

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Journal of Corporate Finance

journal homepage: www.elsevier.com/locate/jcorpfin

Air pollution and analyst information production

Chao Kevin Li^a, Jin-hui Luo^{b,*}, Naomi S. Soderstrom^c^a School of Accounting, Business School, UNSW Sydney, Australia^b Department of Accounting, School of Management, Xiamen University, China^c Department of Accounting, Faculty of Business and Economics, University of Melbourne, Australia

ARTICLE INFO

Keywords:

Particulate matter (PM) pollution

Air pollution

Analysts

Forecast timeliness

Forecast boldness

Work productivity

ABSTRACT

Recent studies investigate the impact of air pollution on labor productivity. We extend this literature by showing that air pollution negatively affects equity analyst information production. Analysts exposed to air pollution are less likely to issue timely forecasts or improve their forecast accuracy. Investigating the underlying mechanism, we find that analysts exposed to air pollution are less likely to provide bold (especially, negatively bold) forecasts. We also find evidence that market pricing is less sensitive to forecast revisions issued by analysts exposed to air pollution. Our results are robust to controlling for firm/analyst and time fixed effects, as well as additional specifications employing difference-in-differences designs and placebo tests.

1. Introduction

In the past five years, particulate matter (PM) pollution has risen 8%, with over 80% of urban residents worldwide currently exposed to dangerous levels of outdoor pollution (Vidal, 2016; Walsh, 2016).¹ Recent economic studies find that outdoor pollution results in a reduction of worker productivity (Chang et al., 2014; Chang et al., 2016; Dong et al., 2019; Zivin and Neidell, 2012). We extend this enquiry by investigating whether outdoor air pollution affects analyst information production for the capital market.

We examine two related issues. First, we investigate whether PM pollution affects the timing and accuracy improvement of analyst earnings forecast revisions in response to firm earnings announcements. Our evidence is consistent with forecasts being less timely and of lower accuracy during polluted periods. Second, we investigate whether this effect is related to analysts having greater difficulty in information production during polluted periods. Consistent with forecasts being less-informed, we find that when analysts are in polluted cities during an earnings announcement, they are less likely to issue bold (especially, negatively bold) forecasts. Forecast boldness reflects the extent to which analysts map new information into revised forecasts (Clement and Tse, 2005; Gleason and Lee, 2003; Hong et al., 2000; Keskek et al., 2014). Our results are thus consistent with analysts reducing the extent to which they incorporate new information into their forecasts when they are forecasting under polluted conditions. Overall, our study advances understanding of the effects of air pollution on the work productivity of highly skilled professionals, providing additional insights about the hidden costs associated with air pollution.

We address our research questions by examining analyst forecast revisions in response to firm earnings announcements in China. This setting is particularly appropriate for our study. In the past two decades China has experienced severe PM pollution, which varies by location and across time. This allows us to detect differences in analyst information production relative to changes in levels of PM pollution for the same firm/analyst. Focusing on forecast revisions is also an important feature of our study. Although analysts may

* Corresponding author.

E-mail addresses: k.li@unsw.edu.au (C.K. Li), jinhui Luo@xmu.edu.cn (J.-h. Luo), naomiss@unimelb.edu.au (N.S. Soderstrom).¹ Throughout the study, we interchangeably use air pollution level, particulate matter pollution, PM pollution and outdoor pollution.<https://doi.org/10.1016/j.jcorpfin.2019.101536>

Received 17 January 2019; Received in revised form 17 September 2019; Accepted 28 October 2019

Available online 02 November 2019

0929-1199/ © 2019 Elsevier B.V. All rights reserved.

make a strategic trade-off between the timeliness and accuracy of their forecasts (Gul and Lundholm, 1995; Guttman, 2010), they face greater pressure to issue forecasts directly following firm earnings announcements (Zhang, 2008). Given this pressure, analysts are unlikely to delay their forecast revisions to reduce the negative effects of air pollution. On the other hand, our use of a short-term window after earnings announcements to capture the effects of PM pollution mitigates the concern that confounding factors other than air pollution are responsible for the changes in analysts forecast revisions that we observe. Taken together, these features allow a strong test of our research questions. Finally, China's air pollution is representative of similar issues in other developing countries in the world (e.g., Brazil, India, Iran). For decades, China has been criticized for developing its economy at the expense of the environment (Christmann and Taylor, 2001). Results from this setting may shed light on indirect costs related to such a country-level strategy, with practical implications for many developing countries/regions worldwide that follow a similar development path.

A large number of studies suggest that air pollution worsens the performance of agents. This effect is due to its negative effects on health (Chang et al., 2014; Chang et al., 2016; Lavy et al., 2014; Suglia et al., 2007; Zivin and Neidell, 2012) and mood (Bakian et al., 2015; Calderon-Garciduenas et al., 2015; Lim et al., 2012; Power et al., 2015; Szyszkowicz, 2007).² Extant research finds that air pollution leads to biased decisions by investors (Dong et al., 2019; Heyes et al., 2016; Li et al., 2019). We extend this research to investigate whether air pollution impacts analyst forecasts, which play an important informational role in capital markets.

Our empirical findings lend support to our hypotheses. Compared to analysts forecasting during periods of low PM pollution, analysts exposed to the highest levels of pollution are 14% less likely to revise their earnings forecasts within a two-day short window of earnings announcements. In addition, for forecast revisions issued during the short-term window of earnings announcements, analysts facing the worst air pollution levels are 4% less likely to improve their forecast accuracy than are analysts experiencing good air quality. In terms of forecast boldness, analysts facing the highest levels of air pollution are 4.8% (4.9%) less likely to issue bold (negatively bold) forecasts compared with analysts experiencing good air quality. When benchmarked against sample means of our dependent variables, these negative impacts are economically significant. All these findings remain robust when we include a battery of control variables at both the firm and analyst levels, as well as firm/analyst/time fixed effects and for alternative model specifications.

To strengthen our identification and rule out the possibility that the observed findings arise from high ability analysts migrating to cities with low air pollution, we conduct difference-in-differences (hereafter diff-in-diff) designs following Li et al., 2019. To identify the treatment sample, we require that for an analyst-firm observation, there is a stark change (at least one standard deviation) in the air quality index (AQI) between two adjacent quarters. Use of a stark quarter-to-quarter change in AQI not only takes the advantage of exogenous variation in air quality but also avoids the possibility that high ability analysts migrate to clean cities in short-term. For each observation in the treatment sample, we choose an observation from the remaining sample that did not experience drastic changes in AQI as the control observation. The control observation must have forecasts issued by an analyst from a different city but following the same firm over the same period as the treatment observation. As a result, the difference between changes in the dependent variables for the treatment sample and changes for the control sample over the same period should reflect the effects of air pollution. Results from the diff-in-diff tests show that analysts exposed to air pollution (clean air) tend to issue less (more) timely, less (more) accurate and less (more) negatively bold forecasts compared with analysts without experiencing drastic air quality changes. Moreover, we conduct two placebo tests by randomizing PM pollution within each analyst-firm and by randomizing cities to each analyst-quarter. Both placebo tests help mitigate the concern that our findings are driven by some time-variant city-level omitted variables. Both tests corroborate our main findings.

To understand whether there are certain factors which could alleviate the negative effects of air pollution on analyst information production, we partition the full sample based on several firm or analyst level characteristics and re-estimate our baseline models. We find that factors related to both analyst supply of forecasts and investor demand for forecasts play a key role. From the supply side, the negative effects of air pollution are more pronounced in the subsample where analysts have a higher workload or where competition among analysts is weak. From the demand side, investors' high level of demand for information following annual earnings announcements attenuates the negative effects of the PM pollution.

We also investigate whether investors understand the reduction in forecast quality when analysts provide forecasts during periods with high levels of air pollution. We examine whether investor response to forecast revisions is moderated by the levels of air pollution faced by analysts. We compute the cumulative abnormal return (CAR) on days t and $t + 1$ relative to the forecast revision date. We then regress the CAR on an interaction term of forecast revision and air quality over the two days preceding the forecast revision date. We find that investors respond less to forecast revisions issued after polluted days, suggesting that the market understands the negative effects of air pollution on analyst information production.

We contribute to three strands of literature. First, we extend research investigating the impact of air pollution on worker productivity (Archsmith et al., 2016; Chang et al., 2014; Chang et al., 2016; Zivin and Neidell, 2012) and research investigating how air pollution affects decisions made by capital market participants (Dong et al., 2019; Heyes et al., 2016; Li et al., 2019). Previous research focuses on output of physical goods, non-financial services or investor trading behaviors. However, the study of how air pollution affects highly skilled professionals (e.g., analysts) is very limited. To the best of our knowledge, the only paper closely related to our study is Dong et al. (2019), who find that air pollution during analyst site visits induces more pessimistic forecasts (forecast bias). We extend Dong et al. (2019) by examining the impact of air pollution on additional dimensions of analyst information production. First, we focus on analyst forecast revisions, which proxy for analyst information production (Clement and Tse, 2005; Keskek et al., 2014), rather than on forecast bias. Second, whereas Dong et al. (2019) examine a window of 15 days, we employ

² See a review paper by Block et al. (2012).

a 2-day time period. Our short-term window mitigates concerns of confounding factors underlying our results.

