Security Analysis and the Collection of Hard and Soft Information*

Azi Ben-Rephael[†] Bruce Carlin[‡] Zhi Da[§] Ryan Israelsen[¶]

Current version: January 2025

Abstract

We use minute-by-minute Bloomberg online status microdata during 2017-2021 to directly study how hard and soft information collection affects equity analyst performance. Collection of hard and soft information are measured by office workday length and propensity to travel, respectively. The measures are validated by examining analysts' behavior during earnings calls, and their coverage and forecast decisions. Both hard and soft information collection improve forecast precision, a causal result that we confirm using the COVID lockdown as an instrument. Soft information collection is positively correlated with likelihood of becoming an All-star analyst. We find that some analysts who travel extensively appear to sacrifice some accuracy to increase their chances of becoming All-stars through stronger relationships.

Keywords: Analyst, Effort Provision, Forecast Accuracy
JEL Classification: G12

*We thank Bingxu Fang, Rawley Heimer, Ron Kaniel, Devin Shanthikumar, Eric So, David Solomon, Andrew Van Buskirk and seminar participants at Alliance Manchester Business School, Boston College, Brigham Young University, HEC Lausanne / EPFL, Hong Kong PolyU, Iowa State University, Lancaster University, MIT, McMaster University, Michigan State University, Nottingham University, Oklahoma State University, Peking University, Renmin University, Rutgers University, Tsinghua University, Washington State University, Western University, York University, Finance Conference at Chapman University, Asian Bureau of Finance and Economic Research 2023 Annual Conference, 5th Future of Financial Information Conference, for helpful comments and suggestions. We also appreciate the advice and counsel from Richard Zansitis (Vice President and General Counsel for Rice University). This study was approved by the IRB at Rice University and use of the data was approved by Bloomberg Finance L.P. A previous version of the paper has been circulated under the title "All in a day's work: What do we learn from Analysts' Bloomberg Usage?"

[†]Rutgers Business School. Email: abenrephael@business.rutgers.edu.
[‡]Rice University, Jones Graduate School of Business. Email: carlin@rice.edu.
[§]University of Notre Dame, Mendoza College of Business. Email: zda@nd.edu.
[¶]Michigan State University, Broad College of Business. Email: israels4@msu.edu.

1 Introduction

Calibrating the importance of hard and soft information in security analysis is typically challenging because its collection is inherently a hidden action. Surely, access to private information is valuable for financial analysts (Green, Jame, Markov, and Subasi, 2014)[1], but characterizing broad cross-sectional trends is typically challenging. Distance measures have been used successfully in a variety of settings, but are likely to be noisy proxies for information collection (Liberti and Petersen, 2019). Except for Malloy (2005), there is still a dearth of evidence that measures the link between the effort market participants employ to collect hard and soft information and the quantity and quality of their security analyses.

In this paper, we analyze the work habits of sell-side analysts *directly* by collecting minute-by-minute Bloomberg usage microdata from September 2017 through March 2021. We study 336 sell-side analysts employed by 42 brokerage firms, and estimate both the time that analysts spend in the office, as well as the time they spend away. This allows us to proxy for their hard and soft information collection and quantify the effect that both types of effort provision have on their ability to forecast earnings and value equities.

Equity analysts use Bloomberg extensively. In our sample, they logged into the platform on 72% of workdays. On those days, they worked actively for more than 8 hours on average, and their pre-market login activities strongly react to overnight information. Among other useful functions, Bloomberg allows analysts to explore financial data, utilize existing analytics and examine research by peer analysts.[2] In addition, it constitutes an online social network community. When individuals sign user agreements with Bloomberg, they are given the opportunity to communicate with each other using the messaging service. As a result, whether a user is actively using the software is publicly observable to all users.

A Bloomberg terminal user's profile page indicates the status of their activity on the platform. A green dot next to an analyst's name indicates that he/she is actively using

[1] See also Soltes (2014), Brown, Call, Clement, and Sharp (2015), Cheng, Du, Wang, and Wang (2016), and Han, Kong, and Liu (2018).

[2] See https://www.bloomberg.com/professional/expertise/analyst

his/her personal account. If the analyst were to become inactive for greater than 15 minutes, the dot would turn yellow. If a user is offline, the dot is red, and if a telephone icon appears, it indicates he/she is using the mobile application.

To analyze the effects of hard information collection, we use an expectation-maximization algorithm to quantify the length of their workday based on Bloomberg usage pattern (Ben-Rephael, Carlin, Da, and Israelsen, 2023). The quarterly measure Average Workday Length (*AWL*) proxies for each analyst's effort to collect and process hard information at work. The average *AWL* in our sample is 9.8 hours. Not surprisingly, *AWL* increased sharply starting during the COVID outbreak in the first quarter of 2020 to almost 11 hours. Note that we do not focus on the intensity or total time of Bloomberg usage in our tests, as we expect analysts to engage in other hard information processing activities at work, such as meetings, working on a spreadsheet, emailing, and reading. Nevertheless, given that analysts are heavy Bloomberg users, we find similar results using their time spent on the platform, as reported in the appendix.

We proxy for soft information collection by using the percentage of workdays when analysts are not on the Bloomberg platform at all (Percentage of Away Days, *PAD*). Each quarter, we define "traveling analysts" as those with a *PAD* above the sample median. Admittedly, there is a possibility that this measures the magnitude of soft information collection with some error. For example, an analyst might be traveling for leisure when they are not using the platform. The results speak against this being a problem. First, the percentage of away days is too high to be consistent with the lack of work. The majority of the analysts in our sample come from the top five brokerage firms, and given the typical work culture of these firms, it seems unlikely that the magnitude of *PAD* is driven purely by leisure. Second, an alternative *PAD* measure where we only count an "away day" if it coincides with an actual event organized by the firm the analyst covers, generates similar results. Last, and most interestingly, we use the COVID lockdown as an instrument, and show that when *PAD* decreased for "traveling" analysts, their forecast precision actually suffered.

We provide three sets of validation tests, further supporting the notion that *AWL* and *PAD* reflect analyst effort to collect hard and soft information, respectively. First, we hand-collect the text of over a half-million analyst statements and questions during Q&A sessions on corporate earnings calls. While many analyst forecasts may reflect a team effort, studying each analyst's inquiry is an individual activity that informs us about a particular analyst. Applying machine learning methods and large language models, we find that higher *PAD* is associated with more specific, qualitative inquiry. In contrast, *AWL* does not show the same pattern. Instead, we find that *AWL* is associated with an increase in quantitative inquiry, whereas *PAD* is not.

Second, we find that analysts with higher *PAD* cover more growth stocks, which may require more soft information, while analysts with higher *AWL* tend to cover larger and mature firms. Third, we find that analysts with higher *PAD*s (*AWL*s) tend to issue more qualitative (quantitative) output, as measured by the ratio of stock recommendations (qualitative) to earnings and price target forecasts (quantitative).

We show that *AWL* and *PAD* are authentic and persistent analyst characteristics. Neither quarter, brokerage-firm, nor sector fixed effects explains more than 15% of their variation. Analyst fixed effects only explain 49.8% and 57.2% of variation in *AWL* and *PAD*. There is a negative correlation between *AWL* and *PAD* ($\rho = -0.23$), and both measures are positively correlated with the number of stocks that analysts cover. Analysts with more experience or who have a high-ranked title are associated with a lower *AWL*. In addition, star analysts or high-ranked analysts are associated with a higher *PAD*. We control for experience and include a seniority indicator in our regressions, in order to isolate the effect coming from analyst effort. Including an analyst fixed effect in our analyses further controls for other persistent analyst characteristics.

We study the relationship between analyst effort and performance, and use two metrics. The first is whether an analyst becomes an Institutional-Investor (II) All-Star. According to (Groysberg, Healy, and Maber, 2011), All-Star analysts earned 61% higher compensation on average and gaining/losing All-Star status was associated with a 16% increase/decrease in pay. Similarly, Brown et al. (2015) report that 67% of their survey respondents rate analysts' standing in rankings/broker votes to be very important. We find that analysts who travel more during the first three quarters of the year are more likely to be voted as a star analyst in quarter 4 by the *Institutional Investor* magazine. In contrast, *AWL* is not significantly related to All-Star status. This evidence suggests that *PAD* likely captures the time that analysts and institutional clients spend together. Interestingly, among analysts who were not star analysts in the previous year, only those in the top *PAD* quintile were significantly more likely to be voted as star analysts in the current year. This suggests perhaps that non-All-Stars may invest in travel, hoping to achieve All-Star status in the new year.

The second analyst performance metric we examine is the accuracy of EPS forecasts. Brown et al. (2015) report that 35% of the surveyed analysts view accuracy as an important determinant of their compensation. Using proprietary salary data, Groysberg, Healy, and Maber (2011) show that poor forecasts are important for termination. In addition, earnings forecast accuracy is probably the most widely studied analyst output in the finance and accounting literature.

Following Clement (1999) and Jame, Johnston, Markov, and Wolfe (2016), we compute a "Proportional Mean Absolute Forecast Error" (*PMAFE*), which compares each analyst's forecast error to those of their peers covering the same earnings announcement. We find *AWL* to be significantly related to improved accuracy. A one standard deviation increase in *AWL* is associated with a reduction of about 2% in *PMAFE*'s standard deviation units. The impact of *PAD* on accuracy is more nuanced and non-monotonic. *PAD* is associated with improved accuracy, but the improvement in accuracy is concentrated in the group of analysts that travel above the median but not in the top 20%. The top 20% have better accuracy than analysts with low *PAD*, but it seems that analysts in the top *PAD* quintile sacrifice some accuracy to make corporate connections that might have other advantages, like becoming an All-Star. These patterns appear to be robust to team effort, which is shown to be important in Fang and Hope (2021).

To establish a causal effect of *AWL* and *PAD*, we use two instruments. The first is the COVID lockdown that exogenously curtailed travel during the first two quarters of 2020. This shock should hurt "traveling" analysts more than their peers. Indeed, we find that analysts whose *PAD*'s exceed the sample median pre-COVID (during the last two quarters of 2019) experienced a significant increase in their *PMAFE*s (or reduction in accuracy) of 11.7%. In addition, the increased relative forecast error is concentrated among faraway firms whose headquarters are at least 300 miles from the "traveling" analyst.

The COVID lockdown is less effective as an instrument for *AWL* since there is no clear ex-ante separation, as is the case for *PAD*. A better instrument that offers such separation is the pre-lockdown commute time, which we estimate using the distance between each analyst's home and corporate address from Google Maps. Analysts who spent a longer time commuting to work during the last two quarters of 2019 would ostensibly save more time by working from home. We find that one-hour commuting time pre-COVID predicts a 1.3 hour increase in *AWL* during the lockdown. Using commuting time as an instrument for increased *AWL*, we find that *AWL* significantly improves the accuracy of the forecasts (a reduction of *PMAFE* of 8.5%).

The importance of hard and soft information in finance cannot be overstated, both for raising capital and the pricing of traded financial assets. While distance measures have been used extensively for the former (e.g., Lerner, 1995; Garmaise and Moskowitz, 2004; Butler, 2008)[3], they are less attractive as a proxy when studying security analysis.[4] This is because information collection is inherently a hidden action. Distance is likely to be a noisy proxy, especially for soft information collection. For example, a distance-based measure would

[3]See Liberti and Petersen (2019) for an excellent review. Distance measures have been used to distinguish hard and soft information collection in equity markets (Coval and Moskowitz, 1999; Ivkovic and Weisbenner, 2005; Loughran and Schultz, 2005), the municipal bond market (Butler, 2008), the venture capital market (Lerner, 1995), the real estate market (Garmaise and Moskowitz, 2004), and in the market for distressed assets (Granja, Matvos, and Seru, 2017). The thesis in these papers is that hard information can be transmitted across distance, whereas soft information cannot.

[4]One exception is Malloy (2005) who finds that analysts located closer to firm headquarters have more accurate forecasts

assume that two analysts in the same location have the same information, which may not be true based on their effort provision. So, our paper contributes to this literature in that we measure information collection more directly.

Our paper also adds to a series of papers that show that collecting soft information is valuable for security analysis. Green et al. (2014) show that access to management at broker-hosted investor conferences leads to analyst recommendation changes that have larger immediate price impacts. Brown et al. (2015) survey 365 analysts and find that private communication with management is more useful to analysts than their own primary research, recent earnings performance, and recent 10-K and 10-Q reports. Cheng et al. (2016) show that analysts who visit corporate sites have better forecast accuracy than others. Han, Kong, and Liu (2018) show that visits to listed companies lead to improvements in forecast accuracy.

The rest of the paper is organized as follows. Section 2 provides information about our data and economic variables. Section 3 characterizes the determinants of *AWL* and *PAD*. Section 4 describes how our measures of hard and soft information affect the performance of analysts. Section 5 describes the use of the COVID lockdown and commuting data as instruments to deal with potential endogeneity. Section 6 concludes.

2 Sample Construction and Analyst Work Habit Measures

This section describes how we construct our sample of sell-side analysts and measures of their hard and soft information collection. Table A.1 provides variable definitions for all variables used in this paper.

2.1 Sample Construction

Bloomberg Usage Data:

When Bloomberg users are assigned accounts, the company records their "status" by default.

Status is either designated as "online", "idle", "offline", or "mobile". When users first log on to the platform, their status changes from offline to online, and it remains that way while they use Bloomberg. However, if they stop using it for 15 minutes, the user's status automatically changes to "idle". Eventually, and depending on the users' settings, a user is logged off after a long period of inactivity. Also, when users are logged in via the "Bloomberg Anywhere" application on their mobile device, the status is listed as "mobile". While using the mobile app, access to an assigned desktop is restricted, so there is no possibility of double counting.

Other users of the platform can detect the status of any other Bloomberg user by employing the "PEOP" function, the "BIO" function, or by directly navigating to a user's profile. A green dot by a user's name indicates that he/she is online and active. Other status indicators are as follows: a red dot means that a user is offline, a yellow dot means that a user is idle, and a gray dot indicates that a user has chosen to be private. If a user is online via the mobile app, a mobile phone icon appears.

Analyst Data:

Since 2017, we have observed and recorded the profile status and the time spent on Bloomberg for a few thousand users who self-identified as "analysts." Some of them are credit analysts, analysts working for buy-side firms, or simply have the title "analyst" without actually being one. We identify 997 sell-side equity analysts among them by cross-referencing them to the IBES recommendation file. We verify that the individuals are the same based on their full names, the brokerage firms and locations.[5] Requiring non-missing IBES output further reduces the number of analysts to 710.

We restrict the sample to analysts who are active on Bloomberg. To be considered as an active Bloomberg user, an analyst needs to have at least one quarter with a quarterly average percent activity greater than 3%. Percent activity is the time in minutes that an

[5] The alternative is to start with all IBES analysts and identify them on Bloomberg. This alternative procedure is less efficient and likely error-prone as the IBES recommendation file only provides the initials of analysts' first names.

analyst is actively logged on, scaled by the number of minutes within a day, so 3% means around 40 minutes of Bloomberg usage per day. This cut-off removes the left tail of the login distribution, which is populated by inactive users. In addition, we require an analyst to be reasonably active in IBES, meaning that they issue at least two earnings forecasts per quarter and cover at least 3 stocks. These minimum Bloomberg and IBES activity filters result in a final sample of 336 analysts across 42 brokerage firms. We also collect all of their recommendations across all US stocks as well as their earnings per share forecasts, across all horizons, long term growth forecasts, and 12-month price target forecasts. Information on star analysts is obtained from *Institutional Investor* Magazine's All-America Research Team rankings.[6]

Overall, our sample includes about 15% of all active IBES analysts in these 42 brokerage firms. The sample attrition mostly comes from the fact that many sell-side analysts do not self identify as "analysts" on Bloomberg. We verify that analysts in our sample are similar to their peers from the same brokerage firm. In other words, this attrition should not impose any systematic bias in our analyses.

Analyst Participation on Earnings Calls:

We quantify various dimensions of analysts' participation on earnings calls. Unlike examining analyst forecasts which may reflect a team effort, analyzing what each analyst asks during earnings calls informs us about that particular analyst. To do this, we collect each analyst's statements and questions made during firms' earnings conference calls using the Refinitiv StreetEvents Transcripts database. We manually identify our sample using the names and brokerage firms from the earnings call transcripts during our sample period, and then collect the text of their statements and questions that they made during the Question and Answer section of the conference calls.

We split the text into sentences and extract two sets of measures from the analysts' statements and questions. The first type of measure we generate based on analyst earnings

[6] We thank An-Ping Lin for sharing his data on star analysts.

call statements is the number of "named entities" included in their language. Named entities are specific people, organizations, places, dates. etc. Hence, the frequency of named entities in a statement is a measure of the specificity of the statement.[7] The inclusion of more specific information in a statement or question indicates more preparation by the analyst. By contrast, little preparation or insight is needed to make general, vague statements or questions.

We use the Python SpaCy natural language processing library to extract named entities from each sentence. More specifically, we use the RoBERTa transformer model within SpaCy which is able to classify words into 18 categories of named entities.[8] For each sentence, we extract and count the number of named entities as well as the total number of words in the sentence. Then, we aggregate to the analyst-quarter frequency by dividing the total number of named entities by the total number of words across all of the analyst's sentences that quarter.

The second measure we generate from the earnings call transcripts is whether each sentence is quantitative in nature. For example, a sentence mentioning (or asking for) a specific accounting value would be considered quantitative, while a question asking if there are any new product opportunities would be considered non-quantitative. We use a Large Language Model (LLM) to classify sentences into these categories for two reasons. First, LLMs provides a consistent classification across sentences. The second reason is due to the extremely large number of sentences that need to be classified – more than half a million. For each sentence, we prompt the OpenAI "gpt-3.5-turbo" model to classify each sentence into one of the two categories and to provide a justification for the classification.[9] Using these classifications, we create quarterly measures of quantitative sentence frequency for each analyst

[7]Named entities have been used to identify the impact of specific versus boilerplate language used by firms and investors in their SEC filings. See, for example, Hope, Hu, and Lu (2016), Cazier, McMullin, and Treu (2021), and Israelsen, Schwartz-Ziv, and Weston (2024).

[8]We use the model en_core_web_trf. See the description at `https://spacy.io/models/en/#en_core_web_trf`. Table A.2 in the appendix includes a few examples of extracted named entities.

