0.051
		(0.19)		(-1.05)
D_Issuance		-0.004		-0.030
		(-0.35)		(-1.07)
MB		-0.000		0.001
		(-0.06)		(0.44)
Lev		-0.034		0.010
		(-0.64)		(0.09)
StdEarn		0.263		0.208
		(1.44)		(1.14)
D_PosChEarn		-0.002		-0.010
		(-0.15)		(-0.42)
Optimism		-0.004		0.011
		(-0.21)		(0.25)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
# Obs	4,110	4,110	2,339	2,339
Adj. R^2	0.444	0.444	0.422	0.423

Continued next page

Table 8 Continued

Panel B. Good news and bad news EPS forecasts

	Good News Forecasts		Bad News Forecasts	
	(1)	(2)	(3)	(4)
Released	**-0.040****	**-0.037****	**-0.002**	**-0.005**
	(-2.36)	**(-1.98)**	**(-0.08)**	**(-0.18)**
RetVol		-0.234		0.017
		(-1.57)		(0.12)
AbsChEPS		-0.078*		-0.052
		(-1.88)		(-1.12)
MktAdjRet		0.012		0.025
		(0.66)		(1.09)
R&D		0.053		0.730
		(0.02)		(1.06)
Size		-0.010		-0.008
		(-0.51)		(-0.52)
Coverage		0.002		0.002
		(0.97)		(0.75)
InstOwn		0.012		0.034
		(0.22)		(0.49)
D_Issuance		-0.008		-0.009
		(-0.30)		(-0.39)
MB		-0.002		0.000
		(-0.76)		(0.10)
Lev		0.027		0.071
		(0.33)		(1.63)
StdEarn		-0.134		-0.184
		(-0.61)		(-0.91)
D_PosChEarn		0.005		0.003
		(0.43)		(0.15)
Optimism		-0.081***		-0.001
		(-3.31)		(-0.03)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
# Obs	2,147	2,147	2,258	2,258
Adj. R^2	0.308	0.309	0.396	0.395

Table 9. Does Satellite Data Contain Valuable Information for Firm Performance?

This table reports the regressions of quarterly sales growth (in %), net income growth (in %), and market-adjusted cumulative abnormal returns around the three-day [0,2] event window of quarterly earnings announcement (in %) on traffic growth of the same quarter. Sales growth or net income growth is defined as year-over-year percentage change of quarterly sales or net income (in %). Traffic growth is defined as the weighted average of quarterly store-level percentage change in car count for each retail firm (in %). The percentage change in car count for each store is calculated as the average daily car count of a store in the current quarter relative to that in the same quarter of the previous year. The weight is the store's relative size within the firm, which is defined as the quarterly average car count of a store divided by the sum of the quarterly average car count of all stores within the firm. Control variables include lagged sales growth (*Lag Sales Growth)*, quarterly stock market returns (*Qret*), and other firm characteristics used in the baseline regression. The *t*-statistics based on standard errors clustered at both the firm and year levels are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively. Coefficients on main variables of interest are in bold.

	Sales Growth	Net Income Growth	CAR[0,2]
	(1)	(2)	(3)
Traffic Growth	**0.015****	**0.169****	**0.013****
	(2.42)	**(2.16)**	**(2.54)**
Lag Sales Growth	0.580***	1.004***	-0.025
	(11.17)	(5.38)	(-1.07)
Qret	0.062***	0.357	-0.015
	(7.76)	(1.57)	(-1.32)
RetVol	-0.073***	-1.208**	-0.028
	(-2.83)	(-2.47)	(-0.49)
AbsChEPS	0.014	-3.465***	0.030
	(0.58)	(-5.14)	(0.52)
MktAdjRet	0.008	0.135	-0.012*
	(1.18)	(1.46)	(-1.73)
R&D	-1.573	-33.330	-3.275
	(-0.69)	(-1.27)	(-0.85)
Size	-0.357	-4.616	-1.747***
	(-0.89)	(-0.92)	(-3.42)
Coverage	-0.131	-0.359	0.028
	(-1.60)	(-0.50)	(0.42)
InstOwn	-0.009	-0.260***	-0.011**
	(-0.65)	(-2.75)	(-2.08)
D_Issuance	-0.066	3.286	0.782**
	(-0.25)	(1.33)	(1.96)
MB	0.681**	2.770***	0.051
	(2.06)	(2.75)	(0.16)
Lev	-0.028*	0.051	-0.013
	(-1.85)	(0.21)	(-0.74)
StdEarn	-0.079	-0.861	-0.157
	(-0.64)	(-0.34)	(-0.91)
D_PosChEarn	0.602***	8.719***	1.583***
	(4.06)	(3.58)	(8.05)
Optimism	-1.018***	-10.037***	-5.044***
	(-4.49)	(-3.69)	(-8.39)
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
# Obs	4,342	4,342	4,342
Adj. R^2	0.625	0.096	0.079

Table 10. Test for Reverse Causality: Can Changes in Management Forecasts Predict the Initiation of Satellite Data Release?