Second, we contribute to the literature linking weather conditions to information production (more broadly, decision making) of market participants (Chen et al., 2017; DeHaan et al., 2017; Heyes et al., 2016; Li et al., 2019; Huang et al., 2018). Air pollution is fundamentally different from the more general notion of weather conditions (e.g., sunshine). Unlike bad weather, air pollution can be directly influenced by policy makers, implying that the costs of air pollution observed in our study might be controllable.³ Further, our results hold after controlling for local weather conditions. Indeed, we find that the effect of air pollution on analyst forecast timeliness is as large as that of weather conditions. Unlike poor weather, however, we find that air pollution significantly affects analyst forecast accuracy and the likelihood of analysts to issue bold forecasts (particularly negatively bold forecasts). This indicates that relative to poor weather, air pollution has large incremental impacts on analyst information production.

Finally, our study adds to the literature on determinants of analyst forecast timeliness and properties of analyst forecasts (Clement and Tse, 2005; Yezege, 2015; Zhang, 2008). Our findings imply that the air pollution may negatively affect functioning of the capital market. Given that air pollution is widespread and a growing problem throughout the world, particularly in emerging countries/regions, our findings provide insights regarding an additional indirect cost of prioritizing economic development over environmental protection.

The remainder of the paper proceeds as follows. Section 2 briefly discusses the air pollution in China. Section 3 summarizes prior literature and proposes hypotheses. Section 4 describes the research design. Section 5 reports empirical findings and Section 6 concludes.

2. Particulate matter pollution in China

The Chinese government started to release the daily air quality index (AQI) in 2009. AQI calculation is based on three criteria air pollutants and follows an algorithm developed by the U.S. Environmental Protection Agency (EPA, 2006). The pollutant that has the highest index determines the AQI on a specific day. In China, the primary form of air pollution is PM, measured as detrimental airborne particulate matter with a diameter smaller than 10 μm (PM10) (one seventh of the width of a human hair).⁴

Detrimental effects of PM are pervasive, since the particles can easily penetrate buildings due to their small size (Vette et al., 2001). To curb the negative effects of PM pollution, most offices keep air filters operating during periods of high pollution. However, the usefulness of the filters is questionable. In a survey regarding the effectiveness of air filters, 72% of respondents did not believe that filters reduce indoor PM pollution.⁵ In addition to concerns about the ineffectiveness of air filters to clear indoor pollution, there are concerns that a widely-used technology, which was developed in Western countries to filter larger pollutants (such as pollen and dust), may be useless in cleaning Chinese PM pollutants.⁶ Given the limited power of air filters, air pollution can affect workers even if they remain in their offices during polluted days (Chang et al., 2016; Dong et al., 2019; Li et al., 2019).

3. Hypothesis development

3.1. The direct influence of air pollution on analyst information production

A large body of literature provides evidence that PM pollution results in physical diseases, such as respiratory conditions and impairment of cognitive performance (Calderon-Garciduenas et al., 2015). Lavy et al. (2014) investigate the effects of fine PM on Israeli high school student performance on tests. Results indicate that test scores and the likelihood of students to enter university are negatively associated with exposure to PM pollution during exam days. Suglia et al. (2007) indicate that a high level of black carbon (a traffic pollutant) decreases intelligence and cognitive function.

PM pollution also has a negative impact on mood. Lim et al. (2012) find that increases in PM level are correlated with symptoms of depression. After controlling for weather conditions, Szyszkowicz (2007) finds that PM pollution is positively associated with daily depression-related emergency room visits. Power et al. (2015) find that exposure to PM2.5 pollution is associated with high symptoms of anxiety. Building upon the negative effects of PM pollution on mood, Bakian et al. (2015) find evidence that exposure to PM pollution increases the incidence of suicide attempts.

Based upon the negative effects of PM pollution on physical health and mood motivate, economic research has examined the effects of air pollution on workers' marginal productivity. Zivin and Neidell (2012) find that U.S. agricultural worker productivity is significantly harmed when there is an increase in ozone pollution levels. Chang et al. (2014) extend this enquiry by examining effects of PM pollution on indoor workers' productivity, focusing on workers at a U.S. pear-packing factory. Their results indicate a negative association between PM2.5 pollution and worker productivity. Chang et al. (2016) provide the first evidence that outdoor PM pollution also affects office workers' productivity. They compare the number of calls answered by call center workers in Shanghai and

³ Market participants can move to locations with better local weather to mitigate the negative effects of bad weather. However, participants' current locations probably reflect consideration of cost/benefit trade-offs and there are likely significant frictions associated with changing locations in the short term.

⁴ Other systems track particles as small as 2.5 μm (PM2.5). For example, the U.S. Environmental Protection Agency has set standards for both PM10 and PM2.5 (see https://www3.epa.gov/ttn/naaqs/standards/pm/s_pm_history.html)

⁵ Details at: <http://view.news.qq.com/original/intouchtoday/n2944.html>

⁶ Details at: http://www.kongqiwu.com/m/index.php?m=default&information_list&catid=23&sCatid=37&id=65

Nantong on high PM pollution versus low PM pollution days and find that workers exposed to high PM pollution answer fewer calls. Further analysis indicates that the reduction in workers' productivity is mainly driven by their reluctance to work during high pollution days. Archsmith et al. (2016) study professional baseball umpires—a setting with highly-skilled and quality-focused employees, who make complex decisions. They find that an increase in Carbon Monoxide/PM2.5 pollution increases the propensity of umpires to make incorrect calls.

An emerging body of studies investigates how the negative mood induced by poor environmental conditions affects investor behavior in the capital market. Heyes et al. (2016) find that one standard deviation increase in PM2.5 is associated with an 11.9% decrease in the same-day return. They argue that this result may be driven by decreased risk tolerance due to pollution-induced changes in mood or cognitive function. DeHaan et al. (2017) document that, earnings announced on bad weather days exhibit high post earnings announcement drift and low trading volume, implying that bad weather reduces investors' ability to digest earnings news. In their study of the impact of bad weather on analysts, DeHaan et al. (2017) find that relative to those exposed to good weather, analysts exposed to bad weather are less likely to revise their forecasts. These analysts also issue more pessimistic forecasts in response to earnings announcements. Chen et al. (2017) document that the negative mood induced by bad weather also leads to pessimism in management forecasts. Focusing on air pollution, Dong et al. (2019) find that analysts who make site visits to companies on polluted days tend to issue more pessimistic forecasts.

Building upon the above literature, we argue that PM pollution may negatively affect the analyst's ability and motivation to process information contained in earnings announcements to provide a forecast. Forecasting is a complex task, for which analysts must consider not only firm-specific but also intra/inter- industry information (Ivkovic and Jegadeesh, 2004; Jacob et al., 1999; Piotroski and Roulstone, 2004). In addition to skill and knowledge, the analyst's physical health, proper cognitive performance, and mood can impact forecasting ability. Moreover, as illustrated by Chang et al. (2016), if PM pollution reduces analysts' willingness to work (i.e., to synthesize information and formulate a forecast), analysts are unlikely to effectively process the information embedded in earnings announcements.

Prior literature suggests that analysts trade off late issuance of more accurate forecasts with early issuance of relatively inaccurate forecasts (Gul and Lundholm, 1995; Guttman, 2010). To compensate for the negative impacts of PM pollution on their information processing, analysts seeking to reduce the effect of pollution on their forecast quality may postpone issuing a revised forecast following earnings announcements. However, since investors may value early, low accuracy forecasts that lead peer analysts' forecasts and provide more timely signals, in the face of high PM pollution, analysts may choose to issue an early forecast even though the pollution has deterred them from fully synthesizing the information in earnings. Based upon this discussion, our hypotheses explore the timing and quality impacts of PM pollution:

H1a. *There is a negative association between PM pollution and the likelihood of analysts issuing forecast revisions immediately after earnings announcements.*

H1b. *Improvement in forecast accuracy from forecast revisions following earnings announcements is lower for analysts facing high PM pollution.*

3.2. Channels underlying PM pollution's effect on analyst forecasts

Hypotheses H1a and H1b build on the assumption that air pollution constrains analysts' information production within a short window after earnings announcements. Consistent with Truman (1994), Gleason and Lee (2003), Clement and Tse (2005) and Keskek et al. (2014), we argue that revised analyst forecasts, benchmarked against their own prior forecasts and peer analyst prior forecast consensus, reflect analyst information production. Gleason and Lee (2003) provide evidence that deviation from the prevailing consensus by an individual analyst (i.e., issuing a bold forecast) is associated with a stronger market reaction than when the analyst merely mimics peer forecasts. Clement and Tse (2005) find that bold forecasts are more accurate than other forecasts, and that herding forecasts are indicative of lower forecasting ability.