[9]Table A.3 in the appendix includes the full prompt we provide the LLM as well as some examples of sentences classified into each category and the justification.

by dividing the total number of quantitative sentences by the total number of the analyst's sentences during a specific quarter.

2.2 Analyst Work Habits Measures

Average Workday Length (*AWL*):

To measure AWL, we use an unsupervised machine learning algorithm - the Gaussian Mixture Model - to quantify analysts' time spent on hard information collection and processing in a given quarter based on their Bloomberg usage patterns. The same methodology was used in Ben-Rephael et al. (2023) and validated there using cellphone geolocation data. In that paper, we measured AWL for top executives (e.g., CEO's and CFO's) in U.S. firms and used it as a proxy for work effort. We showed that AWL is associated with higher firm value and that long-short portfolios using computed AWL earned abnormal risk-adjusted returns.

In Ben-Rephael et al. (2023), we showed that our results were robust to using other distributional measures, but that AWL proxies for effort provision in a very intuitive way. Figure 1 illustrates the algorithm for a specific analyst-quarter observation. In the figure, the blue bars represent relative usage patterns throughout each workday during the quarter. The overall usage pattern resembles the mixture of two normal distributions: one in the morning and one after lunch. This pattern holds generally across most analysts. Clearly, the usage pattern is not derived from a distribution, per se, but we use this observation to construct our Average Workday Length (AWL) measure based on a mixture of normal distributions as follows.

For each analyst and quarter, we calculate the probability P^j_{min} as the percentage of the time that an analyst is actively using the platform during all workdays in that specific quaerter, where $j \in J \equiv$\{12:00 am, 11:59 pm\}. Then, using these relative frequencies, we construct a pdf by computing $p^i_{min} = P^i_{min} / \sum_J P^j_{min}$. By construction, $\sum_J p^j_{min} = 1$. This pdf captures the likelihood of the time of the analyst's terminal activity during the quarter. We then assume that the constructed distribution is a mixture of two normal distributions

$k \in \{1, 2\}$, each with mean μ_k and variance σ_k^2, where $\mu_2 > \mu_1$. This captures the notion that analysts' work habits may differ before and after lunch. As mentioned, a dip in activity around lunchtime is very frequent in our sample.

For the mixed distribution, there is a probability q that any realization is drawn from distribution 1 and probability $(1 - q)$ that it was drawn from distribution 2. The mixed distribution has mean $\mu_{1,2}$ and variance $\sigma_{1,2}^2$, which can be measured for each analyst. We also have the following relationships:

$$\mu_{1,2} = q\mu_1 + (1 - q)\mu_2 \tag{1}$$

$$\sigma_{1,2}^2 = q\sigma_1^2 + (1 - q)\sigma_2^2 + q(1 - q)(\mu_2 - \mu_1)^2 \tag{2}$$

Using these two equations, we perform an expectation-maximization (EM) algorithm to estimate all five parameters for each analyst $(q, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$.

The EM algorithm consists of two steps: the estimation step (E-Step) and the maximization step (M-Step). In the E-Step, the expectation of the log-likelihood function is calculated for a given set of parameters. In the M-Step, the parameters are re-chosen in order to maximize the expectation. The process continues, iterating between the E-Step and the M-Step until the sequence converges. In our case, the likelihood function involves the likelihood of observing the data given that there are two unobservable Gaussian distributions generating the data. We implement the procedure using the skikit-learn libarary for Python.[10]

Returning to the example in Figure 1, we see the estimated Gaussian Mixture Model pdf in red as well as the two underlying Gaussian distributions in orange for this analyst-quarter observation. The dashed vertical bars are the estimated means of the two distributions. The two black lines represent the beginning and end of the AWL measure, or the interval $(\mu_1 - \sigma_1, \mu_2 + \sigma_2)$.[11] For this example, AWL is 9.12 hours.

Since *AWL* is measured using Bloomberg usage patterns, it naturally captures the average

[10] We use the sklearn.mixture.GaussianMixture method with a convergence threshold of 0.001 and K-Means clustering to initialize the parameters.

[11] An alternative *AWL* can be computed as the length of an interval that covers the middle 90% of the usage distribution. We confirm that such an alternative measure gives similar results.

time spent on hard information collection and processing per day in that quarter (when the analyst is not traveling). Note that the measure does not require the analyst to be active on Bloomberg for the entire 9.12 hours. The analyst could also be collecting and processing hard information by reading periodicals, doing spreadsheet modeling, or meeting with colleagues. Assuming that the analyst generally logs in to Bloomberg near the start of their workday and logs off near the end, the *AWL* measure also captures these other non-Bloomberg work activities.

Since analysts in our sample spend a non-trivial amount of time on Bloomberg, we also consider an *intensive* usage measure. The measure, *LnActive*, is calculated as the natural logarithm of the average daily minutes of active Bloomberg usage in a quarter. Table A.5 in the Appendix confirms that the main results are similar if we replace *AWL* with *LnActive*.

Percentage Away Day (*PAD*):

To quantify the extent of soft information collection that requires travel, we count the days when the analyst does not log in to Bloomberg at all. We first define a daily dummy variable that receives the value of one if an analyst is not logged in to Bloomberg during that day, and zero otherwise. Then, we average the dummy variable within a quarter to compute the Percentage Away Days (*PAD*).

Clearly, *PAD* measures analysts' work-related travel with some error. While analysts in our sample are heavy Bloomberg users, it is still possible that on some days, analysts may work in the office without using Bloomberg at all. In addition, even if they are away from the office, there is no guarantee that they are traveling for work-related reasons rather than vacationing. To the extent that analysts have similar total numbers of annual vacation days, the cross-sectional variation in *PAD* should still reveal differences across analysts in their soft information collection effort.

If anything, this bias works against our finding a benefit to being away from the office. But, as we show later in the paper, high levels of *PAD* are associated with a higher probability of becoming a star analyst, indicating that this does not capture systematic noise or leisure.

More importantly, we use the travel restriction during the COVID lockdown as an instrument and show that fewer days away led to less accurate EPS forecasts for analysts who tend to be away from the office. In addition, in the Appendix, we repeat the main tests using a percentage away measure that takes into account information events (*EvPAD*). Specifically, *EvPAD* is calculated using away days that coincide with brokerage and firm events for stocks the analyst covers. Since *EvPAD* is only based on firm and brokerage firms' events, it does not capture other interactions with institutional investors or other firm site visits. Thus, we view it as a lower bound for information-gathering activities and focus on *PAD*.

Throughout the paper, we report results using *PAD*. In addition, when we explore analyst performance (the probability of being a star analyst and accuracy), we also present results using two sets of dummy variables that allow us to capture non-linearities and present a more nuanced view. First, we identify traveling analysts as those whose *PAD* is above the median (*PAD_HIGH*). Second, to better understand the dynamics of the *PAD_HIGH* group, we slice *PAD_HIGH* into analysts in the top quintile of the *PAD* distribution (*PAD_TOP*) and a middle group (*PAD_MED*). Overall we conjecture that traveling analysts are more likely to specialize in acquiring soft information from attending events organized by the firms, meeting management face-to-face, and visiting sites and institutional investors. In contrast, analysts with low *PAD*s are more likely to rely on hard information when making forecasts.

Using earnings call data, section 3.3 presents supporting evidence that *PAD* and *AWL* are related to soft and hard information collection, respectively. For example, we find analysts with higher *PAD*s to focus more on specific information that may have been acquired on the road, visiting management, investors, and other market participants. Similarly, we find analysts with higher *AWL*s to focus on hard, quantitative information.

2.3 Summary Statistics

Table 1 provides summary statistics of analyst output during the sample period. In Panel A we report statistics for the Bloomberg sample. The sample includes 2,874 analyst-quarter observations with 336 distinct analysts from 42 brokerage firms. In Panel B we contrast the Bloomberg sample with a comparable I/B/E/S analyst sample (the comparison sample). To be included in the comparison sample, we require an analyst to cover at least 3 stocks, to be on I/B/E/S for at least four quarters, and to belong to one of the 42 brokerage firms in our Bloomberg sample. The comparison sample includes 1,854 distinct analysts and 16,239 analyst-quarter observations.

Starting with Bloomberg analysts, we find that the average number of unique stocks covered over the previous four quarters is 17.85. The number of unique industries based on GICS 6-digit codes is 3. The average number of Q1 (Y1) forecasts in a given quarter is 23.1 (24.79). This is based on 16.07 unique stocks, where 77% of the forecasts are for common stocks (Share code 10 or 11). Other forecasts include long-term growth with an average of 5.67 forecasts, stock recommendations with an average of 3.28 recommendations, price targets with an average of 11.8, and all other forecasts with an average of 140.1 forecasts. The number of stock recommendations and price targets is lower than the number of earnings forecasts, with an average of 3.28 and 11.81, respectively.

Panel B reports each group averages together with their differences and associated p-values. Overall, the comparison reveals that Bloomberg analysts are more active than those in the comparison sample, but the differences are not large. For example, Bloomberg analysts cover 2 more stocks and issue 1.75 more quarterly forecasts, on average. Bloomberg analysts also issue 0.4 (1.36) more recommendations (price targets). These differences come from the fact that active Bloomberg analysts in our sample are more likely to come from larger brokerage firms. Indeed, 55% of them come from the largest 5 brokerage firms. These firms have more resources to assign Bloomberg accounts to individual analysts so our effort measures are less likely to reflect shared Bloomberg terminal usage by a team of analysts.

Finally, both groups display better accuracy than analysts who are not in the same 42 brokerage firms.[12] This is consistent with the fact that larger brokerage firms have

[12] The forecast accuracy measure is defined in details in Section 3.3. It is normalized so the most accurate forecast takes the value of -1 while a median forecast takes the value of 0.

more resources leading to more accurate forecasts. Interestingly, the Bloomberg group displays higher portfolio accuracy relative to the comparison group on an equally weighted basis. However, these differences shrink and are no longer statistically significant on a value-weighted basis, based on stock market capitalization.

Next, Table 2 reports summary statistics of analysts log-in activity on Bloomberg (Panel A), together with the log-in based measures (Panel B), and their correlation matrix (Panel C). Panel A indicates that, on average, analysts are logged in to the terminal on 71.7% of the work days. Analysts are active on average 362 minutes (6 hours) per day, which amounts to 30.14 hours per week.

Providing more granular information, Figure 2 depicts the average time spent on Bloomberg by day-of-the-week and holidays. As in Panel A of Table 2, the daily time spent on the terminal is around 6 hours, but it drops to 5 hours on Fridays. The log-in activity is small during weekends and holidays. In addition, Graph A of Figure 3 plots the average daily minute activity across analysts in a given quarter over time. There is a sharp increase in the minutes spent on the platform starting the first quarter of 2020 (the COVID period).

Panel B of Table 2 provides statistics of the log-in based measures of analyst work habits (*AWL* and *PAD*). The average *AWL* during the sample period is around 9.8 hours with a tight distribution. Eighty percent of the time, *AWL*s range from 8 hours to 12 hours. We can see a shift in the distribution during the COVID period, which was affected by work-from-home. As for *PAD*, the average is 0.283. Compared to *AWL*, the distribution of *PAD* is wider, with the 10th percentile of 0.033 and the 90th percentile of 0.656. In a similar manner, we document a shift in the distribution of *PAD* during the COVID period, when traveling was restricted. We utilize the differences in *AWL* and *PAD* during the pre-COVID and COVID periods in our analysis and identification strategies.

For emphasis, *AWL* is different from intensity of Bloomberg usage. Using intraday distribution of Bloomberg usage within a quarter, *AWL* aims to measure the typical length of analyst' workday in that quarter, without assuming Bloomberg usage throughout the day.

We measure the intensity of Bloomberg usage using *LnCondActive*, defined as the natural logarithm of the average daily minutes of active Bloomberg usage conditioning on days with Bloomberg activity in a quarter. The correlation between *AWL* and *LnCondActive*, while positive, is only 0.25. Alternatively, one can focus on the average daily usage of Bloomberg during the quarter as a measure of gathering information. Thus, we also define *LnActive*, as the natural logarithm of the average daily minutes of active Bloomberg usage across all days during the quarter. The correlation between *AWL* and *LnActive* is 0.27, which is not different from the one reported with *LnCondActive*. Finally, the correlation between *AWL* and *PAD* is negative, but not huge ($\rho = -0.23$). This suggests that hard and soft information collection effort are not perfect substitutes for each other.

Graphs A-C of Figure 3 provide additional information at the quarterly level. Similar to the minutes spent on the terminal, *AWL* has increased from around 9.5 hours during the early part of the sample to more than 10.5 hours during the COVID period. In a similar manner, *PAD* dropped significantly from Q1 of 2020.

Finally, Figure 4 depicts the log-in measures averages based on stock coverage deciles. In particular, we rank analyst-quarter observations based on the number of stocks that an analyst covered during the recent year. Decile 1 (10) refers to the lowest (highest) number of stocks covered. It is probably not surprising that *PAD* generally increases with the number of stocks covered. For *AWL*, we also observe a positive relation with the stock coverage beyond the first three coverage deciles. In our empirical tests, we control for such mechanical correlations with coverage $\times$ time fixed effects, whenever possible.

3 Determinants of *AWL* and *PAD*

3.1 Other Analyst Characteristics

In this subsection, we first explore how much of the variation in *AWL* and *PAD* is explained by time (year-quarter), analyst, industry coverage, and broker fixed effects. We then regress

AWL and *PAD* on a battery of analyst characteristics obtained from FINRA's BrokerCheck website, LinkedIn, and Facebook.

Almost every analyst in our sample is registered with FINRA BrokerCheck. These records include the full name (including middle name as well as other names used) of each analyst as well as their work histories, the locations of their branch offices, and which FINRA Qualification Exams the analysts have passed. The full name and work history from FINRA help us locate LinkedIn accounts, which provide educational background, and Facebook accounts, which help identify whether analysts have children.

Panel A of Table 3 indicates that analyst fixed effects are the most important determinant in explaining the variation in both *AWL* and *PAD*, with an R-squared of 43.2% and 51.5%, respectively. So, *AWL* and *PAD* both appear to be independent and authentic analyst characteristics. Next, broker fixed-effects explain 8.2% and 11.5% of the variation in *AWL* and *PAD*, which is consistent with workplace culture. Both analyst characteristics also change over time, with time fixed-effects explaining 5.1% and 9.0% of the variation in *AWL* and *PAD*. The time variation is in part due to the COVID lockdown as evident in Figure 3. Finally, industry fixed effects, based on the analyst's main covered GICS6 industry, explain around 8.6% and 6.7% of the variation in *AWL* and *PAD*, suggesting that information collection effort differs based on the type of stocks that the analysts are covering.

The analyst characteristics reported in Panel B of Table 3 reveal that analyst time on I/B/E/S (*IBES Years*), seniority (*High Rank Indicator*), and being a star analysts are three important determinants of *AWL* and *PAD*. An increase in years in the I/B/E/S sample leads to a significant reduction in *AWL*. *PAD* on the other hand, exhibits a positive sign, but the effect is not statistically significant. Second, greater seniority leads to a lower *AWL* and a higher *PAD*. We, therefore, control for both *IBES Years* and *High Rank Indicator* in subsequent analyses when we relate *AWL* and *PAD* to analyst performance. Finally, we find that being a star analyst is positively associated with *PAD* but not *AWL*. This is consistent with the fact that analyst ranking depends on interactions with institutional investors, who are the ones ultimately voting on analysts.

Other work experience variables such as total work experience (*Work Experience*) and the number of jobs that an analyst had switched (*# Jobs FINRA*) are not statistically nor economically significant. In addition, variables such as NYC location, MBA degree, gender, children, and qualifying exam do not load significantly or consistently across the *AWL* and *PAD* specifications. These variables only add around 0.003- 0.027 to the R-squared. Finally, including brokerage firm fixed effects does not alter these findings, but adds between 0.051-0.113 to the R-Squared.

3.2 Login Activity and Market Information

As mentioned, Bloomberg allows analysts to explore financial data, utilize existing analytics, and examine research by peer analysts. In this subsection, we provide evidence on this link by exploring Bloomberg analysts' login activity in response to market events concerning the stocks they cover (hard information). We show that analysts increase their login activity in response to public information about the stocks they cover. To study this link, we focus on login activity between 7-9 am (the pre-open period), which is more likely to reflect analysts' processing of overnight news. Table 4 reports the findings.

We find that analysts increase their login activity if stocks they cover are in the top decile based on abnormal trading volume over the previous day. Also, various measures of news (RavenPack News Analytics) indicate that analysts increase their login behavior if stocks they cover have fundamental news – either after-market-close of the previous day or before-market-open of the current day. This is particularly strong for earnings news, where analysts respond to both stock level news and industry news. For example, a one standard deviation increase in the number of stocks with before-market-open earnings news leads to a $(0.43 \times 0.079 =)$ 0.034 increase in abnormal login activity. Since the average login activity during 7-9 am is around 0.269, this means an increase of 12.6%. Finally, the pre-market login activity is positively correlated with *AWL* (a correlation of 0.24), which highlights the link between *AWL* and analyst effort to collect and process hard information.

3.3 Information Collection

In this subsection, we present three tests that examine the link between analysts' *AWL*, *PAD*, and the type of information they collect. The first set of tests utilizes information from analyst participation in earnings conference calls. The second test explores the relation between *AWL*, *PAD*, and analyst stock coverage decisions. The third test explores the relation between *AWL*, *PAD*, and the types of analyst output.

3.3.1 Earnings Calls Participation

We use advanced NLP algorithms and ChatGPT to systematically analyze different features of analysts' discussions during earnings conference calls using all transcripts of stocks covered in our sample. Our measures are at the analyst-quarter level based on all stocks covered by the analyst.

The first measure, *% Named Entities*, uses NLP algorithms to identify name entities that capture the specificity of analysts' discussions. Named entities include categories such as event, location, person, organization, product, and facility that are aimed to capture how specific their statements are. We remove named entities that capture pure quantities (cardinal, percent, quantity, and money) to distinguish from our quantitative measure. For each analyst and quarter, our measure, *% Named Entities*, is the percentage an analyst words made up by non-quantity named entities across all earnings calls for the analyst that quarter. The second measure, *% of Quantitative Sentences*, uses ChatGPT to identify analyst sentences that are classified as quantitative. The measure is the percentage of quantitative sentences to the total number of sentences across all earnings calls for the analyst that quarter.