This table reports the following regression for the five-year period before and after the initiation year of satellite data release for the treatment sample of retail firms:

$$ReleaseInitiation_{i,t} = \alpha_t + \beta \Delta Ln(\#MgmtFcst + 1)_{i,t-1} + \gamma \Delta \boldsymbol{z_{i,t}} + \epsilon_{i,t},$$

where $ReleaseInitiation_{it}$ takes a value of one if year *t* is the initiation year of satellite data release for firm i, and zero otherwise. $\Delta Ln(\#MgmtFcst + 1)_{it-1}$ is the lagged change in the number of management forecasts relative to the previous year. $\Delta \boldsymbol{z_{i,t}}$ includes the change of all control variables used in our previous main regression, where the change is measured relative to the previous year. Since the specification of first difference removes unobserved firm-specific fixed effects, only year fixed effects are included in the regression. Columns 1-2 present the results based on the OLS regression and columns 3-4 based on the logit regression. The *t*-statistics based on standard errors clustered at both the firm and year levels are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively. Coefficients on main variables of interest are in bold.

	OLS		**Logit**	
	(1)	(2)	(3)	(4)
ΔLn(#MgmtFcst+1)	**-0.002**	**-0.001**	**-0.019**	**0.015**
	(-0.20)	**(-0.11)**	**(-0.19)**	**(0.09)**
ΔRetVol		0.032		0.784
		(0.14)		(0.21)
ΔAbsChEPS		0.017		0.376
		(0.38)		(0.28)
ΔMktAdjRet		-0.002		-0.038
		(-0.29)		(-0.18)
ΔR&D		-9.189		-252.567
		(-1.60)		(-1.04)
ΔSize		0.015		0.257
		(0.36)		(0.47)
ΔCoverage		0.010**		0.116***
		(2.24)		(3.40)
ΔInstOwn		-0.014		-0.261
		(-0.22)		(-0.33)
ΔD_Issuance		-0.044		-0.478
		(-1.20)		(-1.23)
ΔMB		-0.000		-0.001
		(-1.46)		(-0.42)
ΔLev		-0.159		-1.538
		(-1.04)		(-0.87)
ΔStdEarn		0.103		1.233
		(0.30)		(0.31)
ΔD_PosChEarn		-0.008		-0.131
		(-0.54)		(-0.87)
ΔOptimism		-0.000		0.202
		(-0.09)		(0.89)
Year FE	Yes	Yes	Yes	Yes
# Obs	926	926	733	733
Adj. R^2(Pseudo R^2)	0.165	0.166	0.154	0.176

Table 11. Address Potential Biases in Staggered DID Regressions due to Treatment Effect Heterogeneity

This table reports the average treatment effect and dynamic treatment effects based on the Callaway and Sant'Anna estimator. The dependent variable is *Ln(#MgmtFcst+1)* in columns 1-4 and *Ln(#MgmtSalesFcst+1)* in columns 5-8. Columns 1-2 and 5-6 estimate the average treatment effect for all cohorts across all periods. Columns 3-4 and 7-8 estimate the dynamic treatment effects for each period relative to the treatment year across all cohorts. For brevity, we only report coefficients from *g*-3 to *g*+3. We use never-treated or not-yet-treated firms as effective control groups. The sample period is from 2008 to 2018. The *t*-statistics based on standard errors clustered at both the firm and year levels are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively. Coefficients on main variables of interest are in bold.

	Ln(#MgmtFcst+1)				*Ln(#MgmtSalesFcst+1)*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATT	**-0.263***** **(-4.27)**	**-0.265***** **(-4.28)**			**-0.227***** **(-3.37)**	**-0.227***** **(-3.37)**		
g-3			**0.000** **(0.00)**	**-0.000** **(-0.01)**			**-0.018** **(-0.32)**	**-0.017** **(-0.30)**
g-2			**0.013** **(0.33)**	**0.012** **(0.31)**			**0.003** **(0.06)**	**-0.000** **(-0.01)**
g-1			**-0.011** **(-0.31)**	**-0.010** **(-0.28)**			**-0.043** **(-0.86)**	**-0.040** **(-0.80)**
g+0			**-0.132***** **(-3.07)**	**-0.133***** **(-3.09)**			**-0.125***** **(-2.68)**	**-0.125***** **(-2.67)**
g+1			**-0.270***** **(-3.26)**	**-0.271***** **(-3.27)**			**-0.191***** **(-2.85)**	**-0.190***** **(-2.83)**
g+2			**-0.354***** **(-3.39)**	**-0.356***** **(-3.39)**			**-0.214**** **(-2.37)**	**-0.213**** **(-2.36)**
g+3			**-0.417***** **(-3.15)**	**-0.421***** **(-3.19)**			**-0.356**** **(-2.55)**	**-0.358**** **(-2.56)**
Control Sample	Never-Treated	Not-Yet-Treated	Never-Treated	Not-Yet-Treated	Never-Treated	Not-Yet-Treated	Never-Treated	Not-Yet-Treated
# Obs	5,466	5,466	5,466	5,466	5,466	5,466	5,466	5,466