Bold forecasts not only reveal the analyst's private information to the market, but also signal the analyst's confidence in forecasting and the analyst's skill (Clement and Tse, 2005; Ramnath et al., 2008). Although both high- and low-ability analysts may issue bold forecasts (Clarke and Subramanian, 2006), career concerns of analysts reduce the incidence of this behavior, since inaccurate forecasts are associated with higher likelihood of career termination (Clement and Tse, 2005; Hong and Kubik, 2003). Further, when forecasts are bold, investors attribute poor forecasting quality to low analyst ability and effort (Kadous et al., 2009). Because pollution negatively affects information synthesis by analysts, we expect analysts experiencing pollution to be less willing to provide bold forecasts. This leads to the following hypothesis:

H2a. *There is a negative association between PM pollution and forecast boldness for forecast revisions issued within a short window after earnings announcements.*

Prior studies find that analysts are likely to issue optimistic forecasts because of conflicts of interest and/or cognitive bias (Beyer et al., 2010; Ramnath et al., 2008). Analysts may want to appear optimistic to facilitate their access to management, to advance their career, or to win investment banking business (Hong and Kubik, 2003; Ke and Yu, 2006; Michaely and Womack, 1999; O'Brien et al., 2005). Cognitive biases can also lead to over-optimism among analysts (Kadous et al., 2006; Sedor, 2002), which may be expressed in a positively bold forecast.

The optimism literature implies asymmetry in analysts' reluctance to issue positively or negatively bold forecasts. Although

pollution can keep analysts from fully synthesizing the information in earnings, the benefits arising from positively bold forecasts (Hong and Kubik, 2003; Ke and Yu, 2006; Michaely and Womack, 1999; O'Brien et al., 2005), can potentially override the reticence of analysts to make positively bold forecasts in the face of pollution.

Issuing negatively bold forecasts is arguably more costly for an analyst than issuing positively bold forecasts since negative information is relatively penalized. Because of the additional risk to the analyst associated with providing a negatively bold forecast, analysts will only issue negatively bold forecasts if results of extensive information processing substantiate the decision. If high PM levels affect the analyst's ability to synthesize information, the analyst is less likely to have processed information sufficiently to substantiate a negatively bold forecast. We summarize the hypothesis as follows:

H2b. *The negative association between PM pollution and forecast boldness for forecast revisions issued within a short window after earnings announcements is stronger for negatively bold forecasts than for positively bold forecasts.*

4. Research design

4.1. Model specification

To test the timeliness and accuracy of analyst forecasts, in model (1) we follow several recent studies and apply a linear probability model to estimate the probability of the analyst (i) issuing earnings forecasts in response to quarterly earnings announcements (j) (DeHaan et al., 2017; Hanlon and Hoopes, 2014; Hoberg et al., 2014).⁷

$$[TL, ACCURACY] = \alpha_0 + \alpha_1 LOG_AQI + \delta X + \theta_j + \lambda_q + QTR_t + \varepsilon \quad (1)$$

where TL equals 1 if an analyst revises his forecasts within two days (i.e. day 0, which is the earnings announcement date and day 1, which is the following day) of a quarterly earnings announcement of firm j , 0 otherwise (Zhang, 2008).⁸ Following Clement and Tse (2005) and Keskek et al. (2014), we define $ACCURACY$ based upon whether an analyst's current forecast improves his own forecast accuracy surrounding earnings announcements. $ACCURACY$ equals 1 if the analyst's revised earnings forecast accuracy is higher than his own prior earnings forecast accuracy, 0 otherwise.⁹ AQI equals the average daily air quality index between the earnings announcement date and the analyst forecast revision date if an earnings forecast is issued within two days of the earnings announcement date. If the earnings forecast is not issued within this short window, we measure AQI over the two-day window (DeHaan et al., 2017). In our study, AQI is measured for each analyst, according to the air pollution level in his own location (city). Higher values of AQI correspond to higher levels of air pollution. Following Li et al., 2019, we use the logarithm of AQI to ensure that the outliers do not disproportionately affect the results. θ_j and λ_q represent firm and year fixed effects respectively, which remove firm-invariant and time-invariant characteristics affecting the likelihood of analysts to issue timely forecasts. QTR_t denotes indicator variables for earnings announcement quarters. Negative coefficients of LOG_AQI are consistent with our expectation that analyst issuance of forecasts is postponed (H1a) and that forecasts issued within the short window of earnings announcements are less likely to represent improvements in accuracy (H2a) when analysts face PM pollution.

The vector X consists of a battery of both analyst- and firm-level control variables. All analyst-level control variables are defined in the year preceding the earnings announcement dates, unless stated otherwise. N_FIRM measures the number of firms the analyst follows. N_IND measures the number of unique industries the analyst follows. $STAR$ equals 1 if the analyst was awarded the title of star analyst by the *NEW FORTUNE* magazine in the preceding calendar year, otherwise 0. $SIZE_BROKER$ measures the number of analysts employed by the analyst's brokerage. EXP is the number of quarters between the analyst's first forecast report date recorded in the database and the current earnings announcement date. N_FOLLOW is the number of analysts who issued at least one report for the firm.

The remaining control variables are at the firm-level. $HOLDPERCT$ is the percentage of shares held by funds. LEV is total liabilities deflated by total assets. $SIZE$ is the logarithm of total market capitalization on the day before the earnings announcement. MB equals

⁷ Throughout the paper, we employ linear probability model (LPM) rather than non-linear model with limited dependent variable (LDV) where the dependent variables are dichotomous (Angrist and Pischke, 2008). There are two advantages in our setting to use LPM: 1) Greene (2004) shows that including fixed effects in a LDV model can severely bias both coefficient and standard error estimators when there is a small 'T' in the panel data (a unique firm only exists for a short period) and this bias persists even for 'T' up to 20. Our sample period spans from 2009 to 2016 which probably suffers from the small 'T' problem; and 2) the linear probability model allows easier interpretation of the coefficient of our interested variable than the non-linear LDV model because the coefficient directly indicates the marginal effect of our interested variable. In addition, to ensure that LPM fits the data well, following Hanlon and Hoopes (2014), we check the range of the fitted values of dependent variables and find that < 5% predicted values of the dependent variables fall outside of [0,1].

⁸ To define the TL measure, we require an interval spanning from the quarter t earnings announcement date to the day before the quarter $t+1$ earnings announcement date as the period where analysts could provide forecasts. We assign 1 to those forecasts issued on the day and one day after quarter t earnings announcement date (the two-day window), and 0 to remaining forecasts. The mean (median) value of this interval is 100 (98) days.

⁹ This measure does not build on the assumption that an analyst accesses peer estimates or that he puts a weight on peer estimates. Earnings forecast accuracy is the distance between the forecasted earnings and actual realized earnings. If one earnings forecast shortens (prolongs) this distance compared with the prior forecast, we assign 1 (0) to the variable $ACCURACY$. This variable is only defined for forecasts issued on or one day after earnings announcement dates. This definition is advantageous since it removes analyst-firm paired fixed effects.

total market capitalization plus the book value of total debt deflated by the book value of total assets. *SALES_GROWTH* is quarter-to-quarter sales revenue difference deflated by the sales revenue for the same quarter last year. *ROA* equals net profit divided by total assets. *LOSS* equals 1 if *ROA* is < 0, otherwise 0. *SUE* is the change in net profit from quarter-to-quarter net profit deflated by the market capitalization on the day before the earnings announcement date.¹⁰ *TRVOL* is the average trading volume over the $[-1, +1]$ window of the three prior earnings announcement dates, scaled by the average trading volume in the non-earnings announcement days for the same three quarters. *RETVAR* is the standard deviation of daily stock returns over the $[-1, +1]$ window of the three earnings announcement dates minus the standard deviation of daily stock returns over non-earnings announcement days for the same three quarters. We cluster by firm to correct for firm-level correlation.¹¹ Given that prior research documents a strong association between forecast horizon and analyst forecast accuracy, we also include *DAYS* and forecast period fixed effects (*FPI*) when the dependent variable is *ACCURACY*. Specifically, *DAYS* is the number of days between the analyst forecast revision date and the forecast period end date. *FPI* denotes the accounting period for which an earnings forecast is made. For example, *FPI* equal to 0 (1) represents that the earnings revision is made for the current (next) accounting period.

To test the effects of air pollution level on the propensity of analysts to issue bold forecasts during the high pollution period, we estimate a linear probability model:

$$[BOLD, NEG_BOLD, POS_BOLD] = \alpha_0 + \alpha_1 LOG_AQI + \delta X + \theta_j + \lambda_q + QTR_t + FPI + \varepsilon \quad (2)$$

Fig. 1 depicts the timeline to construct the three dependent variables in model (2). Before revising earnings forecasts in response to earnings announcements, an analyst may have two reference estimates to start with: 1) his own most recent forecast, and 2) the consensus forecasts issued by all peer analysts in the year preceding his forecast revision (constructed using all peer analysts' latest forecasts). To address the potential for stale information contained in the two reference estimates, we require that the analyst's prior forecast and the peer consensus forecast must be issued within one year preceding earnings announcements. Following Clement and Tse (2005), we define *BOLD* as 1 if an analyst revises his forecast below or above both reference estimates, 0 otherwise. *NEG_BOLD* equals 1 if the analyst's revised earnings forecast is below his own prior forecast and peer consensus forecasts, otherwise 0. *POS_BOLD* equals 1 if the analyst's revised earnings forecast is above his own prior forecast and peer consensus forecasts, otherwise 0. All control variables have been previously defined. To be consistent with H2a and H2b, we expect negative α_1 s when the dependent variables are *BOLD* and *NEG_BOLD*, respectively.