We explore the relation between *AWL* and *PAD* and subsequent quarter earnings call participation. Table 5 reports the results. Given that traveling was restricted during 2020 and companies implemented work-from-home policies, we report results for the entire sample period ("ALL") and for the pre-COVID period (2017-2019). Panel A reports the results for the percentage of named entities. To ease economic interpretation, we Z-Score adjust (a mean of zero and a standard deviation of one) both the dependent variable and independent variables of interest. The coefficient estimates represent the effect of 1 standard deviation of X in terms of standard deviation units of Y.

PAD coefficients are positive and statistically significant and indicate that a higher *PAD* is associated with discussions that include more specific information. The results are economically significant and hold with and without the inclusion of analyst fixed effects. A one standard deviation increase in *PAD* is associated with a 5.4% - 13.6% increase in specific discussions in terms of the dependent variable standard deviation units. In contrast, *AWL* does not show the same patterns. The results are neither statistically nor economically significant. The fact that *PAD* is associated with specific discussions is consistent with the conjecture that high *PAD* analysts travel to collect specific/soft information from various market participants and convey that information when speaking with management.

Panel B explores the relation between *PAD* and *AWL* and the percentage of quantitative sentences. The picture that emerges from the analysis is that *AWL* is associated with an increase in quantitative questions, whereas *PAD* is not. A one standard deviation increase in *AWL* is associated with a 4.4% - 5.2% increase in quantitative discussions in terms of the dependent variable standard deviation units. This suggests that *AWL* is associated with the collection of hard information. These results echo the findings of the results reported in Panel A. They are consistent with the notion that *AWL* is associated with more quantitative-oriented information, while *PAD* is associated with more specific/soft information. *AWL* becomes insignificant in the last two columns when Analyst FEs are included, suggesting that "quantitativeness" is a persistent analyst characteristic.

3.3.2 Coverage Decisions

Next, we are interested to learn if *AWL* and *PAD* are associated with different stock coverage decisions. To this end, each quarter, we rank all the stocks in our sample into quintiles based on selected firm characteristics. Then, for each analyst and quarter, we calculate the stock market cap weighted average of each ranking across all the stocks covered by the analyst. We then run quarterly panel regressions excluding analyst fixed effects to capture cross-sectional differences in coverage across analysts.

Table 6 report the results where we Z-score both the dependent and independent variables of interest. We find that *AWL* loads positively on firm age, market cap, and price, and negatively on illiquidity. Thus, analysts with higher *AWL* prefer to cover larger, mature firms that are more liquid. This may suggest that mature and large firms may rely less on soft information. We further find that *PAD* also loads positively on size and negatively on illiquidity, but also loads positively on growth (inverse of BM) and momentum. This suggests that analysts with higher *PAD* tilt their coverage toward growth stocks, which may require more soft information. The positive relation between *PAD* and stocks whose value recently appreciated may be driven by institutional demand for momentum stocks, which traveling analysts with presumably better institutional relations accommodate.

3.3.3 Analyst Output

In the last test, we explore the relation between *AWL*, *PAD*, and subsequent quarter output decisions. In particular, analysts issue earnings forecasts, price target forecasts, and recommendations. Earnings forecasts and price targets that have specific numbers are considered to be hard information in nature, while recommendations, which are broader, are considered to be soft information in nature (e.g., Green et al., 2014).

To capture analyst output decisions based on hard and soft attributes, for each analyst and quarter, we calculate the ratio between all the recommendations the analyst issued to the sum of all earnings forecasts and price targets. We then explore the relation between the

AWL, *PAD*, and the analyst's subsequent output decisions using quarterly panel regressions. Overall, we find that an increase in *AWL* is associated with a reduction of this ratio, while *PAD* is associated with an increase in this ratio. Thus, an analyst with a higher *PAD* will issue more recommendations over the subsequent quarter, where the results are more significant once analyst fixed effects are included.

4 Analysts' Information Collection and Performance

Career concerns play an important role, and there are many dimensions that analysts care about (Brown et al., 2015). One important dimension that affects compensation is being ranked as a top analyst. For example, according to (Groysberg, Healy, and Maber, 2011), All-Star analysts earned 61% higher compensation on average and gaining/losing All-Star status was associated with a 16% increase/decrease in pay. Similarly, Brown et al. (2015) report that 67% of their survey respondents rate analysts' standing in rankings/broker votes to be very important. We therefore examine Institutional-Investor (II) All-Star status as an important outcome variable. The All-Star status is individual analyst specific.

Another important dimension is accuracy, where Harford, Jiang, Wang, and Xie (2019) show that career concerns shape effort allocation and stock accuracy. Brown et al. (2015) report that 35% of the surveyed analysts view accuracy as an important determinant of their compensation. Using proprietary salary data, Groysberg, Healy, and Maber (2011) show that poor forecasts are important for termination. Motivated by these studies, we examine analysts' forecast accuracy, which is also the most widely studied analyst performance metric in the finance and accounting literature.

In this section, we explore how *AWL* and *PAD* are associated with both the probability of being ranked as a star analyst and the accuracy of analysts' quarterly forecasts. We acknowledge that analysts optimize across these dimensions to achieve their career goals.

4.1 The Probability of Being a Star Analyst

We explore how *AWL* and *PAD* affect the probability of being ranked as a star analyst. Since the rankings are done in Q4 in each year, we explore the relation between being ranked as a star in year t and the averages of *AWL* and *PAD* in Q1-Q3 of year t. To explore the potential non-linearity of *PAD*, besides *PAD*, we also include in the analysis dummy variables based on the distribution of average *PAD* in Q1-Q3 of year t.

We start with a median split between low and high *PAD*. Then we zoom in on the high *PAD* group. In particular, we define *PAD_HIGH* as a dummy variable that receives the value of one if the *PAD* average is above the median of the average *PAD* distribution in year t, and zero otherwise. In a similar manner, *PAD_MED* (*PAD_TOP*) are dummy variable that receives the value of one if the average of *PAD* is between the 50^{th} and 80^{th} percentiles (above the 80^{th} percentile) of the distribution of the average *PAD* in year t, and zero otherwise. Since we use the averages of *PAD* and *AWL* during Q1-Q3, we limit our analysis to 2018-2020, where we have full information.

Table 8 reports the results. Since we employ a linear probability model, the dependent variable has a natural economic interpretation. Thus, we only Z-score adjust *AWL*, and *PAD*. Panel A of Table 8 reveals that average *PAD* is associated with a higher probability of being ranked as a star analyst, especially if he/she was not ranked as a star analyst in the previous year. These findings are robust to the inclusion of brokerage firm fixed effects and across sub-samples. In contrast, average *AWL* is mostly insignificant, in particular when brokerage firm fixed effects are included. This evidence also suggests that institutional investors value soft information. The results further indicate that a one standard deviation increase in average *PAD* is associated with 4.3% - 10.2% increase in the probability of being a star analyst.

To further explore the non-linearity of *PAD*, we analyze our set of *PAD* dummy variables. Panel B indicates that *PAD_HIGH* is statistically and economically significant. Thus, analysts ranked above the median have a higher chance of becoming star analysts. The economic magnitude of *PAD_HIGH* suggests that most of the effect comes from the *PAD_HIGH* analysts.

To further explore the high group, we slice *PAD_HIGH* into *PAD_MED* and *PAD_TOP*, and report the results in Panel C. The results indicate that the top group *PAD* (i.e., the top quintile) exhibits a higher probability of becoming star analysts than the base or the middle group. First, this alleviates the concern that high values of *PAD* are contaminated with noise in estimation. Second, this shows that taking career concerns into account, analysts in the top group have a higher chance of being ranked as top analysts. For example, Columns 6 and 8 indicate that most of the effect is concentrated in the top group. Interestingly, the results for non-stars ("NS y-1") indicate that only *PAD_TOP* is significant, suggesting that these analysts travel a considerable amount of time, hoping to be ranked as star analysts for the first time.

4.2 Analysts' Forecast Accuracy

After establishing a strong link between *PAD* and the probability of being a star analyst, we turn to explore the relation between analyst hard and soft information collection efforts and forecast accuracy. We follow the same structure of Table 8, where we start with a linear relation between accuracy and *PAD*, and then focus on our set of *PAD* dummy variables, where we rank analysts based on their *PAD* values each year.

We employ the accuracy measure suggested by Clement (1999) and Jame, Johnston, Markov, and Wolfe (2016) and calculate the "Proportional Mean Absolute Forecast Error" ($PMAFE$) defined as $(AFE_{i,j,t} - \overline{AFE_{j,t}})$ / $\overline{AFE_{j,t}}$. In particular, for each analyst i and firm j, we calculate the analyst's quarterly equally-weighted forecast errors average based on all earnings forecasts initiated during the quarter. We then calculate the absolute value of the average forecasts errors. We repeat the calculation for all analysts on I/B/E/S covering the stock during that quarter and calculate the stock's quarterly mean absolute forecasts errors. The measure has a minimum is at -1 (most accurate relative to peers) and a maximum around 3 (the least accurate analyst). At zero, the analyst's accuracy is similar to that of its peers. The measure has a standard deviation of 0.53. In absolute terms ($|PMAFE|$) the measure has a mean of 0.39.

We run regressions at the analyst-quarter-stock level. The regressions include firm fixed effects, Coverage × Time fixed effects, and with or without analyst fixed effects. In addition, we control for various analyst and firm characteristics. In particular, we include how early the analyst forecast is relative to its peers (*Early Forecast*), past analyst accuracy (*Ave Q1 PMAFE t-4_t-1*), experience and seniority as captures by *IBES Years* and *High Rank Indicator*, number of quarterly forecasts and industries covered (*# Q1 EPS Forecasts*, and *# of GICS6 Industries*), firm size, firm book-to-market, return volatility, and institutional holdings.

Table 9 reports the results. We Z-score both the dependent and independent variables of interest. Panel A of Table 9 reports the results using *AWL* and *PAD*. *AWL* exhibits a significant negative relation with accuracy, suggesting that hither *AWL* is associated with improved accuracy. In terms of economic significance, a one standard deviation increase in *AWL* is associated with a reduction of about 2% in *PMAFE*'s standard deviation units. In contrast, while *PAD*'s coefficient estimates are negative, they are insignificant. This stands in contrast to all of our other tests where *PAD* is a significant factor in explaining analysts' behavior. Given the evidence on career concerns, we conjecture that this is driven by non-linearities in *PAD*, which we turn to explore using our *PAD* dummy variables.

Panel B of Table 9 starts with the relation between accuracy and *PAD_HIGH*. This immediately reveals that analysts with *PAD* above the median exhibit significant improvement in accuracy. During a pre-COVID period, *PAD_HIGH* is associated with a 3.3% improvement in accuracy in terms of its standard deviation units. The results are also economically significant when analyst fixed effects are included, but the statistical significance weakens. To further understand this relation in Panel C of Table 9, we explore *PAD_MED* and *PAD_TOP*. Strikingly. we find that most of the relation is concentrated in *PAD_MED*. That is, most of the improvement in accuracy is concentrated in the group of analysts that travel above the median but not in the top 20%. In particular, *PAD_MED* is associated with a 4.5% improvement in accuracy.

Overall, the results indicate that both hard and soft information seem to contribute to forecast accuracy. Moreover, the combined results of the star analysis and accuracy suggest that analysts in the top 20% are willing to sacrifice part of their accuracy at the expense of making connections and increasing their probability of becoming star analysts. This non-linearity explains why the relation between *PAD* and accuracy breaks down.

Finally, Fang and Hope (2021) show that equity research reports are often prepared by a team of analysts. We, as is standard in the analyst literature, focus on the lead analyst, who is recorded in I/B/E/S. Nevertheless, in Table A.6, we repeat the analysis conducted in Table 9 after controlling for team effort. In the baseline version, we measure team effort using the average *AWL* of peer analysts from the same brokerage firm covering the same industry. For about 9.8% of the lead analysts stock-quarter observations, the team members (signed on the report) are also in our sample, so we can measure their team effort using the average *AWL* of their actual team members for a given stock in a given quarter, resulting in an augmented team effort measure. We confirm that our results are robust to controlling for team effort.

5 Causal Evidence from the COVID Lockdown

The COVID-19 pandemic changed the work habits of many people. During the first two quarters of 2020, much of the country (and the world) was under stay-at-home mandates. Many in-person conferences, meetings, and other events were canceled. Our minute-by-minute Bloomberg online status data uniquely allows us to examine how sell-side equity analysts changed their work habits during that period. In addition, to the extent that the shocks to their work habits are largely exogenous, we can establish a causal relation when studying the resulting changes in the quantity and quality of their outputs.

For this section, we focus on the period 2019Q3-2020Q2 and keep all analysts with 4 quarters of data. We match the analysts' names with records on FINRA BrokerCheck, LinkedIn, Facebook, and other sources. From their online profiles, we estimate personal characteristics such as age, gender, and whether they have young children.

Almost every analyst in our sample is registered with FINRA BrokerCheck. These records include the full name (including middle name as well as other names used) of each analyst, as well as their work histories and the locations of their branch offices. After we identify the full name and work history of each analyst, we manually search through the Mergent Intellect database, which includes address histories for hundreds of millions of people in the US. These address histories combined with the work/school histories in the FINRA and LinkedIn data allow us to uniquely identify individuals in the Mergent data, which ultimately helps us identify home addresses of almost every analyst in our data during our sample period.

We then calculate the typical commute time between home and work using Google Maps. Google Maps provides typical travel times between points at any hour of the day. We measure minimum travel times between home and work at 7:00 am on workdays. We keep the minimum time based on foot, car, public transport, and bicycle travel. Figure 5 illustrates how we collect this information using a fictitious home address (to preserve the anonymity of the analysts in our sample). These filters leave us with 102 identified analysts with full information. Of these 102 analysts, 87 are from the New York area, 7 are from San Francisco, 6 are from Houston, and 2 are from Chicago.

The soft information production channel was effectively shut down during much of 2020Q1-2020Q2. The COVID-lockdown made it harder for analysts to travel. Even if they could travel, there was little soft information they could extract from in-person interactions as most conferences and meetings had been moved online. Intuitively, this negative information shock should be larger for traveling analysts, who we can uniquely identify using their PAD pre-COVID. In what follows, we use the pre-COVID PAD to instrument the shock to soft information production during the COVID lockdown.

5.1 Pre-COVID *PAD* Identification Strategy and Analyst Accuracy

Table 10 examines the causal impact of *PAD* on forecast outcomes in a standard difference-in-difference setting. The treatment group consists of analysts with above-median *PAD* pre-COVID (2019Q3-2019Q4). The control group contains the remaining analysts who rarely traveled pre-COVID. Following the analysis conducted in Table 9, we also compare the analysts in the top quintile (i.e., *PAD_TOP*) and those between the 50^{th} and 80^{th} percentile (*PAD_MED*) as additional treatment groups. The POST dummy equals 1 for 2020Q1-2020Q2 and 0 for 2019Q3-2019Q4. The coefficient on the interaction term (*TREATMENT* × *POST*) identifies the impact of *PAD* on forecast outcomes. As in Table 9, we examine the relative forecast accuracy of analysts' quarterly forecasts as measured by *PMAFE*.

Panel A of Table 10 analyzes the groups of analysts above and below the median. Focusing on the treatment effect (*TREATMENT*), consistent with the full sample results in 9, traveling analysts forecasts are slightly more accurate (though not significant). Focusing on the post effect (*POST*), with all analysts locked down at home, the accuracy measure *PMAFE* is not significantly affected since it is a relative accuracy measure (which should not change over time on average). Finally, focusing on the interaction term (*TREATMENT* × *POST*), we find that the accuracy of the treatment group (relative to their peers) decreases significantly, as reflected in a significant increase in *PMAFE* of 11.7%. Column 5 shows that the effect is driven by firms whose headquarters are located at least 300 miles away, and thus, are more affected by travel restrictions. The result provides causal evidence that soft information extracted by traveling analysts increase forecast accuracy.

In Panel B of Table 10, we explore the *MID* and *TOP* groups. The *TREATMENT* × *POST* coefficients are positive and significant for both, where the difference is larger for the *TOP* group. We know that analysts in the *TOP* group care more about the II All-Star analyst status. Achieving such a status during a lockdown may require more effort from them. As a result, their accuracy suffers more. Finally, in Figure 6, we plot the coefficient estimates of the treatment and control groups for each period and their differences with confidence intervals.

5.2 Commute Time to Work Identification Strategy

We now turn our attention to *AWL*. Graph B of Figure 3 shows that the average analyst in our sample experiences a one hour increase in his *AWL* after the COVID lockdown. Unlike the reduction in *PAD* which is completely exogenous and beyond any analyst's control, the increase in *AWL* during the lockdown could reflect an analyst's conscientious choices, which may in turn affect their forecast outcomes.

In Panel A of Table 11, we run cross-sectional regressions of changes in *AWL* (from 2019Q3-2019Q4 and 2020Q1-2020Q2) on various analyst characteristics measured pre-COVID. Analyst characteristics include the pre-COVID analyst commute time, the analyst age, a female analyst indicator, an indicator for an analyst with kids under 18-years old, and a few other analyst characteristics reported in Panel B of Table 3 such as yeas in I/B/E/S, MBA degree, work experience, and analyst rank.

The average analyst age in the pre-COVID analyzed sample is 44, where the youngest analyst is 30 years old, and the oldest is 62 years old. The pre-COVID sample also includes 10 female analysts and 19 analysts with kids under 18 years old. Both Du (2023) and Li and Wang (2024) document that female analysts, especially those with young children are more negatively affected by the COVID lockdown. By observing their *AWL*s, we can precisely quantify the impact of analysts' personal characteristics on changing workday length.

Table 11 Panel A presents clear evidence that the only significant predictor of analysts' changing *AWL* during COVID lockdown is their commuting time pre-COVID. The result is very intuitive. COVID lockdown makes commuting to the office impossible, and analysts can spend the time saved from commuting on work. Indeed, Table 11 suggests that one hour saved from not commuting leads to a workday that is 1.3 to 1.4 hours longer. Such a strong and positive relation between pre-COVID commute time and change in *AWL* during the lockdown is evident in the decile bin scatter plot in Figure 7. Importantly, the commute time is measured pre-COVID and, therefore, cannot be affected by events during the COVID pandeimc, so it provides a nice instrument for the change in AWL during the lockdown.