4.2. Data source

4.2.1. Analyst location

We extract the list of analyst names recorded at the end of each reporting date by CSMAR for all brokers in China. We search each analyst name in the HUIBO terminal (<http://www.hibor.com.cn>) to find all analyst reports issued by the analyst.¹² These reports include office landline phone numbers. The first three/four digits of landline phone numbers are area codes, which can be used to identify the analyst's location (city). Given that each analyst issues multiple forecast reports in a year, it is difficult for us to search each single analyst report. Instead, we assign the opening and closing dates for each analyst as the 1st of January and 31st December respectively. In each year, we search two reports, the one issued after, but closest to, 1st January (early report) and the one issued before, but closest to, 31st December (late report). We assign the early report to 1st January and the late report to 31st December in each year and collect the two office landline phone numbers appearing on the two reports. To be included in our analysis, we require that the analyst-year observations do not exhibit changes in their phone number's area code within a single year.¹³

4.2.2. Air quality index

We download the daily air quality index (*AQI*) from the website of Ministry of Environmental Protection of the People's Republic of China (<http://www.mep.gov.cn/>). The *AQI* has been archived on this website for major cities in China since 2010. Hourly air quality data, similar to the data used in previous literature from other countries, is unavailable for China, so we employ daily air quality data in our study.

4.3. Sampling

We retrieve analyst forecast data from the CSMAR and merge it with quarterly earnings announcements. Based on the forecast revision date, each firm-quarter-analyst-report-forecast period observation is between the announcement date of quarter t and quarter $t + 1$. This process leads to an initial sample of 592,706 firm-quarter-analyst-report-forecast period (254,621 firm-quarter-analyst-report) observations. We first drop the forecasts for a forecast period issued by an analyst after his first forecast for the same period that is in response to quarter t earnings announcement (124,091 observations). We delete observations without location information (149,036 observations) or observations lacking *AQI* information (80,281 observations). The missing analyst and firm

¹⁰ We follow a random walk assumption for the earnings generation process (Ball and Brown, 1968; Bernard and Thomas, 1989).

¹¹ When we cluster by firm and analyst or firm and city-quarter, results remain qualitatively the same.

¹² HUIBO terminal has archived a large volume of analyst reports (more than 1 million) which are widely used by investors in the Chinese market.

¹³ We drop analysts with changes in phone numbers from our sample because it is very difficult to determine on which date they changed location.

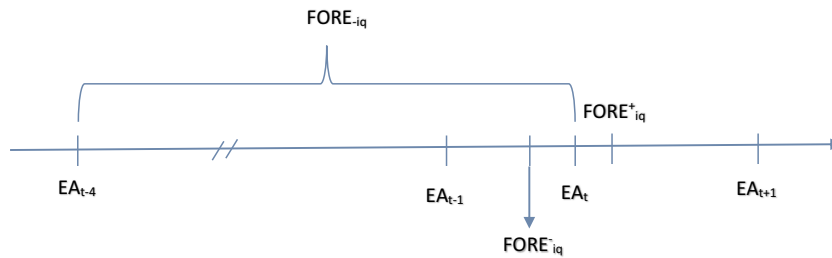


Fig. 1. Timeline for the research design.

Notes: This figure presents the timeline for our research design. EAs represent the earnings announcement dates for quarterly earnings. $FORE_{iq}^+$ is analyst i 's forecast for year q in response to the earnings announcement of quarter t . $FORE_{iq}^-$ is analyst i 's latest forecast for year q before the earnings announcement of quarter t . $FORE_{iq}^-$ is the consensus forecast for year q before the earning announcement of quarter t by all analysts except analyst i . To calculate this consensus forecast, we keep all forecasts for year q issued between the earnings announcement of quarter $t-4$ and that of quarter t . If the analyst issues multiple forecasts during this period for year q , we retain only the latest forecast.

level control variables reduce the sample by 1504 and 30,941 observations respectively. We also drop observations with missing information in calculating accuracy improvement or boldness (114,450 observations). The resulting sample size comprises 92,403 firm-quarter-analyst-report-forecast period observations, which corresponds to 43,195 firm-quarter-analyst-reports.¹⁴ After excluding 914 singletons from the firm-quarter-analyst-report sample, the remaining 42,281 observations are used to test the H1a, consisting of 1982 firms, 1487 analysts, 16,699 firm quarters and 9593 analyst-quarters.¹⁵

Within the 92,403 observations, there are 59,640 firm-quarter-analyst-report-forecast period observations generated by analysts on the day of and one day after earnings announcement dates. We drop 333 singletons, arriving at a final sample of 59,307 observations, which we use to examine H1b, H2a and H2b. This sample includes 1860 unique firms, 1483 analysts, 12,733 firm-quarter earnings announcements and 7488 analyst-quarters. Table 1 summarizes the above sampling procedures.

Table 2 describes the distribution of analysts and air pollution across cities. Panel A indicates that analysts in our sample are spread over 19 cities in China. Shanghai is the location for the largest number of analysts, accounting for 45.12% of our sample, with Beijing, Shenzhen, Nanjing, Guangzhou and Wuhan also hosting a large number of analysts. The skewed distribution of analysts (a few cities hosting the majority of analysts) raises two challenges to our research design. First, the analyst level AQI may exhibit strong positive correlations at the city level, which may downwardly bias the estimated standard errors (Petersen, 2009). To deal with this challenge, we include analyst fixed effect in our main models to mitigate the city level correlation.¹⁶ In addition, later in robustness tests, we cluster the standard errors by city. Second, a firm-quarter-forecast period observation might be followed by multiple analysts from the same city. In other words, there is no variation within the firm-quarter-forecast period observation. Lack of such variation may mean that any observed differences in dependent variables are simply due to differences across firm-quarter observations. To address this concern, we employ diff-in-diff models by constructing a sample in which each firm-quarter-forecast period observation is followed by multiple analysts from cities with varying AQI values.

Panel B Table 2 summarizes variation in PM levels in the top 6 cities in China (i.e., Shanghai, Beijing, Shenzhen, Nanjing, Guangzhou, and Wuhan) during our sample period of 2009–2016. Overall, there is a large degree of variation in the amount of pollution across the cities where analysts are located. During the six years of our sample period, there are 158 days with AQI between 50 and 100 in all six cities simultaneously; 209 days have a high AQI in only one city. There are only 9 days with AQI index higher than 100 simultaneously in all six cities, but 701 days with such a high AQI in only one city. During our sample period, there are 435 (1477), 899 (1298), 68 (1254), 697 (1611), 258 (1594), and 800 (1482) days with AQI higher than 100 (higher than 50 but no higher than 100) in Shanghai, Beijing, Shenzhen, Nanjing, Guangzhou, and Wuhan respectively.

5. Empirical results

5.1. Summary statistics

Table 3, panel A provides summary statistics for the dependent variables in models (1) and (2) and for control variables in model (1).¹⁷ 55.52% of analyst forecasts are issued with the two-day window of earnings announcements in China, which is higher than the corresponding fraction (40.32%) in the U.S. market (Zhang, 2008). Analysts deviate from both their own prior forecasts and from

¹⁴ Similar to in the US, analysts in China issue several period forecasts in one report.

¹⁵ Correia (2015) shows that the existence of singletons may bias estimation of standard errors. Singleton groups refers to groups with only one observation. In our setting, adding analyst fixed effects when there is an analyst with only one observation can lead to over-stated significance of the coefficients.

¹⁶ We do not cluster standard errors by city in our main research design is because the number of cities is too small (lower than 40) to correct the correlation within cities according to Petersen (2009).

¹⁷ We do not tabulate the summary statistics for control variables in model (2) is because they are not materially different those reported for model (1).

Table 1
Sampling process.

	# Firm-quarter-analyst-report-forecast period	# Firm-quarter-analyst-report
Analyst forecasts merged with the quarterly earnings announcement dates (forecast period end between 31/12/2010–31/12/2016)	592,706	254,621
Less: forecasts issued after an analyst's first revision in response to the current quarter earnings announcement	124,091	51,330
Less: observations without location information	149,036	73,550
Less: observations without air quality information	80,281	30,927
Less: observations without analyst level control variables	1504	574
Less: observations without firm level control variables	30,941	12,568
Less: observations without accuracy improvement/boldness measure	114,450	42,477
Remaining forecasts issued between two quarterly announcements	92,403	43,195
Forecasts issued on and one day after earnings announcement dates	59,640	
Less: singletons	333	914
Remaining forecasts	59,307	42,281
Firm observations	1860	1982
Firm quarter announcements	12,733	16,699
Analyst observations	1484	1487
Analyst quarter observation	7488	9593

Table 2
Statistics of the variation in analysts and *AQ* across locations.