Building on the relation between the COVID lockdown and commute-time-saved, in Table 11 Panel B we examine the causal impact of AWL on forecast outcomes in a difference-in-difference setting, very similar to that in Table 10. The treatment group ($TREATMENT$) consists of analysts with below-median commute time pre-COVID (2019Q3-2019Q4) who are predicted to have a higher increase in AWL during COVID lockdown. The control group contains the remaining analysts with above-median commute time pre-COVID. The post dummy ($POST$) equals 1 for 2020Q1-2020Q2 and 0 for 2019Q3-2019Q4. The coefficient on the interaction term ($TREATMENT \times POST$) identifies the impact of AWL on forecast outcomes.

The treatment effect is not significant, suggesting that commuting time does not affect forecast outcomes pre-COVID. The post effect again suggests a significant increase. $PMAFE$, being a relative forecast accuracy measure, does not change for an average analyst. Finally, focusing on the interaction term, we find that analysts with a long commute time pre-COVID experience an improvement in their accuracy. Specifically, their accuracy (relative to their peers) increases significantly, as reflected in a significant decrease in $PMAFE$ of 8.5%. This result provides causal evidence that a longer workday length increases both the quantity and quality of forecasts. Finally, as a placebo test, we repeat the analysis for far and near firms. This should not be relevant for AWL, which doesn't rely on soft information gathering. Consistent with this conjecture and in contrast to Table 10, we do not find any differences between firms whose headquarters are located far or near and analyst locations.

6 Conclusion

Despite the importance of equity analysts, we still know relatively little about how they spend their working hours. In this paper, we take advantage of their minute-by-minute Bloomberg usage data to quantify two dimensions of their work habits: their average workday length to measure hard information collection and processing; and the extent of their travels to measure their soft information acquisition. We find that hard and soft information collection improves analysts' output on several dimensions, including the likelihood of becoming star analysts and the accuracy of their earnings forecasts. Our measures also reveal novel insights on the trade-off among analysts' objectives. For example, we find that analysts who travel extensively are willing to sacrifice accuracy in order to increase their chance of becoming star analysts.

Our findings related to the COVID lockdown speak to the recent debate on the benefit and cost of working-from-home (WFH). At least in the case of equity analysts, we find WFH to increase effort provision by eliminating work commute, which in turn improves the quality of the forecasts. On the downside, WFH hurts soft information production based on decreased in-person interaction and reduces forecast accuracy.

More broadly, we uncover another hidden effort problem which is ubiquitous in economics. We are able to characterize analysts' information collection without changing their behavior, and link their effort to outcomes that can be objectively and precisely measured.

References

Ben-Rephael, A., B. I. Carlin, Z. Da, and R. D. Israelsen. 2023. Uncovering the hidden effort problem. *Journal of Finance* forthcoming.

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

Butler, A. W. 2008. Distance still matters: Evidence from municipal bond underwriting. *Review of Financial Studies* 21:763–84.

Cazier, R. A., J. L. McMullin, and J. S. Treu. 2021. Are lengthy and boilerplate risk factor disclosures inadequate? an examination of judicial and regulatory assessments of risk factor language. *The Accounting Review* 96:131–55.

Cheng, Q., F. Du, X. Wang, and Y. Wang. 2016. Seeing is believing: analysts' corporate site visits. *Review of Accounting Studies* 21:1245–86.

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

Coval, J. D., and T. J. Moskowitz. 1999. Home bias at home: Local equity preference in domestic portfolios. *Journal of Finance* 54:2045–73.

Du, M. 2023. Locked-in at home: The gender difference in analyst forecasts after the covid-19 school closures. *Journal of Accounting and Economics* 76.

Fang, B., and O.-K. Hope. 2021. Analyst teams. *Review of Accounting Studies* 26:425–67.

Garmaise, M. J., and T. J. Moskowitz. 2004. Confronting information asymmetries: Evidence from real estate markets. *Review of Financial Studies* 17:405–37.

Granja, J., G. Matvos, and A. Seru. 2017. Selling failed banks. *Journal of Finance* 72:1723–84.

Green, T. C., R. Jame, S. Markov, and M. Subasi. 2014. Access to management and the informativeness of analyst research. *Journal of Financial Economics* 114:239–55.

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

Han, B., D. Kong, and S. Liu. 2018. Do analysts gain an informational advantage by visiting listed companies? *Contemporary Accounting Research* 35:1843–67.

Harford, J., F. Jiang, R. Wang, and F. Xie. 2019. Analyst career concerns, effort allocation, and firms' information environment. *The Review of Financial Studies* 32:2179–224.

Hope, O.-K., D. Hu, and H. Lu. 2016. The benefits of specific risk-factor disclosures. *Review of Accounting Studies* 21:1005–45.

Israelsen, R. D., M. Schwartz-Ziv, and J. Weston. 2024. Block diversity and governance. *Review of Corporate Finance Studies* forthcoming.

Ivkovic, Z., and S. Weisbenner. 2005. Local does as local is: Information content of the geography of individual investors' common stock investments. *Journal of Finance* 60:267–306.

Jame, R., R. Johnston, S. Markov, and M. C. Wolfe. 2016. The value of crowdsourced earnings forecasts. *Journal of Accounting Research* 54:1077–110.

Lerner, J. 1995. Venture capitalists and the oversight of private firms. *Journal of Finance* 50:301–18.

Li, F. W., and B. Wang. 2024. The gender effects of covid-19 on equity analysts. *Review of Accounting Studies* forthcoming.

Liberti, J. M., and M. A. Petersen. 2019. Information: Hard and soft. *Review of Corporate Financial Studies* 8:1–41.

Loughran, T., and P. Schultz. 2005. Liquidity: Urban versus rural firms. *Journal of Financial Economics* 78:341–74.

Malloy, C. 2005. The geography of equity analysis. *Journal of Finance* 60:719–65.

Soltes, E. 2014. Private interaction between firm management and sell-side analysts. *Journal of Accounting Research* 52:245–72.

Figure 1: Average Workday Length Example

This figure provides an example of the AWL measure for an analyst-quarter observation. The blue bars represent the empirical probability density function based on activity on Bloomberg. The red curve is the estimated Gaussian Mixture Model pdf using the iterative Expectation-Maximization (EM) algorithm. The two orange curves are the two underlying Gaussian pdfs. The dashed vertical bars are the estimated means of the two distributions. The two black lines represent the beginning and end of the AWL measure, or the interval $(\mu_1 - \sigma_1, \mu_2 + \sigma_2)$.

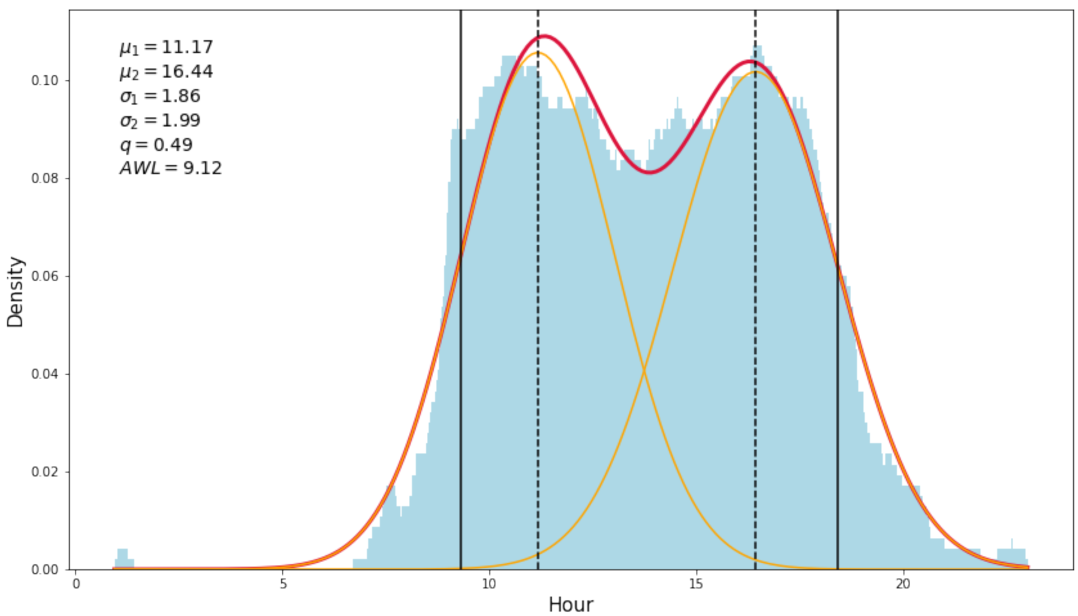


Figure 2: Minutes Active on Terminal based on Day-of-the-Week and Holidays

This figure depicts the average time spent on the Bloomberg terminal by day-of-the-week and Holidays. The sample period is from September 2017 to March 2021.

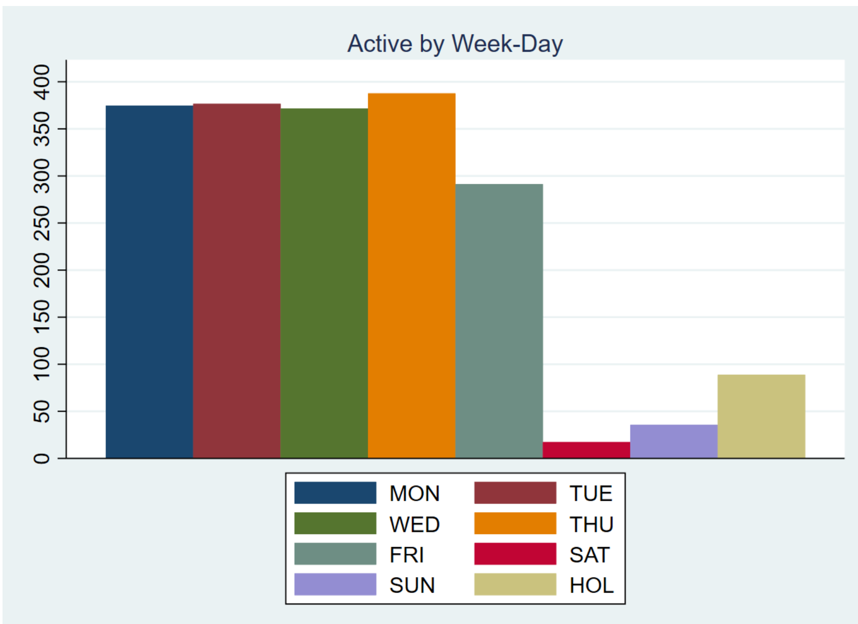
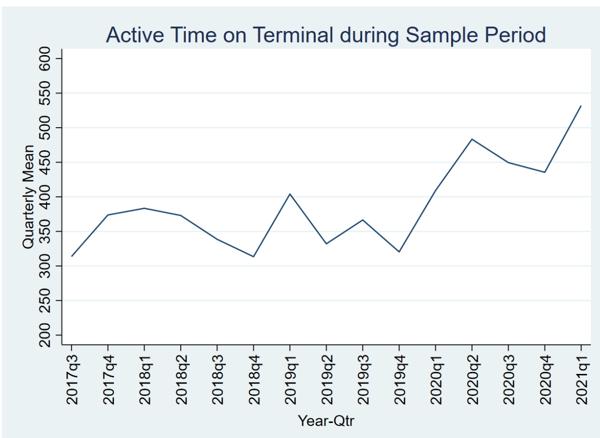


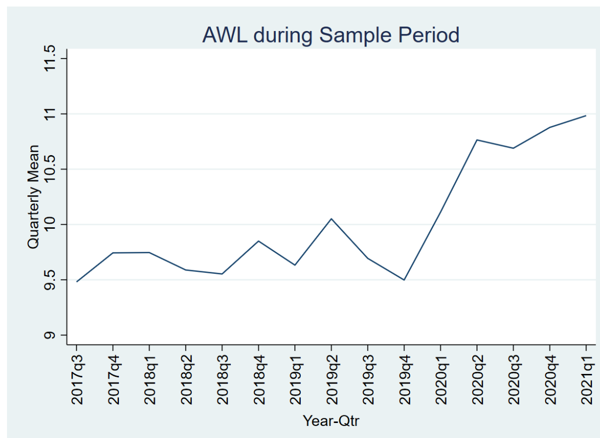
Figure 3: *AWL*, *Minutes Active*, and *PAD* during Sample Period

This figure depicts the quarterly cross-analyst averages of the various log-in measures over the sample period. The measures are: *Minutes Active*, *AWL*, and *PAD*. See Table A.1 and Table 1 for details about variable and sample definitions. The sample period is from September 2017 to March 2021.

Panel A: *Minutes Active*



Panel B: *AWL*



Panel C: *PAD*

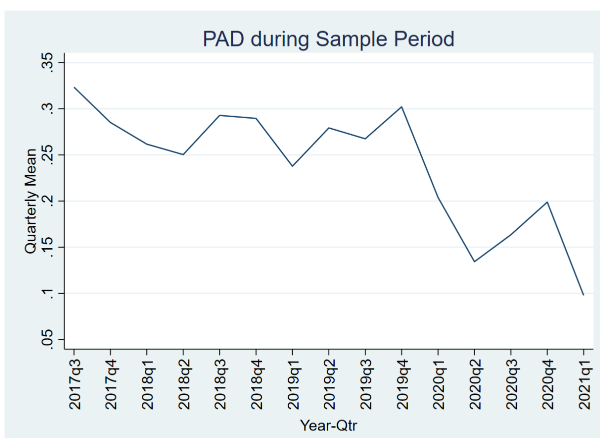
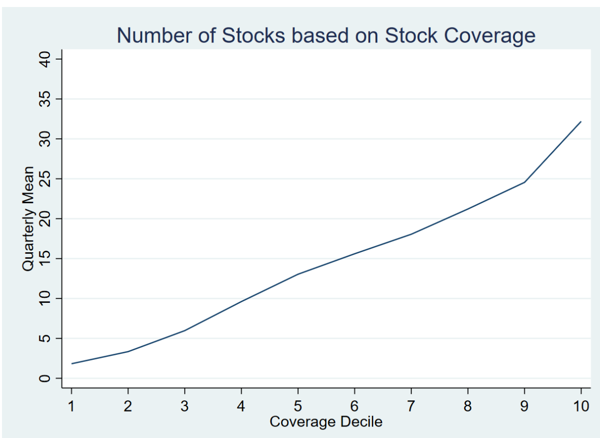


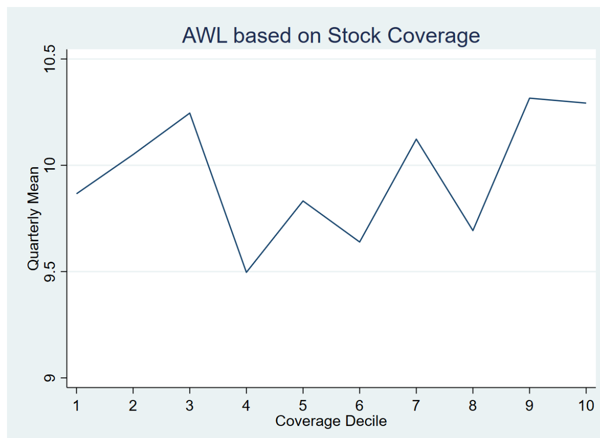
Figure 4: Stocks, *AWL*, *Minutes Active*, and *PAD* based on Coverage

This figure provides statistics based on stock-coverage deciles. The sample period is from September 2017 to March 2021. Each year and quarter, we rank all analysts in our sample into deciles based on the number of stocks they cover over the previous 4 quarters. Graph A plots the average number of stocks covered per decile. Graph B plots the average *AWL*. Graph C plots the average time on Bloomberg terminal conditioning on days with terminal activity ("Conditional Active"), and Graph D plots the average *PAD*.

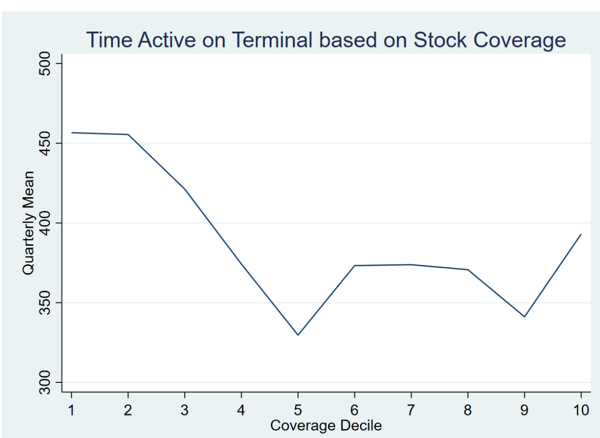
Panel A: Number of Stocks



Panel B: *AWL*



Panel C: Conditional Active Time on Terminal



Panel D: *PAD*

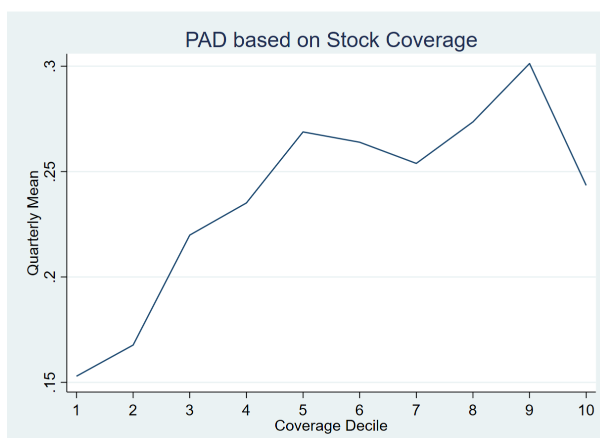


Figure 5: Measuring Commute Time - Example

This figure provides a fictitious (to preserve anonymity) example of how we measure commute time for a given analyst. Using Google Maps, we measure the minimum typical travel time between home and work at 7:00 am on a workday. The figure illustrates this for public transit – in this case, 23 minutes – but we collect the same information for automobile, bicycle, and foot travel. Commute time is then the minimum travel time across these various options. We verify the home address and work address of the analysts using data from FINRA BrokerCheck, Mergent Intellect, and LinkedIn.