Panel A Location distribution of analysts in our sample.					
City	No. of observations	% of observations	City	No. of observations	% of observations
Beijing	496	29.07	Shanghai	770	45.13
Changsha	11	0.64	Shenzhen	262	15.36
Dalian	2	0.12	Suzhou	6	0.35
Dongguan	8	0.47	Taiyuan	8	0.47
Guangzhou	24	1.41	Tianjin	12	0.70
Hangzhou	5	0.29	Wuxi	13	0.76
Hefei	4	0.23	Wuhan	15	0.88
Jinan	2	0.12	Xian	3	0.18
Nanjing	51	2.99	Chongqing	13	0.76
Nanning	1	0.06	Total	1706	100

Panel B Statistics of extreme event days with high <i>AQ</i> across top 6 cities during the period of 2009–2016 in China.		
	50 < <i>AQ</i> ≤ 100	<i>AQ</i> > 100
# Days in 6 cities	158	9
# Days in 5 cities	393	57
# Days in 4 cities	659	123
# Days in 3 cities	649	271
# Days in 2 cities	459	394
# Days in 1 city	209	701
# Days in Shanghai	1477	435
# Days in Beijing	1298	899
# Days in Shenzhen	1254	68
# Days in Nanjing	1611	697
# Days in Guangzhou	1594	258
# Days in Wuhan	1482	800

Panel A of this table displays the location distribution of analysts in our sample observations. Panel B presents the numbers of extreme event days with high *AQ* during our sample period of 2009–2016 across top 6 cities in terms of the number of analysts in China.

other peer consensus forecasts for 49.57% of the forecasts, which is much lower than the ratio reported by [Clement and Tse \(2005\)](#) in the US. Most of the bold forecasts (0.3182/0.4957 = 64%) are negatively bold, in which analysts downwardly revise their forecasts. Relative to analysts' immediately prior forecast, fewer than half (mean of *ACCURACY* = 45.41%) of the forecasts provided in a short window immediately following earnings announcements represent a reduction in forecast errors.¹⁸ Mean and median *AQ* are 77 and

¹⁸ In our sample, analysts do not update their forecasts for 30.08% of their available opportunities. Taken together, on average analysts increase their forecast accuracy over time.

Table 3
Summary statistics.

Panel A: Descriptive statistics for all variables							
Variables	Mean	Std.	P25	P50	P75	N	
<i>TL</i>	0.5552	0.4970	0	1	1	42,281	
<i>ACCURACY</i>	0.4551	0.4980	0	0	1	59,307	
<i>BOLD</i>	0.4957	0.5000	0	0	1	59,307	
<i>NEG_BOLD</i>	0.3182	0.4658	0	0	1	59,307	
<i>POS_BOLD</i>	0.1774	0.3821	0	0	0	59,307	
<i>AQI</i>	76.6295	39.9647	52	67	91	42,281	
<i>LOG_AQI</i>	4.2286	0.4623	3.9512	4.2047	4.5109	42,281	
<i>N_FIRM</i>	2.2845	0.8023	1.7918	2.3026	2.8332	42,281	
<i>N_IND</i>	1.2226	0.4825	0.6931	1.0986	1.6094	42,281	
<i>SIZE_BROKER</i>	3.6204	0.5388	3.3322	3.6889	3.9703	42,281	
<i>EXP</i>	2.8090	0.6464	2.4849	2.8904	3.2958	42,281	
<i>N_FOLLOW</i>	2.8549	0.6380	2.4849	2.9444	3.3322	42,281	
<i>STAR</i>	0.2564	0.4366	0	0	1	42,281	
<i>DAYS</i>	6.0842	0.6168	5.5053	6.2500	6.5944	59,307	
<i>HOLDPERCT</i>	0.0637	0.0748	0.0074	0.0364	0.0933	42,281	
<i>ROA</i>	0.0440	0.0400	0.0145	0.0334	0.0626	42,281	
<i>LEV</i>	0.4584	0.2271	0.2776	0.4510	0.6264	42,281	
<i>SALES_GROWTH</i>	0.2875	0.4579	0.0608	0.2022	0.3846	42,281	
<i>TRVOL</i>	1.1633	0.3254	0.9396	1.1200	1.3388	42,281	
<i>RETVAR</i>	0.0012	0.0091	-0.0044	0.0007	0.0067	42,281	
<i>MB</i>	2.7472	1.8728	1.3870	2.1779	3.4588	42,281	
<i>SUE</i>	0.0046	0.0174	0.0002	0.0035	0.0092	42,281	
<i>SIZE</i>	16.4226	1.1397	15.6016	16.2346	17.0745	42,281	
<i>ANNUAL</i>	0.1866	0.3896	0	0	0	42,281	
<i>LOSS</i>	0.0295	0.1692	0	0	0	42,281	
<i>WIND</i>	2.0014	2.0504	0	2	4	37,499	
<i>CLOUD_RAIN</i>	0.5946	0.4910	0	1	1	37,507	

Panel B: Forecast timeliness, boldness and accuracy improvement across AQI quintiles						
AQI quintile	Mean_LOG_AQI	<i>TL</i>	<i>ACCURACY</i>	<i>BOLD</i>	<i>NEG_BOLD</i>	<i>POS_BOLD</i>
0	3.5927	0.6210	0.4835	0.5270	0.3471	0.1799
1	4.0103	0.5535	0.4656	0.5109	0.3344	0.1765
2	4.2058	0.5854	0.4474	0.4928	0.3157	0.1771
3	4.4404	0.5355	0.4355	0.4670	0.2947	0.1723
4	4.8938	0.4805	0.4436	0.4806	0.2993	0.1813
Q4-Q0		-0.1405*	-0.0399***	-0.0464***	-0.0478***	0.0014
t-statistic		-18.5567	-6.1629	-7.1492	-7.8810	0.2869

Panel A of this table presents summary statistics of all variables used in the paper. Panel B presents the means of all four dependent variables across AQI quintiles, with quintile 0 (1) representing the lowest (highest) PM pollution observations. ***, **, * represent statistical significance at the 1%, 5%, 10% level (two-tailed), respectively. See the appendix for variable definitions.

67, which is highly comparable to those reported by Li et al., 2019, although those papers cover a slightly different period.¹⁹

Regarding control variables, a representative Chinese analyst follows 12 firms and 3 industries, which is lower than his peers in the U.S. (Clement and Tse, 2005). On average, a brokerage employs 37 ($e^{3.6204}$) analysts. Each firm-year observation has an average following of 17 analysts. Mean and median *EXP* are 17 and 18 quarters respectively. 26% forecasts are issued by star analysts. Mean *TRVOL* is 16% higher during earnings announcement periods than during non-earnings announcement days. However, stock return volatility (*RETVAR*) does not appear to be affected by earnings announcements. 18.66% of earnings announcements are related to annual earnings. The fraction of loss firms is only 2.95%, which is much lower than in the U.S.

Table 3, panel B presents the mean statistics for all four dependent variables across *LOG_AQI* quintiles, with quintile 0 (1) representing the lowest (highest) PM pollution periods. There is a generally declining trend for each of the four variables from the least to most polluted periods.²⁰ For instance, 62.10% of analyst forecasts are issued within the two-day window of earnings announcements in the least polluted period, whereas this percentage drops to only 48.05% in the most polluted period. Compared with the lowest AQI quintile, where 48.35% of analyst forecasts experience improvement in accuracy; accuracy improvements for analysts in the highest AQI quintile are much lower (44.36%). Mean *BOLD* and *NEG_BOLD* are 52.70% (48.06%) and 34.71% (29.93%) in the lowest (highest) AQI quintile respectively. Positively bold forecasts do not show a similar pattern, however, with no discernable

¹⁹ Mean and median AQI in Li et al. (2019) are 80 and 70 respectively.

²⁰ The generally monotonic trend provides support for our linear model specification.

differences in their incidence across AQI quintiles. The differences in *TL*, *ACCURACY*, *BOLD* and *NEG_BOLD* between the lowest and highest quintiles are significant, not only statistically (all with $p < .0001$) but also economically, with a difference from the mean between the extreme quintiles of 14%, 4%, 5%, and 5% respectively.