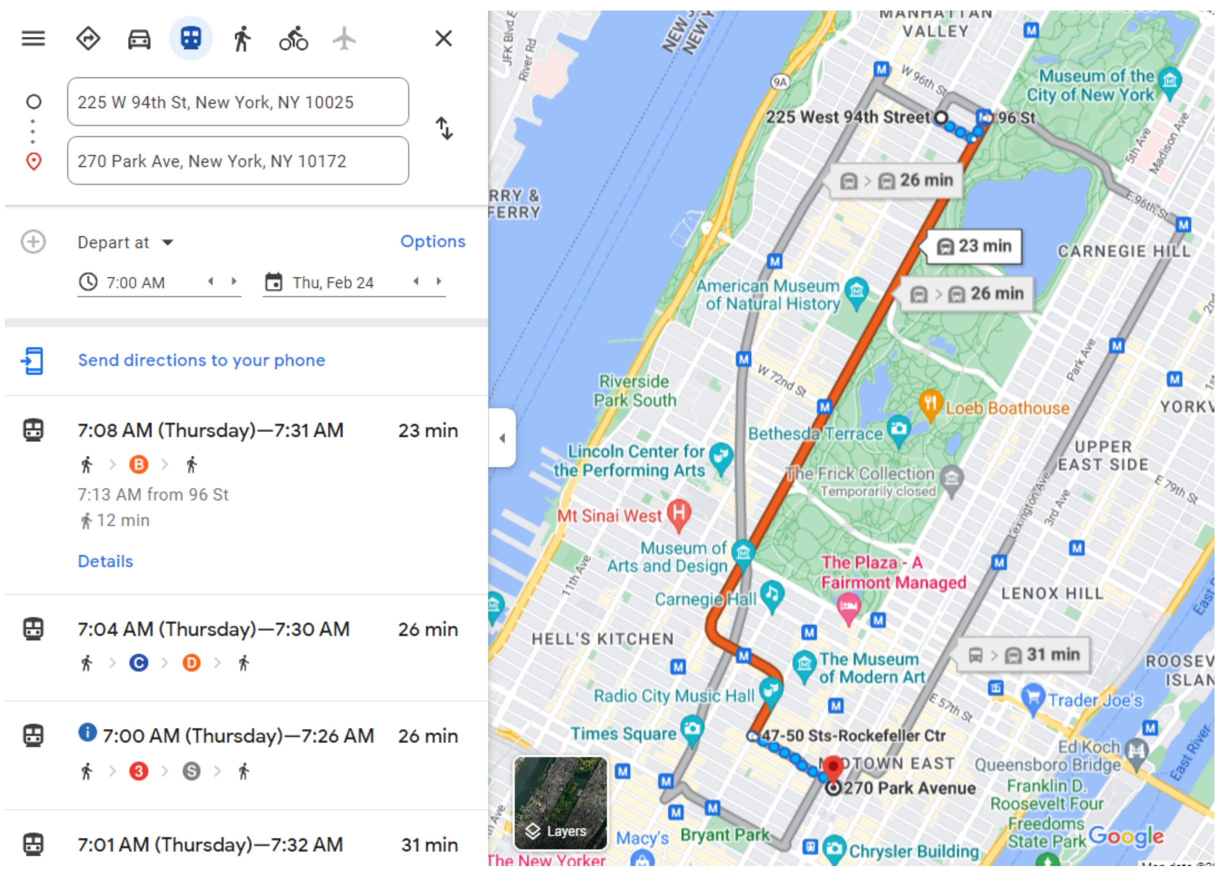
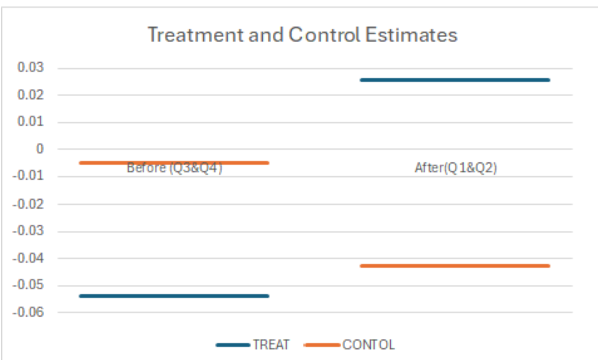


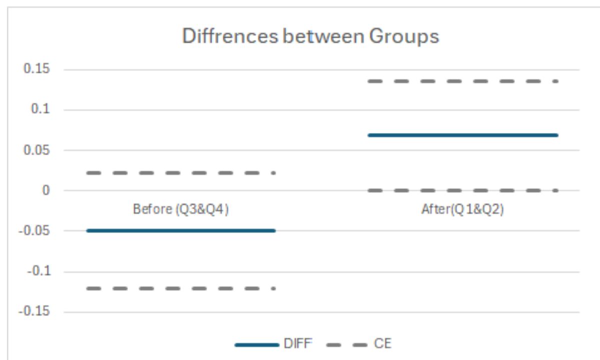
Figure 6: DID Estimates and Differences

This figure provides the estimates of the treatment and control groups and their differences from Table 10 identification strategy. Graphs A and B report results for the above and below median groups. Graphs C and D report results for the *PAD_TOP* and below median groups.

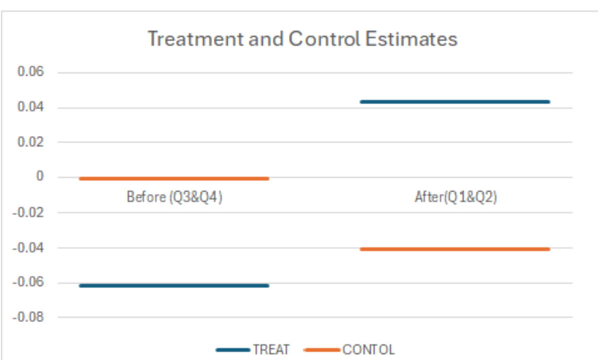
Panel A: Medians - Treatment and Control Estimates



Panel B: Medians - Treatment and Control Differences



Panel C: *PAD_TOP* and Bottom50 - Treatment and Control Estimates



Panel D: *PAD_TOP* and Bottom50 - Treatment and Control Differences

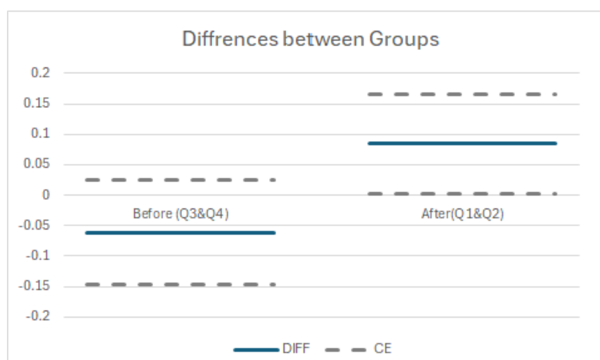


Figure 7: Changes in AWL and Commute Time Saved

This figure illustrates the relation between AWL and commute-time-saved reported in Table 11, where changes in AWL (Q1-Q2 of 2020 minus Q3-Q4 of 2019) are plotted against commute-time-saved deciles. The x-axis reports the average commute time saved for each decile, whereas the y-axis reports the corresponding average change in AWL.

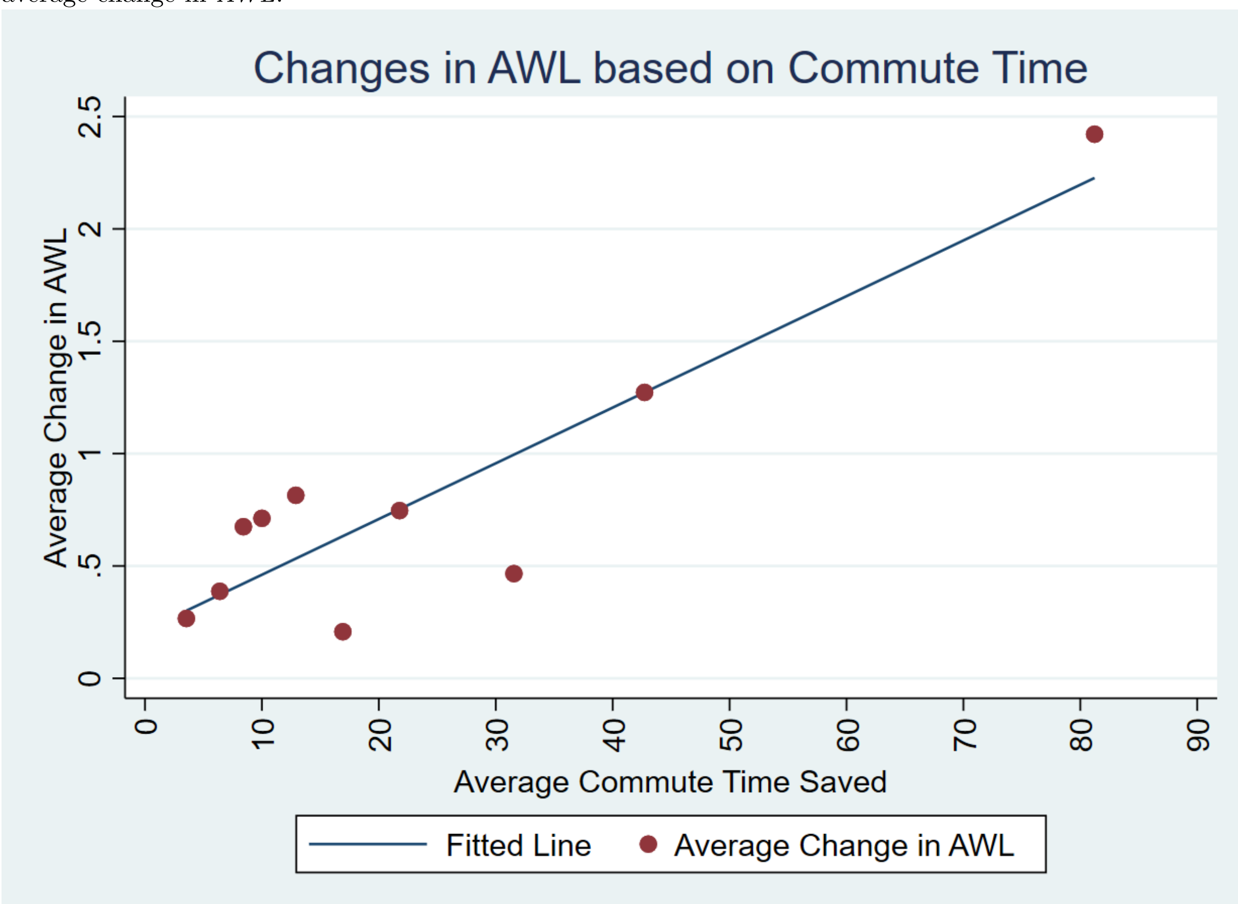


Table 1: Summary stats of analyst output

This table reports summary statistics of analyst output for the sample of Active Bloomberg analysts analyzed in this study (Bloomberg sample) and their comparison sample. The active analysts' sample includes 336 analysts and 42 brokerage firms, with over 2,874 analyst-quarter observations. To be included in the comparison sample, we require an analyst to cover at least three stocks, to be on I/B/E/S for at least four quarters, and to belong to one of the 42 brokerage firms in our Bloomberg sample. The comparison sample includes 1,854 analysts over 16,239 analyst-quarter observations. See Table A.1 for details about variable definitions. The sample period is from September 2017 to March 2021. To be considered as an active Bloomberg user, an analyst needs to have at least one quarter with a quarterly average percent activity greater than 3%. Percent activity is the time in minutes that an analyst is actively logged to the terminal scaled by the number of minutes within a day. This cut-off removes the left tail of the log-in distribution, which is populated by inactive users. In addition, we require an analyst to have at least two earnings forecasts per quarter, and to cover at least 3 stocks. Panel A reports the mean, median, standard deviation, and other percentiles of the Bloomberg sample. Panel B compares the Bloomberg sample with the comparison sample. We report each group's averages, their differences, and associated p-values. Standard errors are clustered by analyst and year-quarter.

Panel A: The Bloomberg Sample Summary Statistics

	Mean	Std. Dev.	10%	25%	Median	75%	90%
# Unique Stocks t-4_t-1	17.848	10.529	4.000	10.000	17.000	25.000	31.000
Ave # Stocks t-4_t-1	15.696	9.384	3.000	7.500	15.500	22.250	27.000
# of GICS6 Industries	2.999	1.969	1.000	2.000	2.000	4.000	6.000
# of Stocks w Q1 EPS Forecasts	16.068	9.354	4.000	8.000	16.000	22.000	28.000
% of Common Stocks	77.070	27.997	28.125	69.231	88.000	96.154	100.000
# Q1 EPS Forecasts	23.079	16.194	5.000	10.000	21.000	32.000	43.000
# Y1 EPS Forecast	24.785	17.414	5.000	11.000	22.000	35.000	47.000
# Long Term Growth Forecasts	5.673	11.281	0.000	0.000	0.000	6.000	20.000
# of Other Forecasts	140.124	133.086	19.000	45.000	101.000	193.000	305.000
# of Stocks w Rec	3.276	3.269	1.000	1.000	2.000	4.000	7.000
# of Rec	2.468	3.343	0.000	0.000	2.000	3.000	6.000
# of non-stale Rec	2.225	3.025	0.000	0.000	1.000	3.000	5.000
# of Stocks w PTG	11.805	7.940	2.000	5.000	11.000	17.000	23.000
# of PTG	15.275	14.429	0.000	4.000	12.000	23.000	34.000
# of Analyst-Quarters	2,874						

Panel B: Mean Differences of the Bloomberg Sample and their Comparison Group

	Bloomberg	Comparison	Mean-Diff	P-value
# Unique Stocks t-4_t-1	17.848	15.7486	2.099	0.011
Ave # Stocks t-4_t-1	15.696	13.7563	1.940	0.008
# of GICS6 Industries	2.999	3.13178	-0.133	0.316
# of Stocks w Q1 EPS Forecasts	16.068	14.359	1.709	0.015
% of Common Stocks	77.07	69.2383	7.832	0.001
# Q1 EPS Forecasts	23.079	21.327	1.752	0.098
# Y1 EPS Forecast	24.785	21.1604	3.625	0.004
# Long Term Growth Forecasts	5.673	1.83447	3.839	0.000
# of Other Forecasts	140.124	125.927	14.197	0.105
# of Stocks w Rec	3.276	2.92485	0.351	0.024
# of Rec	2.468	2.03171	0.436	0.007
# of non-stale Rec	2.225	1.77345	0.452	0.003
# of Stocks w PTG	11.805	10.5826	1.222	0.029
# of PTG	15.275	13.9109	1.364	0.200
AveQtrAccuracy	-0.030	-0.017	-0.012	0.045
AveQtrAccuracy_VW	-0.025	-0.019	-0.006	0.322
# of Analysts	336	1,854		
# of Analyst-Quarters	2,874	16,239		

Table 2: Summary stats of analyst Bloomberg log-in activity and *AWL* measures

This table reports summary statistics of analysts' log-in activity on the Bloomberg terminal (Panel A), together with the log-in based measures (Panel B), and their correlation martix (Panel C). See Table A.1 and Table 1 for details about variable and sample definitions.

Panel A: Log-in Statistics

	Mean	Std. Dev.	10%	25%	Median	75%	90%
% of Workdays with Bloomberg Activity	0.717	0.246	0.344	0.611	0.786	0.902	0.967
Active (minutes per day)	361.711	198.075	87.190	235.902	362.169	477.891	588.000
Conditional Active (on active days)	475.638	188.910	285.829	382.333	472.765	552.520	650.085
Active - hours per Week	30.143	16.506	7.266	19.658	30.181	39.824	49.000
# of Analyst-Quarters	2,874						

Panel B: *AWL* and *PAD* statistics

	Mean	Std. Dev.	10%	25%	Median	75%	90%
ALL							
AWL	9.805	2.028	7.966	8.830	9.732	10.873	12.074
PAD	0.283	0.246	0.033	0.098	0.214	0.389	0.656
Pre-COVID							
AWL	9.532	1.913	7.840	8.662	9.421	10.462	11.678
PAD	0.316	0.234	0.067	0.145	0.256	0.419	0.654
COVID							
AWL	10.461	2.142	8.527	9.421	10.480	11.586	12.763
PAD	0.205	0.254	0.016	0.033	0.100	0.246	0.667
# of Analyst-Quarters	2,874	2,029	845				

Panel C: Correlation matrix

	(1)	(2)	(3)
(1) *AWL*	1.00		
(2) *PAD*	-0.23	1.00	
(3) *LnActive*	0.25	-0.37	1.00

Table 3: *AWL* and *PAD* explained by Fixed-Effect and Analyst Characteristic

This table reports results from panel regressions of *AWL* and *PAD* on various fixed effects and analyst characteristics. Panel A reports the explained variation of our *AWL* and *PAD* measures by time, analyst, brokerage firm, and main GICS6 industry using fixed effect regressions. Panel B regresses the *AWL* and *PAD* measures on analyst characteristics obtained from various sources. In Panel B, the standard errors are clustered by analysts reported in parentheses below the coefficient estimates See Table A.1 and Table 1 for details about variable and sample definitions. The sample period is from September 2017 to March 2021. We keep analyst-quarter observations that meet the required quarterly login activity filter. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

Panel A: *AWL* and *PAD* Variation Explained by Fixed Effects

	AWL				PAD			
	(1) TIME	(2) ANALYST	(3) BROKER	(4) INDUSTRY	(5) TIME	(6) ANALYST	(7) BROKER	(8) INDUSTRY
Constant	9.346***	10.940***	10.797***	10.069***	0.335***	0.145	0.801***	0.263***
	(68.06)	(12.40)	(12.43)	(65.05)	(21.37)	(1.53)	(8.05)	(14.40)
$Adj.R^2$	0.051	0.432	0.082	0.086	0.090	0.515	0.115	0.067
Observations	2,874	2,874	2,874	2,872	2,874	2,874	2,874	2,872

Panel B: *AWL*, *PAD* and Analyst Characteristics

	AWL				*PAD*			
	(1) *AWL*	(2) *AWL*	(3) *AWL*	(4) *AWL*	(5) *PAD*	(6) *PAD*	(7) *PAD*	(8) *PAD*
IBES Years	-0.044***	-0.040***	-0.044***	-0.034*	0.002	0.002	0.002	0.002
	(-2.85)	(-2.63)	(-2.84)	(-1.95)	(1.21)	(1.17)	(1.27)	(1.31)
High Rank Indicator	-0.460**	-0.495***	-0.538***	-0.394**	0.051***	0.051***	0.054***	0.028
	(-2.55)	(-2.80)	(-3.09)	(-2.23)	(2.84)	(2.79)	(2.93)	(1.47)
STAR	0.219	0.118	0.134	-0.165	0.036**	0.038**	0.037**	0.062***
	(1.24)	(0.65)	(0.74)	(-0.81)	(2.26)	(2.32)	(2.28)	(3.60)
Work Experience	0.007	-0.000	-0.000	-0.012	-0.000	0.000	0.000	0.001
	(0.42)	(-0.02)	(-0.03)	(-0.69)	(-0.08)	(0.10)	(0.10)	(0.58)
# Jobs FINRA	-0.025	-0.025	-0.035	-0.048	0.004	0.004	0.005	0.008
	(-0.54)	(-0.55)	(-0.76)	(-0.92)	(0.82)	(0.86)	(0.96)	(1.60)
Ave Q1 PMAFE t-4_t-1	0.045	0.059	0.081	-0.083	0.020	0.019	0.018	0.020
	(0.09)	(0.13)	(0.17)	(-0.19)	(0.45)	(0.42)	(0.39)	(0.47)
NYC Indicator		0.313*	0.348*	0.191		-0.004	-0.006	-0.018
		(1.73)	(1.96)	(0.86)		(-0.20)	(-0.32)	(-0.68)
MBA Indicator		0.272	0.304	0.554		-0.023	-0.025	-0.058**
		(0.56)	(0.63)	(1.23)		(-0.57)	(-0.63)	(-2.08)
Female Indicator		0.070	0.080	-0.040		0.007	0.006	0.006
		(0.31)	(0.35)	(-0.17)		(0.31)	(0.28)	(0.28)
Children Indicator		0.373	0.392	0.144		-0.077	-0.078	-0.065
		(0.72)	(0.75)	(0.26)		(-1.32)	(-1.35)	(-0.99)
Principal Exam			0.383*	0.189			-0.025	-0.033
			(1.68)	(0.76)			(-1.06)	(-1.31)
Coverage x Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Brokerage Firm FE	NO	NO	NO	YES	NO	NO	NO	YES
Analyst Cluster	YES	YES	YES	YES	YES	YES	YES	YES
Observations	2,493	2,493	2,493	2,491	2,493	2,493	2,493	2,491
R^2	0.194	0.212	0.217	0.268	0.295	0.297	0.298	0.411

Table 4: Analysts Pre-Open Daily Abnormal Login Activity

This table reports results from daily panel regressions of analysts' abnormal login activity from 7 am to 9 am on various market and information events variables. Specifically, for each analyst and half an hour during 7-9 am, we have an indicator that is equal to one if an analyst is logged in to the Bloomberg terminal. To capture an analyst's abnormal login activity, for each day and half an hour interval, we remove the analyst's day-interval average sample activity. This is comparable to including day and interval fixed effects in a regression. We then calculate the de-trended averages during the pre-open period. We further construct a battery of analyst-specific explanatory variables based on the set of stocks that an analyst covers in her portfolio during a given year-quarter. These variables include extreme market activity and news coverage. See Table A.1 and Table 1 for details about variable and sample definitions. The sample period is from September 2017 to March 2021. We keep analyst-quarter observations that meet the required quarterly login activity filter. Standard errors are double clustered by analyst and date reported in parentheses below the coefficient estimates. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

	Analysts Average LogIn Activity During 7-9 AM					
	(1)	(2)	(3)	(4)	(5)	(6)
# Stocks in AbnVOl Decile t-1	0.007***	0.007***	0.006***	0.006***	0.005***	0.005***
	(6.44)	(6.45)	(5.90)	(5.88)	(5.19)	(5.21)
# Stocks in AbsExtRet Decile t-1	0.001	0.001	0.001	0.001	0.001	0.001
	(1.03)	(1.04)	(0.82)	(0.82)	(0.95)	(0.96)
# Stock with AMC News t-1	0.005***			0.004***	0.002	
	(3.33)			(2.65)	(1.33)	
# Stock with AMC Earn News t-1		0.008***				0.001
		(2.80)				(0.22)
# Stock with AMC AR News t-1		-0.013				-0.012
		(-1.52)				(-1.30)
# Stock with BMO News t			0.013***	0.013***		
			(9.05)	(9.05)		
# Stock with BMO Earn News t					0.079***	0.079***
					(12.31)	(12.33)
# Stock with BMO AR News t					0.004***	0.004***
					(3.15)	(3.16)
# Max Industry Earn BMO News Pressure t					0.074***	0.074***
					(3.66)	(3.66)
Analyst FE	YES	YES	YES	YES	YES	YES
Date FE	YES	YES	YES	YES	YES	YES
Coverage FE	YES	YES	YES	YES	YES	YES
Date Cluster	YES	YES	YES	YES	YES	YES
Analyst Cluster	YES	YES	YES	YES	YES	YES
Observations	141,472	141,472	141,472	141,472	141,472	141,472
R^2	0.138	0.138	0.140	0.140	0.149	0.149

Table 5: *AWL*, *PAD*, and Subsequent Earnings Calls Participation

This table reports results from quarterly panel regressions of measures that capture different features of analysts' discussions in subsequent quarter earnings calls on *AWL*, and *PAD*. We apply machine learning algorithms to all earnings conference call transcripts in our sample to systematically capture various aspects of analyst participation. Our measures are at the analyst quarter level. The first measure,*% Named Entities*, uses NLP algorithms to identify name entities that capture the specificity of analysts' discussions. Named entities include events, locations, organizations, people, products, etc., that capture specific information. We remove from the list named entities that capture pure quantities (cardinal, percent, quantity, and money). The measure is calculated for each analyst and quarter as the percentage of non-quantity named entities words to total words across all earnings calls for the analyst (Panel A). The second measure, *% of Quantitative Sentences*, uses ChatGPT to identify analyst sentences that are classified as quantitative. The measure is calculated as the percentage of quantitative sentences to the total sentences across all earnings calls for the given analyst (Panels B. "ALL' refers to the full sample. "Pre-COVID" refers to 2017-2019. (Z) stands for a Z-Score adjustment (a mean of zero and a standard deviation of one). We Z-Score adjust both the dependent variable and independent variables of interest. See Table A.1 and Table 1 for details about variable and sample definitions. Control variables include: *IBES Years*, *High Rank Indicator*, *Ave # Q1 EPS Forecasts t-4_t-1*, *Ave # of Industries t-4_t-1*, *LnAveSize*, *LnAveBM*, *AveInstHold*, and *AveStdDev.Ret*. The sample period is from September 2017 to March 2021. We keep analyst-quarter observations that meet the required quarterly login activity filter. Standard errors are clustered by analysts reported in parentheses below the coefficient estimates. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

Panel A: Percentage of Named Entities

	ALL	Pre-COVID	ALL	Pre-COVID
	(1)	(2)	(3)	(4)
AWL(Z)	0.034	0.032	0.014	0.007
	(1.50)	(1.04)	(0.46)	(0.14)
PAD(Z)	0.054*	0.084**	0.084**	0.136***
	(1.95)	(2.40)	(2.23)	(2.90)
Lag % Named Entities (Z)	0.457***	0.402***	-0.090***	-0.143***
	(11.64)	(9.98)	(-2.82)	(-4.12)
# Earn Calls	0.020**	0.029**	0.004	0.019
	(2.46)	(2.56)	(0.49)	(1.41)
Controls	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES
Analyst FE	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES
Observations	2,212	1,558	2,179	1,519
R^2	0.347	0.294	0.616	0.599

Panel B: Percentage of Quantitative Sentences

	ALL	Pre-COVID	ALL	Pre-COVID
	(1)	(2)	(3)	(4)
AWL(Z)	0.052**	0.044*	0.021	-0.014
	(2.38)	(1.77)	(0.66)	(-0.30)
PAD(Z)	0.035	0.009	-0.006	-0.023
	(1.30)	(0.25)	(-0.21)	(-0.56)
Lag % Quant(Z)	0.529***	0.522***	-0.021	-0.033
	(14.75)	(12.49)	(-0.73)	(-0.89)
# Earn Calls	0.005	0.005	0.003	0.010
	(1.19)	(0.96)	(0.52)	(1.32)
Controls	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES
Analyst FE	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES
Observations	2,212	1,558	2,179	1,519
R^2	0.358	0.347	0.609	0.605

Table 6: *AWL*, *PAD*, and Stock Coverage Decisions

This table reports results from quarterly panel regressions of analysts' average firm characteristic portfolio ranking on *AWL*, and *PAD*. Each quarter, we rank all the stocks in our sample into quintiles based on selected firm characteristics. Then for each analyst and quarter, we calculate the stock market cap weighted average of each ranking across all the stocks covered by the analyst. AGE is the firm number of years on CRSP, SIZE is the stock market cap, PRC is the stock price, ILLIQ is the stock AMIHUD illiquidity measure, BM is the stock book-to-market ratio, MOM is the stock return momentum over the past 12 months, IDIOVOL is the stock daily idiosyncratic volatility measured over the past 252 days, and SKEW is the stock daily skewness measured over the past 252 days. (Z) stands for a Z-Score adjustment (a mean of zero and a standard deviation of one). We Z-Score adjust both the dependent variable and independent variables of interest. See Table A.1 and Table 1 for details about variable and sample definitions. The sample period is from September 2017 to March 2021. We keep analyst-quarter observations that meet the required quarterly login activity filter. Standard errors are clustered by analysts reported in parentheses below the coefficient estimates. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

	AGE	SIZE	PRC	ILLIQ	BM	MOM	IDIOVOL	SKEW
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AWL(Z)	0.089*	0.085**	0.156***	-0.092**	-0.018	0.028	-0.072*	-0.005
	(1.82)	(2.06)	(3.12)	(-2.30)	(-0.33)	(0.78)	(-1.74)	(-0.17)
PAD(Z)	0.006	0.088**	0.071	-0.074*	-0.087*	0.113***	-0.033	0.035
	(0.13)	(2.06)	(1.51)	(-1.69)	(-1.79)	(3.03)	(-0.73)	(1.06)
IBES Years	0.035***	0.050***	0.027***	-0.053***	-0.001	-0.001	-0.042***	-0.007
	(3.93)	(5.84)	(3.08)	(-6.04)	(-0.05)	(-0.10)	(-4.66)	(-1.07)
High Rank Indicator	0.044	0.051	0.094	-0.053	-0.126	0.086	-0.077	0.017
	(0.37)	(0.46)	(0.80)	(-0.48)	(-0.94)	(0.99)	(-0.64)	(0.24)
Ave # of Industries t-4_t-1	-0.061**	-0.003	0.080***	-0.006	-0.164***	0.168***	-0.064**	0.032*
	(-2.06)	(-0.12)	(3.09)	(-0.24)	(-5.70)	(7.51)	(-2.09)	(1.78)
Ave Q1 PMAFE t-4_t-1	-0.056	-0.308	-0.787**	0.837**	-0.203	-0.276	0.694**	-0.143
	(-0.14)	(-0.79)	(-2.16)	(2.17)	(-0.51)	(-0.93)	(2.11)	(-0.39)
Coverage x Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Analyst Cluster	YES	YES	YES	YES	YES	YES	YES	YES
Observations	2,540	2,540	2,540	2,540	2,295	2,537	2,538	2,538
R^2	0.148	0.258	0.156	0.273	0.135	0.146	0.144	0.081

Table 7: *AWL*, *PAD*, and Subsequent Output Decisions

This table reports results from quarterly panel regressions of analyst subsequent quarter output decisions on *AWL*, and *PAD*. Specifically, we focus on the ratio between "soft" and "hard"" output, measured for each analyst and quarter as the ratio of all stock recommendations to all quarterly earnings forecasts and price target forecasts. "ALL' refers to the full sample. "Pre-COVID" refers to 2017-2019. (Z) stands for a Z-Score adjustment (a mean of zero and a standard deviation of one). We Z-Score adjust both the dependent variable and independent variables of interest. See Table A.1 and Table 1 for details about variable and sample definitions. The sample period is from September 2017 to March 2021. We keep analyst-quarter observations that meet the required quarterly login activity filter. Standard errors are clustered by analysts reported in parentheses below the coefficient estimates. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

	ALL	Pre-COVID	ALL	Pre-COVID
	(1)	(2)	(3)	(4)
AWL(Z)	-0.045*	-0.026	-0.054*	-0.086**
	(-1.70)	(-0.95)	(-1.82)	(-2.43)
PAD(Z)	-0.023	-0.025	0.055*	0.076*
	(-0.85)	(-0.82)	(1.66)	(1.89)
# Q1 EPS Forecasts	-0.010***	-0.012***	-0.006*	-0.003
	(-4.54)	(-4.83)	(-1.96)	(-0.83)
IBES Years	-0.004	-0.006	0.574*	2.296*
	(-0.88)	(-1.35)	(1.91)	(1.67)
High Rank Indicator	0.077	0.068		
	(1.08)	(0.89)		
Ave Q1 PMAFE t-4_t-1	0.135	0.247	-0.143	-0.154
	(0.59)	(0.99)	(-0.46)	(-0.38)
Ave # of Industries t-4_t-1	0.008	0.011	-0.050	-0.021
	(0.45)	(0.64)	(-1.03)	(-0.35)
Coverage x Time FE	YES	YES	YES	YES
Analyst FE	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES
Observations	2,234	1,651	2,215	1,619
R^2	0.115	0.099	0.343	0.347

Table 8: Probability of Being a Star Analyst

This table reports results from panel regressions of a star analyst indicator on *AWL*,*PAD*, and ranked-based *PAD* measures controlling for various fixed effects and analyst characteristics. We employ a linear probability model where a dummy variable of being a star analyst in Q4 of year t is regressed on average *AWL*, average *PAD*. To explore the potential non-linearity of *PAD*, we include in the analysis dummy variables based on the distribution of average *PAD* in Q1-Q3 of year t. In particular, *PAD_HIGH* is a dummy variable that receives the value of one if the average of *PAD* during Q1-Q3 during year t is above the median of average *PAD* distribution, and zero otherwise. In a similar manner, *PAD_MED* (*PAD_TOP*) is a dummy variable that receives the value of one if the average of *PAD* during Q1-Q3 is between the 50^{th} and 80^{th} percentiles (above the 80^{th} percentile) of the distribution in a given year, and zero otherwise. Since we use the averages of *PAD* and *AWL* during Q1-Q3, we limit our analysis to 2018-2020, where we have full information. Columns 1—4 (5–8)include all observations (focus on the Pre-VOVID period). The even columns focus on a sub sample where the analyst was not elected as a star analyst in the previous year (NS y-1). Standard errors are clustered by analyst reported in parentheses below the coefficient estimates. See Table A.1 and Table 1 for details about variable and sample definitions. "ALL' refers to the full sample. "Pre-COVID" refers to 2017-2019. (Z) stands for a Z-Score adjustment (a mean of zero and a standard deviation of one). We Z-Score adjust the indipendent variables of interest. We keep analyst-quarter observations that meet the required quarterly login activity filter. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

Panel A: *AWL* and *PAD*

	ALL				Pre-COVID			
	ALL	NS y-1	ALL	NS y-1	ALL	NS y-1	ALL	NS y-1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ave AWL Q1-Q3(Z)	0.042*	0.013	0.009	-0.007	0.051*	0.026	0.011	0.000
	(1.75)	(0.77)	(0.42)	(-0.37)	(1.94)	(1.35)	(0.42)	(0.00)
Ave PAD Q1-Q3(Z)	0.067**	0.043*	0.087***	0.063***	0.085***	0.059**	0.102***	0.081**
	(2.44)	(1.97)	(3.44)	(2.72)	(2.69)	(2.22)	(3.10)	(2.46)
Ave Q1 PMAFE t-4_t-1	-0.102	-0.006	-0.057	0.006	-0.160	-0.103	-0.070	0.006
	(-0.72)	(-0.06)	(-0.40)	(0.06)	(-0.92)	(-0.82)	(-0.42)	(0.05)
IBES Years	0.023***	0.005	0.020***	0.009**	0.021***	0.007*	0.021***	0.010*
	(4.96)	(1.30)	(4.78)	(2.12)	(4.37)	(1.67)	(4.20)	(1.91)
High Rank Indicator	0.076	0.049	0.150***	0.105	0.104	0.089	0.168***	0.145*
	(1.29)	(0.96)	(2.82)	(1.65)	(1.64)	(1.55)	(2.74)	(1.96)
Work Experience	-0.003	-0.005	-0.007	-0.008*	-0.006	-0.006	-0.010	-0.009**
	(-0.54)	(-1.27)	(-1.31)	(-1.96)	(-0.95)	(-1.52)	(-1.61)	(-2.03)
# Jobs FINRA	-0.036***	-0.004	-0.038***	-0.006	-0.030**	-0.003	-0.038**	-0.003
	(-2.67)	(-0.39)	(-2.90)	(-0.64)	(-2.06)	(-0.29)	(-2.54)	(-0.28)
NYC Indicator	0.154***	0.056	0.028	0.009	0.190***	0.089**	0.077	0.062
	(2.66)	(1.49)	(0.42)	(0.15)	(3.06)	(2.30)	(1.06)	(1.10)
MBA Indicator	0.139	0.041	0.133	0.041	0.147	0.087	0.128	0.089
	(1.24)	(0.45)	(1.19)	(0.46)	(1.35)	(0.84)	(1.12)	(0.92)
Female Indicator	-0.005	-0.026	-0.038	-0.051	0.043	-0.002	-0.024	-0.030
	(-0.07)	(-0.62)	(-0.72)	(-1.11)	(0.62)	(-0.04)	(-0.36)	(-0.52)
Children Indicator	0.019	0.022	-0.198	-0.138	0.007	0.021	-0.155	-0.073
	(0.25)	(0.39)	(-1.55)	(-1.55)	(0.08)	(0.30)	(-0.97)	(-0.73)
Principal Exam	-0.046	0.004	-0.004	-0.007	-0.123	-0.102	-0.073	-0.144
	(-0.57)	(0.05)	(-0.06)	(-0.09)	(-1.46)	(-1.50)	(-0.78)	(-1.42)
Pre-COVID FE	NO	NO	NO	NO	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Brokerage Firm FE	NO	NO	YES	YES	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES	YES	YES	YES	YES
Observations	514	339	507	331	362	246	353	236
R^2	0.420	0.170	0.600	0.341	0.432	0.205	0.591	0.364