5.2. Multivariate regression results

5.2.1. Air pollution and analyst forecast properties

Table 4 reports results for estimating model (1) where we examine how PM pollution affects the timeliness and accuracy of analyst forecasts. Column (1) Panel A presents the base model, which does not include any control variables or fixed effects. For this model, the coefficient of *LOG_AQI* is -0.0975 and is highly significant ($p < .0001$). The model reported in column (2) includes firm, year and quarter fixed effects. In column (3), we include additional control variables. Column (4) presents results where we remove the effects of analyst invariant characteristics (e.g., gender and education background). None of these additions results in large changes to the coefficient of *LOG_AQI* and the coefficients are all highly significant ($p < .0001$). In column (5), when we control for weather effects, the coefficient of *LOG_AQI* remains negative and significant. All models provide results consistent with the level of PM pollution being significantly negatively associated with analysts' propensity to issue forecasts within the two-day window of earnings announcements. The coefficients of the two weather related variables suggest that bad weather is negatively related to analysts' likelihood to revise earnings forecasts immediately after earnings announcements, which is consistent with the findings in prior literature (DeHaan et al., 2017), but does not affect our inferences regarding air pollution.

Our results are economically significant. Based upon the magnitude of the coefficients of all independent variables reported in column (5) of Panel A, we first set *LOG_AQI* to its 75th percentile value, other continuous variables to their mean values and all indicator values to be zeros. The estimated baseline probability of issuing timely forecasts is 62.32%. When we change the *LOG_AQI* to its 25th percentile value and keep all other variables constant, the re-estimated probability of issuing timely forecasts becomes 68.42%. As a result, an increase in *LOG_AQI* from its first to third quartile leads to a 6.10% (68.42%-62.32%) decrease in the probability of issuing timely forecasts (*TL*). In contrast, an increase in *WIND* from its first to third quartile leads to a 0.60% decrease in the probability of issuing timely forecasts. The coefficient of *CLOUD_RAIN* indicates that compared with a sunny day, analysts are 11.25% less likely to issue a timely forecast revision on a cloudy/rainy day. We note that these results are not exactly comparable, since *CLOUD_RAIN* is a dichotomous variable and *AQI* is continuous.²¹ Overall, we find evidence supportive of H1a, that high PM pollution reduces analysts' propensity to issue timely forecasts. More importantly, this effect is distinct from the previously documented weather effect.

Turning to the control variables, broker size, the number of analysts following the firm, the number of firms followed by the analyst, firm profitability and firm size are positively associated with timely forecasts. Leverage, earnings surprise, and growth prospects are negatively associated with timely forecasts.

Panel B of Table 4 provides evidence consistent with H1b, that analysts are less likely to improve their forecast accuracy in periods with high PM pollution. As before, we first report a model excluding controls and then add different levels of control variables. Across all models, the coefficient of *ACCURACY* is negative and statistically significant. Analyst coverage, firm size, earning surprise, and firm growth prospect are negatively associated with analyst tendency to improve their forecast accuracy. Analyst experience, analyst following, and early forecasts are positively associated with the presence of improvement in forecast accuracy. These results are robust to analyst fixed effects and weather-related variables. In particular, we fail to find significant results for either weather-related variable, suggesting that weather conditions do not affect improvements in forecast accuracy. Following the aforementioned approach, the estimated baseline probability of experiencing improvement in forecast accuracy is 46.11%. When we decrease the *LOG_AQI* from its 75th to 25th percentile value, the probability increases to 48.63%.

5.2.2. Air pollution and analyst information production

Table 5 presents the results of estimating model (2) for our tests of H2a, that there is a negative association between PM pollution and forecast boldness for forecast revisions issued within a short window after earnings announcements. Panel A presents results for *BOLD* as the dependent variable. As before, we first estimate the models without control variables and in consecutive models, add different types of control variables. In all cases, the coefficient of *LOG_AQI* is negative and significant. Column (3) reports results including firm fixed effects, where the coefficient of *LOG_AQI* is -0.0169 ($p < 0.05$). Further results indicate that analysts with wide coverage are less likely to issue bold forecasts. Analyst experience, analyst following, and early forecasts (relative to the forecast period end) are positively associated with the presence of bold forecasts. Firm size, earnings surprise, growth prospect, and institutional shareholdings all exhibit negative associations with the presence of bold forecasts. These results are robust to inclusion of both analyst fixed effects and weather-related control variables (reported in column (5)). With respect to economic significance, an increase in *LOG_AQI* from the first to the third quartile results in a 3% decrease in the proportion of forecasts that are bold.²² Results fail to provide evidence that weather conditions are associated with analyst forecast boldness around earnings announcements.

²¹ The coefficient of the dichotomous variable *CLOUD_RAIN* may be driven by the effect of extreme poor weather days. If we treat *CLOUD_RAIN* as a 'continuous' variable and assess the economic effects, one standard deviation change in *CLOUD_RAIN* triggers 11.11% standard deviation change in *TL*. As a result, we interpret the effect of air pollution to be as large as that of weather in affecting analysts' issuance of timely forecasts.

²² When we investigate the relation between accuracy and boldness of forecasts, we find that air pollution not only directly reduces the likelihood of bold forecasts, but, also impairs the accuracy of the bold forecasts.

Table 4
Impact of air pollution on analyst forecast timeliness and accuracy.

Panel A AQI and timeliness of analyst forecasts (H1a)					
Variables	(1)	(2)	(3)	(4)	(5)
	TL	TL	TL	TL	TL
LOG_AQI	-0.0975*** (-16.89)	-0.0894*** (-14.28)	-0.0868*** (-13.77)	-0.0966*** (-14.63)	-0.1095*** (-15.05)
LOG_N_FIRM			0.0490*** (9.05)	0.0622*** (8.44)	0.0542*** (6.71)
LOG_N_IND			-0.0112 (-1.18)	0.0115 (0.93)	0.0100 (0.74)
LOG_SIZE_BROKER			0.0309*** (5.78)	0.0269** (1.99)	-0.0115 (-0.79)
LOG_EXP			-0.0057 (-1.31)	0.0492*** (2.70)	0.0372* (1.92)
LOG_N_FOLLOW			0.0272*** (2.88)	0.0295*** (4.50)	0.0280*** (4.20)
STAR			-0.0064 (-0.95)	-0.0337*** (-2.65)	-0.0408*** (-2.94)
HOLDPERCT			0.0363 (0.50)	0.0443 (0.83)	-0.0119 (-0.21)
ROA			0.3527** (2.12)	0.2538** (2.12)	0.1818 (1.43)
LEV			-0.1347*** (-3.13)	-0.0394* (-1.84)	-0.0373* (-1.73)
SALES_GROWTH			0.0018 (0.22)	0.0114* (1.88)	0.0161** (2.48)
LOSS			-0.0006 (-0.03)	-0.0215 (-1.28)	-0.0266 (-1.53)
SUE			-0.7189*** (-3.58)	-0.3998** (-2.15)	-0.5149** (-2.43)
SIZE			0.0707*** (5.55)	-0.0032 (-0.67)	-0.0017 (-0.35)
MB			-0.0218*** (-5.25)	-0.0030 (-1.31)	-0.0033 (-1.38)
TRVOL			0.0159 (1.42)	0.0016 (0.16)	-0.0048 (-0.45)
RETVAR			-0.1097 (-0.27)	-0.0140 (-0.04)	0.3795 (1.04)
WIND					-0.0039* (-1.85)
CLOUD_RAIN					-0.1125*** (-19.21)
Observations	42,281	42,281	42,281	42,281	37,465
Year	2010–2016	2010–2016	2010–2016	2010–2016	2010–2016
R-squared	0.008	0.138	0.145	0.217	0.235
Firm FE	NO	YES	YES	NO	NO
Year FE	NO	YES	YES	YES	YES
Quarter FE	NO	YES	YES	YES	YES
Analyst FE				YES	YES

Panel B AQI and analyst forecast accuracy (H1b).

Variables	(1)	(2)	(3)	(4)	(5)
	ACCURACY	ACCURACY	ACCURACY	ACCURACY	ACCURACY
LOG_AQI	-0.0340*** (-5.42)	-0.0212*** (-3.16)	-0.0216*** (-3.21)	-0.0175** (-2.31)	-0.0246*** (-2.94)
LOG_N_FIRM			-0.0110** (-1.96)	-0.0315*** (-3.52)	-0.0340*** (-3.49)
LOG_N_IND			-0.0086 (-0.88)	-0.0033 (-0.22)	-0.0014 (-0.09)
LOG_SIZE_BROKER			-0.0007 (-0.12)	-0.0225 (-1.41)	-0.0330* (-1.95)
LOG_EXP			0.0120*** (2.63)	0.0423** (1.97)	0.0537** (2.12)
LOG_N_FOLLOW			0.0536***	0.0149**	0.0117

(continued on next page)

Table 4 (continued)