Panel B: *AWL* and *PAD_HIGH*

	ALL				Pre-COVID			
	ALL	NS y-1	ALL	NS y-1	ALL	NS y-1	ALL	NS y-1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ave AWL Q1-Q3(Z)	0.039	0.012	0.007	-0.004	0.045*	0.022	0.006	-0.001
	(1.64)	(0.69)	(0.29)	(-0.23)	(1.75)	(1.06)	(0.23)	(-0.03)
PAD_HIGH	0.118**	0.066*	0.156***	0.110**	0.139**	0.081*	0.170***	0.136**
	(2.41)	(1.73)	(3.40)	(2.50)	(2.46)	(1.71)	(3.10)	(2.26)
Pre-COVID FE	NO	NO	NO	NO	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Brokerage Firm FE	NO	NO	YES	YES	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES	YES	YES	YES	YES
Observations	517	341	510	333	365	248	356	238
R^2	0.423	0.168	0.603	0.343	0.435	0.199	0.594	0.364

Panel C: *AWL*, *PAD_MED*, and *PAD_TOP*

	ALL				Pre-COVID			
	ALL	NS y-1	ALL	NS y-1	ALL	NS y-1	ALL	NS y-1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ave AWL Q1-Q3(Z)	0.039	0.013	0.008	-0.003	0.047*	0.027	0.009	0.004
	(1.61)	(0.77)	(0.35)	(-0.14)	(1.83)	(1.39)	(0.32)	(0.16)
PAD_MED	0.119**	0.057	0.149***	0.096*	0.124**	0.047	0.160***	0.110
	(2.21)	(1.23)	(3.03)	(1.86)	(2.04)	(0.82)	(2.76)	(1.61)
PAD_TOP	0.115*	0.084*	0.173***	0.141***	0.176**	0.156**	0.202***	0.214**
	(1.84)	(1.73)	(2.96)	(2.84)	(2.45)	(2.41)	(2.78)	(2.53)
Pre-COVID FE	NO	NO	NO	NO	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Brokerage Firm FE	NO	NO	YES	YES	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES	YES	YES	YES	YES
Observations	517	341	510	333	365	248	356	238
R^2	0.423	0.169	0.603	0.345	0.436	0.208	0.595	0.370

Table 9: Analyst Stock Level Accuracy Regressions

This table reports results from panel regressions of analyst Q1 forecast accuracy on *AWL*, *PAD*, ranked-based *PAD* measures, and other control variables. In particular, to explore the potential non-linearity of *PAD* we also include in the analysis dummy variables based on the distribution of *PAD*. In particular, *PAD_HIGH* is a dummy variable that receives the value of one if *PAD* is above the median of the *PAD* distribution in year *t*, and zero otherwise. In a similar manner, *PAD_MED* (*PAD_TOP*) is a dummy variable that receives the value of one if *PAD* is between the 50^{th} and 80^{th} percentiles (above the 80^{th} percentile) of the distribution in year *t*, and zero otherwise. *PMAFE* is the Analyst quarterly forecast accuracy measure based on Clement (1999) and Jame, Johnston, Markov, and Wolfe (2016). We require at least two analysts to issue earnings forecasts in a given quarter. "ALL' refers to the full sample. "Pre-COVID" refers to 2017-2019. (Z) stands for a Z-Score adjustment (a mean of zero and a standard deviation of one). We Z-Score adjust both the dependent variable and independent variables of interest. See Table A.1 and Table 1 for details about variable and sample definitions. The sample period is from September 2017 to March 2021. We keep analyst-quarter observations that meet the required quarterly login activity filter. Standard errors are clustered by analysts reported in parentheses below the coefficient estimates. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

Panel A: *AWL* and *PAD*

	ALL	PreCOVID	ALL	PreCOVID
	(1)	(2)	(3)	(4)
AWL(Z)	-0.022***	-0.019**	-0.016*	-0.011
	(-2.97)	(-2.30)	(-1.71)	(-0.99)
PAD(Z)	-0.003	-0.007	-0.004	-0.006
	(-0.41)	(-0.79)	(-0.39)	(-0.49)
Ave Q1 PMAFE t-4_t-1	0.439***	0.430***	-0.435***	-0.700***
	(5.85)	(5.04)	(-4.87)	(-6.17)
Early Forecast	0.001***	0.001*	0.001**	0.001
	(2.85)	(1.71)	(2.22)	(1.10)
IBES Years	0.002	0.002	-0.054	-0.517
	(1.35)	(1.43)	(-0.36)	(-0.99)
High Rank Indicator	-0.007	-0.006		
	(-0.43)	(-0.31)		
# Q1 EPS Forecasts	0.003***	0.002***	0.003***	0.003**
	(4.72)	(2.99)	(3.81)	(1.97)
# of GICS6 Industries	0.006	0.013*	0.004	-0.001
	(1.07)	(1.72)	(0.42)	(-0.05)
LnSize	-0.005	-0.002	-0.011	-0.008
	(-0.22)	(-0.07)	(-0.44)	(-0.21)
LnBM	0.007	0.014	0.002	0.010
	(0.51)	(0.64)	(0.14)	(0.45)
StdDev.Ret	0.375	0.486	0.195	0.272
	(0.55)	(0.25)	(0.29)	(0.14)
InstHold	0.045	0.090	0.048	0.082
	(0.99)	(1.49)	(1.03)	(1.32)
AMIHUD	-0.022	-0.027	-0.023	-0.029
	(-0.91)	(-1.28)	(-0.94)	(-1.32)
Firm FE	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES
Analyst FE	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES
Observations	36,711	25,888	36,710	25,887
R^2	0.091	0.108	0.108	0.129

Panel B: *AWL* and *PAD_HIGH*

	ALL	Pre-COVID	ALL	Pre-COVID
	(1)	(2)	(3)	(4)
AWL(Z)	-0.023***	-0.020**	-0.016*	-0.011
	(-3.19)	(-2.45)	(-1.75)	(-0.97)
PAD_HIGH	-0.025**	-0.033**	-0.022	-0.021
	(-2.03)	(-2.26)	(-1.45)	(-1.22)
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES
Analyst FE	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES
Observations	36,711	25,888	36,710	25,887
R^2	0.091	0.108	0.108	0.129

Panel C: *AWL*, *PAD_MED*, and *PAD_TOP*

	ALL	Pre-COVID	ALL	Pre-COVID
	(1)	(2)	(3)	(4)
AWL(Z)	-0.022***	-0.018**	-0.016*	-0.011
	(-3.03)	(-2.17)	(-1.72)	(-0.96)
PAD_MED	-0.033**	-0.045***	-0.026	-0.027
	(-2.43)	(-2.79)	(-1.63)	(-1.51)
PAD_TOP	-0.009	-0.009	-0.009	0.003
	(-0.52)	(-0.42)	(-0.40)	(0.09)
IBES Years	0.002	0.002	-0.047	-0.505
	(1.36)	(1.43)	(-0.32)	(-0.95)
High Rank Indicator	-0.006	-0.004	0.000	0.000
	(-0.34)	(-0.21)	(.)	(.)
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES
Analyst FE	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES
Observations	36,711	25,888	36,710	25,887
R^2	0.091	0.108	0.108	0.129

Table 10: *PAD* and COVID Lockdown Identification Strategy

This table reports results from panel regressions of analyst accuracy on *PAD* and other control variables using a difference-in-difference identification strategy. We focus on the period Q3-2019 to Q2-2020 and use the exogenous drop in *PAD* due to the COVID lockdown as a shock to analysts' ability to travel. We keep all analysts with full 4-quarter data and information about the analysts' home and work locations. This results in 102 unique analysts. We then calculate the average *PAD* during Q3 and Q4 of 2019 as a measure for the potential drop in *PAD*. In panel A, the treatment group includes analysts with *PAD* values above the median. In panel B the treatment group includes analysts with *PAD* values between the 50^{th} and 80^{th} percentiles (MID) and above the 80^{th} percentile (TOP). The pre- (post) period includes Q3-Q4 (Q1-Q2) of 2019(2020). *TREATMENT* $\times$ *POST* captures the potential difference in the drop in *PAD* between the treatment and the control group. All observations are at the analyst-quarter level. Consequently, *PMAFE* is the value-weighted average of the analyst accuracy measure across all stocks covered based on the stock market cap. FAR and NEAR are *PMAFE* averages for sub-groups on stocks that the analyst covers based on the distance between the analyst's home address and the covered firm headquarters. FAR (NEAR) refers to stocks that their headquarters is above (up to) 300 miles. See Table A.1 and Table 1 for details about variable and sample definitions. We keep analyst-quarter observations that meet the required quarterly login activity filter. Standard errors are clustered by analysts reported in parentheses below the coefficient estimates. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

Panel A: Above and Below Median

	ALL	FAR	NEAR
	(1)	(2)	(3)
TREATMENT	-0.049	-0.070	-0.041
	(-1.34)	(-1.64)	(-0.63)
POST	-0.038	-0.046	0.067
	(-0.79)	(-0.87)	(0.89)
TREATMENT $\times$ *POST*	0.117**	0.128**	-0.017
	(2.47)	(2.29)	(-0.19)
Ave # Stocks t-4_t-1	0.006*	0.008**	-0.001
	(1.79)	(2.08)	(-0.09)
Ave # of Industries t-4_t-1	-0.008	-0.004	0.003
	(-1.21)	(-0.61)	(0.24)
IBES Years	-0.002	-0.004	0.001
	(-0.78)	(-1.44)	(0.42)
Ave Q1 PMAFE t-4_t-1	-0.015	-0.076	-0.040
	(-0.14)	(-0.61)	(-0.19)
Coverage FE	YES	YES	YES
Location FE	YES	YES	YES
Analyst Cluster	YES	YES	YES
Observations	407	380	327
AdjR^2	0.036	0.042	0.030

Panel B: MID and TOP relative to Bottom

	BOT50 & 50-80 PCT			BOT50 & 80-100 PCT		
	(1) ALL	(2) FAR	(3) NEAR	(4) ALL	(5) FAR	(6) NEAR
TREATMENT	-0.052	-0.085*	-0.020	-0.060	-0.043	-0.122
	(-1.28)	(-1.75)	(-0.30)	(-1.39)	(-0.80)	(-1.33)
POST	-0.036	-0.051	0.073	-0.040	-0.026	0.057
	(-0.73)	(-0.94)	(0.94)	(-0.77)	(-0.45)	(0.70)
TREATMENT $\times$ *POST*	0.098*	0.108*	-0.073	0.145**	0.161**	0.073
	(1.86)	(1.75)	(-0.85)	(2.54)	(2.21)	(0.58)
Ave # Stocks t-4_t-1	0.006	0.010*	0.005	0.004	0.004	-0.000
	(1.53)	(1.81)	(0.86)	(1.20)	(1.04)	(-0.01)
Ave # of Industries t-4_t-1	-0.002	-0.002	0.015	-0.016**	-0.019**	0.017
	(-0.18)	(-0.16)	(1.28)	(-2.03)	(-2.14)	(0.97)
IBES Years	-0.001	-0.004	0.009**	0.001	-0.000	-0.002
	(-0.33)	(-1.32)	(2.47)	(0.21)	(-0.14)	(-0.37)
Ave Q1 PMAFE t-4_t-1	-0.026	-0.172	-0.193	-0.136	-0.108	0.264
	(-0.20)	(-1.44)	(-0.74)	(-1.03)	(-0.75)	(1.13)
Coverage FE	YES	YES	YES	YES	YES	YES
Location FE	YES	YES	YES	YES	YES	YES
Analyst Cluster	YES	YES	YES	YES	YES	YES
Observations	327	309	260	283	262	218
AdjR^2	0.014	0.047	0.050	0.026	0.012	0.040

Table 11: *AWL* and Commute Time Saved Identification Strategy

This table reports results from panel regressions of analyst output and accuracy measures on *AWL* and other control variables using a difference-in-difference identification strategy. We focus on the period Q3-2019 to Q2-2020 and use the COVID lockdown as a positive shock to analyst *AWL* due to saved commute time to work. We keep all analysts with full 4-quarter data and information about home and work locations. This results in 102 unique analysts. To reduce noise we remove the min and max values of analysts' commute time, which results in a final sample of 99 analysts. Panel A reports the relation between changes in *AWL*(in minutes) and commute time saved. In Panel B, we build on this relation and report difference-in-difference analysis. The treatment (control) group includes the analysts with time saved above (below) the median. The pre- (post) period includes Q3-Q4 (Q1-Q2) of 2019(2020). All observations are at the analyst-quarter level. Consequently, *PMAFE* is the value-weighted average of the analyst accuracy measure across all stocks covered based on the stock market cap. FAR and NEAR are *PMAFE* averages for sub-groups on stocks that the analyst covers based on the distance between the analyst's home address and the covered firm headquarters. FAR (NEAR) refers to stocks that their headquarters is above (up to) 300 miles. See Table A.1 and Table 1 for details about variable and sample definitions. We keep analyst-quarter observations that meet the required quarterly login activity filter. Standard errors are clustered by analysts reported in parentheses below the coefficient estimates. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

Panel A: *AWL* and Commute Time

	Changes in *AWL* in Minutes							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Commute-Time-Saved	1.314***	1.318***	1.328***	1.394***	1.387***	1.309***	1.320***	1.315***
	(2.90)	(2.92)	(2.87)	(2.88)	(2.94)	(2.86)	(2.75)	(2.75)
AGE		-0.097	-0.064	-0.128	-0.049	-0.094	-0.164	-0.135
		(-0.16)	(-0.11)	(-0.21)	(-0.05)	(-0.11)	(-0.22)	(-0.18)
Young Kids Indicator			-17.834	-16.855	-16.806	-23.829	-24.399	-24.713
			(-1.00)	(-0.95)	(-0.94)	(-1.36)	(-1.32)	(-1.32)
Female Indicator				20.286	20.087	21.879	20.122	18.216
				(1.06)	(1.04)	(1.12)	(0.90)	(0.78)
IBES Years					-0.198	-1.250	-1.326	-1.266
					(-0.16)	(-0.70)	(-0.67)	(-0.65)
Work Experience						3.017	3.089	3.260
						(1.40)	(1.37)	(1.39)
MBA Indicator						59.568	60.919	59.671
						(1.08)	(1.12)	(1.09)
# Jobs FINRA						3.136	3.279	3.679
						(0.71)	(0.70)	(0.73)
High Rank Indicator							5.332	6.025
							(0.22)	(0.25)
Principal Exam								-13.248
								(-0.76)
White SE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	102	102	102	102	102	102	102	102
AdjR^2	0.136	0.128	0.126	0.123	0.114	0.132	0.123	0.116

Panel B: Accuracy

	ALL	FAR	NEAR
	(1)	(2)	(3)
TREATMENT	0.046	0.035	0.069
	(1.52)	(0.90)	(1.22)
POST	0.060	0.043	0.109
	(1.36)	(0.77)	(1.55)
TREATMENT $\times$ *POST*	-0.085*	-0.060	-0.081
	(-1.75)	(-1.04)	(-1.00)
Ave # Stocks t-4_t-1	0.006*	0.009**	-0.000
	(1.89)	(2.23)	(-0.04)
Ave # of Industries t-4_t-1	-0.009*	-0.007	-0.003
	(-1.85)	(-1.20)	(-0.25)
IBES Years	-0.002	-0.004	0.002
	(-0.83)	(-1.63)	(0.50)
Ave Q1 PMAFE t-4_t-1	-0.012	-0.086	-0.005
	(-0.11)	(-0.66)	(-0.02)
Firm FE	YES	YES	YES
Coverage FE	YES	YES	YES
Location FE	YES	YES	YES
Analyst Cluster	YES	YES	YES
Observations	395	368	315
AdjR^2	0.032	0.033	0.033

A Appendix—Variable Definitions and Additional Tests

Table A.1 provides the variable definitions. Table A.2 and Table A.3 provide examples of named entities and quantitative statements used in Table 5's analysis. Tables A.4 - A.6 provide additional extensions.

In our main tests, we use *PAD* to proxy for analysts' Percentage Away Days (PAD) to quantify the extent of soft information collection that requires travel. We implicitly assume that analysts, given the nature of their work, do not engage in leisure and travel to meet institutional investors and engage in firm and other information-gathering activities. Nevertheless, in this appendix, we repeat the main tests (Section 4) using a percentage away measure that takes into account information events (*EvPAD*).

Specifically, *EvPAD* is calculated using away days that coincide with brokerage and firm events for stocks the analyst covers, which we collect using the Bloomberg Terminal's Corporate Events Calendar function (EVTS). This includes events hosted by firms such as Analyst Days and Investor Conferences. We report the results in Table A.4. The results indicate that *EvPAD* is associated with a higher probability of becoming a star analysis. *EvPAD* is also associated with higher accuracy, where the results are statistically significant once analyst fixed effects are included. We want to point out that since *EvPAD* is only based on firm and brokerage firms' events, it does not capture other interactions with institutional investors or other firm site visits. Thus, we view it as a lower bound for information-gathering activities.

Next, in our main tests, we use *AWL* to proxy for analysts' general effort provision or work ethics. The use of *AWL* is justified because analysts can engage in other productive activities at work rather than spending time on the Bloomberg terminal. Nevertheless, since analysts' terminal usage is not trivial, in this appendix, we repeat the main tests (Section 4) using an *intensive* usage measure that captures the analyst's minutes spent on the Bloomberg terminal. The measure, *LnActive*, is calculated as the natural logarithm of the average daily minutes of active Bloomberg usage in a quarter. We report the results in Table A.5. Overall, the results using *LnActive* are broadly consistent with the results using *AWL*.

In Table A.6, we repeat the analysis conducted in Table 9 controlling for Brokerage Firm Peers (team effort). Interestingly, they exhibit negative and somewhat significant coefficients, which suggests that the team effect is associated with higher accuracy. But, importantly, including a control for team effort doesn't alter our findings.