Panel B AQI and analyst forecast accuracy (H1b).					
Variables	(1)	(2)	(3)	(4)	(5)
	ACCURACY	ACCURACY	ACCURACY	ACCURACY	ACCURACY
			(4.65)	(2.01)	(1.46)
STAR			0.0031 (0.46)	-0.0262* (-1.74)	-0.0331* (-1.93)
LOG_DAYS			0.0525*** (3.31)	0.0658*** (4.04)	0.0831*** (4.74)
HOLDPERCT			-0.1133 (-1.54)	-0.2026*** (-3.63)	-0.2039*** (-3.31)
ROA			-0.0777 (-0.40)	-0.0379 (-0.27)	0.0094 (0.06)
LEV			0.0220 (0.43)	0.0012 (0.05)	-0.0048 (-0.19)
SALES_GROWTH			-0.0306*** (-3.03)	-0.0248*** (-3.12)	-0.0309*** (-3.43)
LOSS			-0.0764*** (-2.67)	-0.0386* (-1.77)	-0.0394* (-1.78)
SUE			-2.4288*** (-8.75)	-2.7268*** (-11.07)	-2.9866*** (-11.18)
SIZE			-0.0541*** (-3.41)	-0.0042 (-0.98)	-0.0052 (-1.12)
MB			-0.0210*** (-3.39)	-0.0126*** (-4.29)	-0.0118*** (-3.78)
TRVOL			0.0049 (0.35)	0.0057 (0.48)	0.0169 (1.35)
RETVAR			0.4678 (0.96)	0.1332 (0.33)	-0.1892 (-0.43)
WIND					-0.0006 (-0.24)
CLOUD_RAIN					0.0037 (0.55)
Observations	59,307	59,307	59,307	59,307	51,828
Year	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016
R-squared	0.001	0.114	0.125	0.151	0.159
Firm FE	NO	YES	YES	NO	NO
Year FE	NO	YES	YES	YES	YES
Quarter FE	NO	YES	YES	YES	YES
FPI FE	NO	YES	YES	YES	YES
Analyst FE				YES	YES

This table presents OLS regression results using *TL* and *ACCURACY* as the dependent variable in Panel A and B respectively. In all regression models, we log transform *AQI*, and all continuous analyst level control variables. In column (1), we do not control any variables or fixed effects. In column (2), we add firm, year and quarter fixed effects to the model. Column (3) presents the full model where we further control both analyst and firm level control variables. In column (4), we replace firm fixed effects with analyst fixed effects. In column (5), we further control for local weather conditions. All standard errors are corrected for firm level correlation. The sample period spans 2010-2016. ***, **, * represent statistical significance at the 1%, 5%, 10% level (two-tailed), respectively. See the appendix for variable definitions.

To provide evidence regarding H2b, that PM pollution has a larger effect on negatively bold rather than positively bold forecasts, we estimate Equation (2) by using *NEG_BOLD* and *POS_BOLD* as dependent variables. Panel B and panel C of Table 5 tabulate the results. The tables present results only for *LOG_AQI*, which is the variable of interest.²³ Similar to panel A, we include additional controls across the models reported in each panel. Results are consistent across all models. Focusing on column (4), the model with analyst fixed effects, results reported in panel B for negatively bold forecasts indicate a negative coefficient of *LOG_AQI* (-0.0204, $p < 0.01$), whereas in the same model, but for positively bold forecasts, column (4) in panel C reports a positive but insignificant coefficient of *LOG_AQI* (0.0052, $p > 0.1$). The economic significance of the coefficient of *LOG_AQI* is column (4) panel B suggests that an increase in *LOG_AQI* from its first to third quartile value leads to a 5% decrease in the proportion of forecasts that are negatively bold. When we compare the coefficients of *LOG_AQI* between columns (4) of Panel B and C, the difference is significant with $p < 0.01$. The difference in the two coefficients provides evidence that the negative effect of PM pollution on analysts' tendency to issue bold forecasts manifests in the effect of PM pollution on negatively bold forecasts.

²³ Complete results are available from the authors.

Table 5
Boldness as a channel underlying the impact of air pollution on analyst forecasts

Panel A: Air pollution and forecast boldness					
Variables	(1)	(2)	(3)	(4)	(5)
	<i>BOLD</i>	<i>BOLD</i>	<i>BOLD</i>	<i>BOLD</i>	<i>BOLD</i>
<i>LOG_AQI</i>	-0.0385*** (-5.95)	-0.0152** (-2.21)	-0.0169** (-2.49)	-0.0151** (-2.00)	-0.0183** (-2.20)
<i>LOG_N_FIRM</i>			-0.0178*** (-3.05)	-0.0487*** (-5.39)	-0.0544*** (-5.48)
<i>LOG_N_IND</i>			-0.0082 (-0.78)	0.0047 (0.32)	0.0089 (0.53)
<i>LOG_SIZE_BROKER</i>			-0.0030 (-0.48)	-0.0194 (-1.21)	-0.0312* (-1.80)
<i>LOG_EXP</i>			0.0099** (2.15)	0.0522** (2.57)	0.0516** (2.16)
<i>LOG_N_FOLLOW</i>			0.0706*** (6.06)	0.0386*** (5.14)	0.0342*** (4.25)
<i>STAR</i>			-0.0061 (-0.82)	-0.0174 (-1.17)	-0.0324* (-1.92)
<i>LOG_DAYS</i>			0.0769*** (5.87)	0.0867*** (6.26)	0.0924*** (6.08)
<i>HOLDPERCT</i>			-0.1687** (-2.37)	-0.1504** (-2.27)	-0.1350* (-1.84)
<i>ROA</i>			0.1956 (1.06)	0.2790** (2.02)	0.1436 (0.95)
<i>LEV</i>			-0.0715 (-1.41)	-0.0435* (-1.88)	-0.0517** (-2.17)
<i>SALES_GROWTH</i>			0.0022 (0.21)	0.0090 (1.14)	0.0035 (0.40)
<i>LOSS</i>			0.0193 (0.75)	0.0202 (1.00)	0.0116 (0.55)
<i>SUE</i>			-2.0594*** (-8.02)	-2.3923*** (-10.27)	-2.7359*** (-10.69)
<i>SIZE</i>			-0.0482*** (-3.17)	-0.0029 (-0.64)	-0.0008 (-0.16)
<i>MB</i>			-0.0145** (-2.55)	-0.0158*** (-5.00)	-0.0153*** (-4.59)
<i>TRVOL</i>			0.0054 (0.43)	0.0068 (0.64)	0.0051 (0.44)
<i>RETVAR</i>			0.2375 (0.49)	0.3701 (0.89)	0.2720 (0.61)
<i>WIND</i>					-0.0011 (-0.40)
<i>CLOUD_RAIN</i>					0.0042 (0.64)
Observations	59,307	59,307	59,307	59,307	51,828
Year	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016
R-squared	0.001	0.126	0.134	0.174	0.183
Firm FE	NO	YES	YES	NO	NO
Year FE	NO	YES	YES	YES	YES
Quarter FE	NO	YES	YES	YES	YES
FPI FE	NO	YES	YES	YES	YES
Analyst FE				YES	YES

Panel B: Air pollution and negative boldness					
Variables	(1)	(2)	(3)	(4)	(5)
	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>
<i>LOG_AQI</i>	-0.0354*** (-4.93)	-0.0205*** (-2.91)	-0.0226*** (-3.40)	-0.0204*** (-2.75)	-0.0286*** (-3.46)
CONTROLS	NO	NO	YES	YES	YES
Observations	59,307	59,307	59,307	59,307	51,828
Year	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016
R-squared	0.001	0.159	0.204	0.219	0.229
Firm FE	NO	YES	YES	NO	NO

(continued on next page)

Table 5 (continued)

Panel B: Air pollution and negative boldness					
Variables	(1)	(2)	(3)	(4)	(5)
	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>
Year FE	NO	YES	YES	YES	YES
Quarter FE	NO	YES	YES	YES	YES
FPI FE	NO	YES	YES	YES	YES
Analyst FE				YES	YES
Panel C: Air pollution and positive boldness					
Variables	(1)	(2)	(3)	(4)	(5)
	<i>POS_BOLD</i>	<i>POS_BOLD</i>	<i>POS_BOLD</i>	<i>POS_BOLD</i>	<i>POS_BOLD</i>
<i>LOG_AQI</i>	-0.0031 (-0.55)	0.0053 (0.96)	0.0057 (1.04)	0.0052 (0.85)	0.0103 (1.60)
CONTROLS	NO	NO	YES	YES	YES
Observations	59,307	59,307	59,307	59,307	51,828
Year	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016
R-squared	0.000	0.141	0.166	0.191	0.175
Firm FE	NO	YES	YES	NO	NO
Year FE	NO	YES	YES	YES	YES
Quarter FE	NO	YES	YES	YES	YES
FPI FE	NO	YES	YES	YES	YES
Analyst FE				YES	YES

Panel A, Panel B and Panel C presents OLS regression results using *BOLD*, *NEG_BOLD*, and *POS_BOLD* as dependent variables, respectively. In all regression models, we log transform *AQI*, and all continuous analyst level control variables. In column (1), we do not control any variables or fixed effects. In column (2), we add firm, year, quarter, and the forecast period fixed effects to the model. Column (3) presents the full model where we further control both analyst and firm level control variables. In column (4), we replace firm fixed effects with analyst fixed effects. In column (5), we further control for local weather conditions. All standard errors are corrected for firm level correlation. The sample period spans 2010-2016. ***, **, * represent statistical significance at the 1%, 5%, 10% level (two-tailed), respectively. See the appendix for variable definitions.