Table A.1: Variable definitions

Variable	Definition
Bloomberg User Data	
User Data	Bloomberg users with assigned accounts have an online "status" by default. This status is either designated as "online", "idle", "offline", or "mobile". When users first log on to the platform, their status changes from offline to online, and it remains that way while they use Bloomberg. However, if they stop using it for 15 minutes, the user's status automatically changes to "idle". Eventually, and depending on the users' settings, a user is logged off after a long period of inactivity. Using this information we construct various work habits measures.
Activity Measures based on Terminal Usage	
% of Workdays with Bloomberg Activity	The quarterly percent of working days with logged-in activity.
Active (minutes per day)	The quarterly average of the daily minutes that an analyst is actively logged-in to her Bloomberg terminal.
Conditional Active (on active days)	The quarterly average of the daily minutes that an analyst is actively logged-in to her Bloomberg terminal conditioning on days with Bloomberg activity.
LnActive	The natural logarithm of *Conditional Active*.
Active - hours per Week	The quarterly average of hours per week that the analyst is logged-in to the terminal.
AWL	For each analyst and year, we know the probability that an analyst is logged on every minute of the day. Using this information we construct a pdf. We then assume that the constructed distribution is a mixture of two normal distributions. This captures the idea that an analyst may have different morning and afternoon work habits. The distance *AWL*measures the difference between the means of the two distributions and adds a standard deviation on each side.
PAD	The quarterly average of a daily dummy variable that receives the value of one if an analyst is not logged in to her Bloomberg terminal during that day, and zero otherwise.
PAD_HIGH	A dummy variable that recieved the value of one if *PAD* is above the median of the sample distribution.

Variable	Definition
Analyst Coverage and Output Measures	
# Unique Stocks t-4_t-1	The number of unique stocks that an analyst covered over the previous four quarters.
Ave # Stocks t-4_t-1	The average number of stocks in a given quarter that an analyst covered over the previous four quarters.
# of GICS6 Industries	The average number of industries that an analyst covered over the previous four quarters. The industries are defined by the GICS six digit codes.
% of Common Stocks	The % of common stocks from all stocks that an analyst covers.
# of Stocks w Q1 EPS Forecasts	The number of stocks that an analyst issued a quarterly forecast for during a given quarter.
# Q1 EPS Forecasts	The number of Q1 earnings forecasts that an analyst issued across all stocks covered in a given quarter.
# Y1 EPS Forecast	The number of Y1 earnings forecasts that an analyst issued across all stocks covered in a given quarter.
# Long Term Growth Forecasts	The number of long-term forecasts that an analyst issued across all stocks covered in a given quarter.
# of Other Forecasts	The number of other earnings forecasts that an analyst issued across all stocks covered in a given quarter.
# of Rec	The number of stock recommendations that an analyst issued across all stocks covered in a given quarter.
# of non-stale Rec	The number of stock recommendation changes that an analyst issued across all stocks covered in a given quarter.
# of PTG	The number of 12-month price target forecasts that an analyst issued across all stocks covered in a given quarter.
Analyst Earnings Forecast Accuracy Measure	
PMAFE	Analyst quarterly forecast accuracy measure based on Clement (1999) and Jame, Johnston, Markov, and Wolfe (2016). The measure (Proportional Mean Absolute Forecast Error) is defined as $(AFE_{i,j,t} - \overline{AFE_{j,t}})\ /\ \overline{AFE_{j,t}}$, which is the absolute forecast error for analyst i's forecast of firm j minus the mean absolute forecast error for firm j in quarter t, divided by the mean absolute forecast error for firm j in quarter t. To calculate the measure, we require at least two analysts covering the stock on I/B/E/S in a given quarter. In particular, for each analyst i and firm j, we calculate the analyst's quarterly equally-weighted forecast errors average based on all earnings forecasts initiated during the quarter. We then calculate the absolute value of the analyst average forecasts errors. We repeat the calculation for all analysts on I/B/E/S covering the stock in that quarter and calculate the stock's quarterly mean absolute forecasts errors.
AveQtrAccuracy	The average of the analyst quarterly forecast accuracy measure ($PMAFE$) across all the stocks covered in a given quarter.
AveQtrAccuracy_VW	The value weighted average of the analyst quarterly forecast accuracy measure ($PMAFE$) across all the stocks covered in a given quarter. The weights are based on the stock's market capitalization.

Variable	Definition
Analyst Forecast Timeliness Measures	
LnTFE	The analyst earnings forecasts timeliness measure, based on the natural logarithm of the time in days from the earnings announcement and the analyst subsequent earnings forecast. Specifically, for each analyst i, stock j and quarter q, we calculate the number of days from the earnings announcement during quarter q and the subsequent analyst earnings forecast. We then calculate the equally-weighted average across all covered stocks.
Analyst Portfolio Based Measures	
# Stocks in AbnVOl Decile t-1	The number of stocks in the analyst's portfolio that are in the top decile of day t-1 abnormal trading volume of CRSP's cross-sectional ranking. Abnormal volume is calculated as the split adjusted daily stock volume divided by the the split adjusted average trading volume over the past 63 trading days.
# Stocks in AbsExtRet Decile t-1	The number of stocks in the analyst's portfolio that are in the top decile of day t-1 absolute return of CRSP's cross-sectional ranking
# Stock with AMC News t-1	The number of stocks in the analyst portfolio that had after-market-close news on day t-1. The news data is obtained from the Dow Jones Edition of RavenPack Analytics from 2017 to August 2020. To ensure that we capture relevant news, we identify news with a relevance score of 100, which ensures that the news is about the firm of interest, from the following news-types: news-flash, hot-news-flash, full article, and press release. To ensure we capture fundamental news we keep the following 13 news categories: acquisitions-mergers, analyst-ratings, assets, credit, credit-ratings, dividends, earnings, equity-actions, labor-issues, legal, marketing, products-services, and partnerships.
# Stock with AMC Earn News t-1	The number of stocks in the analyst portfolio that had after-market-close earnings news on day t-1.
# Stock with AMC AR News t-1	The number of stocks in the analyst portfolio that had after-market-close analyst rating news on day t-1.
# Stock with BMO News t	The number of stocks in the analyst portfolio that had before-market-open news on day t.
# Stock with BMO Earn News t	The number of stocks in the analyst portfolio that had before-market-open earnings news on day t.
# Stock with BMO AR News t	The number of stocks in the analyst portfolio that had before-market-open analyst rating news on day t.
# Max Industry Earn BMO News Pressure t	We construct an industry earnings news pressure variable, calculated as the market-cap value-weighted earnings news dummy across all CRSP's stocks in a specific Fama-French 48 industry. We then take the maximum across all the industries that are covered by the analyst.
Analyst Additional Characteristic Based Measures	
Data	We manually obtain analyst characteristics data from FINRA's BrokerCheck website, LinkedIn and Facebook.
High Rank Indicator	A dummy variable that received a value of one if the analyst specifies a managing director (high rank) title in his public profiles, and zero otherwise.
STAR	A dummy variable that received a value of one if the analyst is ranked as a star analysis in year t by Institutional Investor All-America Research Team, and zero otherwise.
Work Experience	The number of work experience in years, obtained from FINRA.
# Jobs FINRA	The number of jobs that an analyst had switched, obtained from FINRA.
NYC Indicator	A dummy variable that received a value of one if the analyst work in New York, and zero otherwise.
MBA Indicator	A dummy variable that received a value of one if the analyst specifies an MBA degree in his public profiles, and zero otherwise.

Variable	Definition
Analyst Additional Characteristic Based Measures (cont'd)	
Principal Exam	A dummy variable that received a value of one if the analyst has taken a principal exam and zero otherwise. Around 10% of the analysts in our sample have taken the principal exam. The information is obtained from FINRA.
AGE	The age of the analyst.
Female Indicator	A dummy variable that received a value of one if the analyst is a female and zero otherwise.
Children Indicator	A dummy variable that received a value of one if an analyst has children, and zero otherwise.
Young Kids Indicator	A dummy variable that received a value of one if an analyst has non-adult children, and zero otherwise.
Commute-Time-Saved	We verify the home address and work address of an analyst using data from FINRA BrokerCheck, Mergent Intellect, and LinkedIn. Using Google Maps, we then measure the minimum typical travel time between home and work at 7:00 am on a workday. Commute time is the minimum travel time across various options (public transit, automobile, bicycle, and foot travel). *Commute-Time-Saved*, is simply the commute time that an analyst saves due to working from home.
Additional Analyst Controls	
IBES Years	The analysts experience measured by the number of years in I/B/E/S.
AveQtrAccuracy	The analyst quarterly *PMAFE* average across all covered stocks.
Ave # Q1 EPS Forecasts t-4_t-1	The average of the quarterly number of earnings forecasts over the previous 12 months.
Ave # of Industries t-4_t-1	The average of the quarterly number of different industries that the analyst covers over the previous 12 months.
Stock Controls and fixed effects	
LnSize	The natural logarithm of the stock market capitalization.
LnBM	The natural logarithm of the stock book-to-market ratio.
BM_Dummy	A dummy variable that receives the value of one if book-to-market information is available, and zero otherwise. We augment book-to-market missing values with zeros.
StdDev.Ret	The standard deviation of stock daily stock returns.
InstHold	The stock quarterly percentage of institutional holdings.
Coverage fixed effects	To control for the number of stocks an analyst covers, every quarter we rank all analysts in our sample by the number of stocks they covered over the previous year into ten deciles. We then use the ranking to include coverage fixed effect.
Time fixed effects	We include time fixed effects in our regressions based on year-qtr pairs.

Table A.2: Named Entity Examples

This table provides examples of named entities extracted from analyst sentences using the RoBERTa transformer model within Python's SpaCy natural processing library. In the exampels, named entities are surrounded with curly braces and followed by a colon and named entity category.

Panel A: Named Entity Examples

"{Steve: PERSON}, can you talk about the product and technological differences between {ADT Pulse: PRODUCT} and {ADT Command: PRODUCT}, and what's the likelihood that {ADT: ORG} might become keen on moving the {Pulse: PRODUCT} subscribers to {ADT Command: PRODUCT}?"

"{Jack: PERSON}, just on the {30-plus: CARDINAL} new live content partnerships in {1Q: DATE}, can you talk a little bit about what you're particularly excited about going forward and perhaps give a little more color on the {FIFA: ORG} {World Cup: EVENT} deal in particular?"

"I think, {Scott: PERSON}, back at {Investor Day: EVENT}, I think you threw out a {$5 million: MONEY} EBIT per {Airbus: ORG} plane figure."

"And my follow-up, {Tim: PERSON}, is could you talk about {Concho's: ORG} efforts in terms of delineating {the Upper Wolfcamp Sands: LOC} and {the Northern Delaware Basin: LOC} as well as your {3rd: ORDINAL} {Bone Spring: LOC} results in {the south: LOC}?"

Table A.3: Quantitative and Non-Quantitative Sentence Examples

This table provides examples of sentences that are labeled as quantitative or non-quantitative by OpenAI's "gpt-3.5-turbo" model. We used the prompt "Your goal is to classify the question/statement from an equity analyst as quantitative or non-quantitative. You will also provide a brief explanation for your classification.", provided the question and requested the output in json format. Panel A provides a few examples of sentences classified as quantitative along with the explanation and Panel B provides examples and explanations for non-quantitative sentences.

Panel A: Quantitative Sentence Examples

"Can you just confirm the remaining authorization is $400 million?"
GPT explanation: 'This question asks for a specific numerical confirmation of the remaining authorization amount, which is $400 million.'

"But are we talking about something like under 70% as a normalized run rate once you're back to a more normal sales environment?"
GPT Explanation: "This question asks for a specific numerical percentage (under 70%) as a normalized run rate, which is a quantitative inquiry."

"And just tell us, I mean, what is the size of your business for the U.S. Navy?"
GPT Explanation: "This question asks for a specific numerical figure, which is the size of the business for the U.S. Navy."

Panel B: Non-Quantitative Sentence Examples

"Do you participate and if so, how?"
GPT Explanation: 'This question asks for a descriptive explanation of how the individual participates, rather than requesting specific numerical data.'

"And perhaps speak to the issue of flaring and how PRRC has positioned Texas and your conversations with them."
GPT Explanation: 'This sentence asks for a discussion or explanation regarding the issue of flaring and the positioning of Texas by PRRC, which is qualitative in nature and does not involve specific numerical data.'

"Are there any opportunities or risks you're looking at there?"
GPT Explanation: 'This question asks for a description of opportunities and risks without requesting specific numerical data.'

Table A.4: Probability of Being a Star Analyst and Accuracy - *EvPAD*

This table repeats the analysis conducted in Table 8 (star) and Table 9 (accuracy), replacing *PAD* with *EvPAD*. We downloaded data from the Bloomberg terminal on firm and other brokerage-hosted events from 2017 to 2021. For each analyst, we match all events with the stock the analyst covers. *EvPAD* aims to capture the percentage of days away from the office that captures information-gathering activities. In particular, we keep away days that coincide with brokerage and firm events collected from the Bloomberg Terminal Corporate Events Calendar function (EVTS) for each company. We allow up to 5 days if there is one event during that window or up to 10 away days if there are multiple events during the window. We remove any other days that do not fit these conditions. Given that *EvPAD* aims to capture in-person interactions, we focus on the pre-COVID sample period. Standard errors are clustered by analysts reported in parentheses below the coefficient estimates. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

Panel A: STAR and *EvPAD*

	Pre-COVID			
	ALL	NS y-1	ALL	NS y-1
	(1)	(2)	(3)	(4)
Ave AWL Q1-Q3(Z)	0.026	0.010	-0.019	-0.022
	(1.07)	(0.52)	(-0.76)	(-1.03)
AveEvPAD Q1-Q3(Z)	0.072**	0.049	0.072**	0.066**
	(2.14)	(1.38)	(2.44)	(1.99)
Brokerage Firm FE	NO	NO	YES	YES
Controls	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES
Analyst Cluster	YES	YES	YES	YES
Observations	362	246	353	235
R^2	0.433	0.202	0.593	0.357

Panel B: Accuracy and *EvPAD*

	Pre-COVID			
	(1)	(2)	(3)	(4)
AWL(Z)	-0.018**	-0.018**	-0.011	-0.009
	(-2.14)	(-2.15)	(-0.84)	(-0.75)
EvPAD(Z)		-0.011		-0.021***
		(-1.51)		(-2.60)
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES
Analyst FE	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES
Observations	25,556	25,556	25,555	25,555
R^2	0.108	0.108	0.129	0.129

Table A.5: Probability of Being a Star Analyst and Accuracy - *LnActive*

This table repeats the analysis conducted in Table 8 (star) and Table 9 (accuracy), replacing *AWL* with *LnActive*. Standard errors are clustered by analysts reported in parentheses below the coefficient estimates. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

Panel A: STAR and *LnActive*

	ALL				Pre-COVID			
	ALL	NS y-1	ALL	NS y-1	ALL	NS y-1	ALL	NS y-1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LnAveActive Q1-Q3(Z)	0.068**	0.036	0.032	0.007	0.079***	0.040*	0.053	0.022
	(2.41)	(1.57)	(1.09)	(0.23)	(2.78)	(1.70)	(1.56)	(0.68)
PAD_MED	0.147**	0.074	0.164***	0.101*	0.162**	0.064	0.192***	0.122*
	(2.52)	(1.47)	(3.26)	(1.86)	(2.42)	(1.05)	(3.16)	(1.74)
PAD_TOP	0.207***	0.142**	0.223***	0.156**	0.286***	0.216***	0.292***	0.260**
	(2.75)	(2.57)	(3.09)	(2.33)	(3.29)	(3.00)	(3.26)	(2.38)
Pre-COVID FE	NO	NO	NO	NO	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Brokerage Firm FE	NO	NO	YES	YES	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES	YES	YES	YES	YES
Observations	515	339	508	331	364	247	355	237
R^2	0.425	0.173	0.604	0.344	0.441	0.210	0.598	0.371

Panel B: Accuracy and *LnActive*

	ALL	Pre-COVID	ALL	Pre-COVID
	(1)	(2)	(3)	(4)
LnActive(Z)	-0.016*	-0.009	-0.028*	-0.005
	(-1.80)	(-0.97)	(-1.76)	(-0.22)
PAD_MED	-0.037***	-0.048***	-0.035**	-0.029
	(-2.62)	(-2.84)	(-2.15)	(-1.53)
PAD_TOP	-0.024	-0.015	-0.038	-0.003
	(-1.14)	(-0.58)	(-1.36)	(-0.08)
CONTROLS FE	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES
Analyst FE	NO	NO	YES	YES
Analyst Cluster	YES	YES	YES	YES
Observations	36,711	25,888	36,710	25,887
R^2	0.091	0.108	0.108	0.129

Table A.6: Analyst Stock Level Accuracy Regressions - Controlling for Brokerage Firm Peers

This table repeats the analysis conducted in Table 9 controlling for Brokerage Firm Peers' *AWL*. The sample period is from September 2017 to March 2021. See Table A.1 and Table 1 for details about variable and sample definitions. *Brokerage-Firm PeerAWL* is the average *AWL* of the brokerage firm in a given year and quarter, excluding the analyst. Using Investext database, we also identified 3,672 stock-analyst-quarter observations for which we have team *AWL* data. *AUG Brokerage-Firm PeerAWL* then, is a variant of *Brokerage-Firm PeerAWL* where we augment *Brokerage-Firm PeerAWL* with the average *AWL* of the Investext identified Bloomberg team analysts that are cosigned on the firm reports. All specifications include brokerage-firm fixed effect. Standard errors are clustered by analysts reported in parentheses below the coefficient estimates. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively.

	ALL		Pre-COVID		ALL		Pre-COVID	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AWL(Z)	-0.023***	-0.022***	-0.019**	-0.019**	-0.015	-0.015	-0.010	-0.009
	(-3.08)	(-3.02)	(-2.25)	(-2.21)	(-1.62)	(-1.59)	(-0.83)	(-0.78)
PAD_MED	-0.033**	-0.033**	-0.045***	-0.046***	-0.025	-0.025	-0.027	-0.027
	(-2.42)	(-2.46)	(-2.80)	(-2.83)	(-1.54)	(-1.59)	(-1.52)	(-1.54)
PAD_TOP	-0.010	-0.010	-0.009	-0.010	-0.005	-0.006	0.004	0.003
	(-0.62)	(-0.61)	(-0.43)	(-0.47)	(-0.24)	(-0.26)	(0.14)	(0.11)
Brokerage-Firm PeerAWL	-0.015*		-0.017*		-0.008		-0.021	
	(-1.69)		(-1.87)		(-0.65)		(-1.60)	
AUG Brokerage-Firm PeerAWL		-0.008		-0.009		-0.003		-0.008
		(-1.03)		(-0.92)		(-0.26)		(-0.64)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES	YES	YES
Coverage x Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Analyst FE	NO	NO	NO	NO	YES	YES	YES	YES
Analyst Cluster	YES	YES	YES	YES	YES	YES	YES	YES
Observations	36,711	36,711	25,888	25,888	36,710	36,710	25,887	25,887
R^2	0.091	0.091	0.108	0.108	0.108	0.108	0.129	0.129