5.3. Difference-in-differences designs

In our setting, there is potential existence of correlated omitted variables that might explain our results. For example, migration economics suggests that people with high ability have more choices about where they work (Chiswick, 1999; Chiquiar and Hanson, 2005). In our setting, if analysts with high ability prefer to work in cities without air pollution, our findings might reflect poorer performance of low ability analysts working in polluted cities. To a certain extent, our battery of analyst attributes and analyst fixed effects address this issue. However, to provide further confidence in our results, we further employ diff-in-diff empirical designs.

Following Li et al., 2019, to reduce the threat of analyst migration in response to air pollution, we take advantage of exogenous short-term stark changes in air quality in the cities where analysts are located. Specifically, we identify analyst-firm-announcement quarter observations for which there were drastic differences in air quality during the short-window forecasting periods between two adjacent quarters (the treatment sample). To be included in this sample, the difference in *AQI* of two adjacent quarter analyst-firm forecasts must be at least 40 (one standard deviation of our full sample *AQI*). Conceptually, within this sample, an analyst-firm forecast potentially receives two types of treatment, one where there is an increase in *AQI* (polluted) and the other one where there is a decrease in *AQI* (clean). Once we identify treatment observations, the change in the timeliness and properties of analyst forecasts between the two earnings announcements should reflect the effect of the difference in air quality.

Given that the timeliness and earnings forecasts properties in our study also change in relation to other analyst-/firm-level characteristics (e.g., analyst experience), we construct a control sample to control for the effects of these characteristics on the dependent variables. For each observation in the treatment sample, we choose one observation from the remaining observations that did not experience drastic changes in *AQI*. The observation chosen must include forecasts issued by an analyst from a different city but following the same firm. This eliminates concerns about firm-level omitted variables driving our findings. For treatment sample observations with multiple control observations, we drop those observations with fundamentally different *AQI* (the difference in *AQI* must be less than 30) in the prior earnings announcement quarter. If there is still more than one control observation left, we choose the observation with the closest *AQI* across quarters. This process ensures that all changes between the prior and current quarter timeliness and forecast performance of the control sample should result from changes in analyst-/firm-level characteristics other than changes in *AQI*. The construction of the diff-in-diff design sample also mitigates concerns arising from the skewed distribution of analysts across cities (i.e. a few cities host the majority of analysts).

To assess the effect of PM pollution on our dependent variables, we compare differences in changes in forecasts of the treatment sample and changes in forecasts of the control sample between the prior and current quarter earnings announcements. For polluted

(clean) observations, we should witness a decrease (increase) in timeliness, forecast accuracy improvement, boldness, and negative boldness. The diff-in-diff model is specified as follows:

$$[TL, ACCURACY\ BOLD, NEG_BOLD] = \alpha_0 + \alpha_1 POLLUTED + \alpha_2 CLEAN + \alpha_3 POST + \alpha_4 POLLUTED * POST + \alpha_5 CLEAN * POST + X\delta + \theta_j + \lambda_q + \varepsilon \quad (3)$$

where *POLLUTED* (*CLEAN*) equals 1 if an analyst-firm observation experiences an increase (decrease) in AQI from the previous quarter forecast, otherwise 0. *POST* equals 1(0) for the current (previous) quarter forecast. To be consistent with our hypotheses, we should find negative α_4 s and/or positive α_5 s.

Table 6, panels A and B provide summary statistics for our main variables used in the diff-in-diff analysis. Panel A shows changes in both AQI and TL across firms between the preceding and current quarter announcements. There are 2,762, 910, and 1,852 analyst-firm observations in the Control, Clean, and Polluted samples respectively. Consistent with our sampling strategy, AQI for the Control sample does not differ across periods, whereas the Clean (Polluted) sample has a significantly negative (positive) change in AQI across periods. Results across the samples for TL are consistent with our earlier results. While there is an insignificant difference in TL across periods for the Control sample, there is a decrease (increase) of 12.09% (3.46%) in TL in the Clean (Polluted) sample, which is consistent with the changes in AQI experienced across the samples.

Table 6, panel B provides a similar univariate analysis, but focuses on forecast accuracy, boldness, and negative boldness. For these analyses, there are 1,316, 421 and 881 observations in the Control, Clean, and Polluted samples respectively. In the Clean (Polluted) sample, ACCURACY increases (decreases) by 11.98% (1.5%) across quarters, BOLD increases (decreases) by 6.07% (3.41%), and NEG_BOLD increases (decreases) by 10.39% (1.5%). Although not all of these results are statistically significant, the overall pattern is consistent with our expectations.

Table 6, panel C provides results of estimating model (3), including controls at the firm-level, analyst-level, and fixed effects for time and firm. In column (1), the coefficients of *POLLUTED* POST* and *CLEAN* POST* are consistent with H1a: an (a) increase (decrease) in AQI leads to less (more) timely earnings forecasts relative to control observations without any change in AQI. The remaining three columns document significant differences in response to changes in AQI across samples for forecast accuracy, boldness, and negative boldness.²⁴ Analysts experiencing cleaner air demonstrate increased forecast accuracy and analysts experiencing polluted environments provide forecasts with decreased boldness and negative boldness. Overall, the diff-in-diff design results corroborate our previous findings.

5.4. Models to address additional endogeneity concerns

Our models employ local air pollution faced by analysts to capture our variable of interest. This variable is potentially correlated with other analyst host city (both time invariant and variant) characteristics which we may not have included in our models. In the next two subsections, we discuss this possibility and our strategies to address these concerns.

5.4.1. Geographic distance between an analyst and the firm followed

Given that a firm might be covered by multiple analysts, the geographic distance between an analyst and the firm covered may contaminate our analyses. In particular, analysts who are closely located to the covered firm may have advantages in acquiring private information. These advantages may allow the analyst to issue early, accurate and bold forecasts. This possibility might compromise our tests if analysts closely located to firms are less affected by air pollution. In extreme circumstances, if only analysts remotely located to firms are subject to air pollution, then the findings we document might be largely driven by geographic distance. Our estimated models include analyses with analyst fixed effects, which mitigates this geographic distance effect if analysts and the firms covered generally do not change their locations. In our data, there is no firm changing headquarters and we have removed analysts with different area codes in their landline phone numbers within a year. In untabulated analyses, we compute the geographic distance between analysts and firms covered and include the distance as a control. Our inferences are the same as reported for our earlier analyses.

5.4.2. Other time-variant characteristics

We conduct two sets of placebo tests to address the concern that our observed results are driven by some unobservable analyst-city-time variant characteristics or time variant characteristics. Conceptually, in the first test, we randomly substitute the AQI of the same analyst-firm from a different short-term forecasting window. In the second, we randomly substitute the AQI from another city during the same short-term forecasting window. The advantage of doing this randomization is that we keep the effects of time variant characteristics in our data structure. If our findings are mainly driven by air pollution, our observed results should be much stronger than the effects estimated from these two placebo tests.

Panel A Table 7 reports the results where we randomize the PM pollution measure at the analyst-firm level. For example, take an analyst who follows a firm for 3 years and issues 12 forecasts in response to the 12 earnings announcements. Given a specific forecast issued after a given quarter, we randomly choose an AQI measure from other 11 quarters. After this randomization process, we re-

²⁴ When we compare the coefficients of *POLLUTED* POST* and *CLEAN* POST* across forecast timeliness, accuracy improvement, boldness, and negative boldness, the difference is significant with at least $p < 0.1$.

Table 6
Difference-in-differences designs

Panel A Summary statistics for change in air pollution and forecast timeliness				
	Control	Clean	Polluted	
<i>AQI</i>				
#observations	2,762	910	1,852	
Post = 0	66.2730	94.7146	56.9514	
post = 1	66.8125	41.0637	131.0400	
<i>Diff</i>	-0.5395	53.6528***	-74.0886***	
<i>TL</i>				
#observations	2,762	910	1,852	
Post = 0	0.5807	0.5473	0.6026	
post = 1	0.5873	0.6681	0.5680	
<i>Diff</i>	-0.0065	-0.1209***	0.0346*	
Panel B Summary statistics for change in air pollution and forecast accuracy, boldness and negative boldness				
	Control	Clean	Polluted	
<i>AQI</i>				
#observations	1,316	421	881	
Post = 0	65.0501	92.7391	55.1044	
post = 1	65.2207	39.1814	126.5870	
<i>Diff</i>	-0.1706	53.5577***	-71.4826***	
<i>ACCURACY</i>				
#observations	1,315	421	879	
Post = 0	0.5502	0.4076	0.4967	
post = 1	0.5053	0.5274	0.5116	
<i>Diff</i>	0.0448*	-0.1198***	-0.0150	
<i>BOLD</i>				
#observations	1,315	421	878	
Post = 0	0.4970	0.4457	0.5737	
post = 1	0.5449	0.5063	0.5395	
<i>Diff</i>	-0.0479**	-0.0607	0.0341	
<i>NEG_BOLD</i>				
#observations	1,315	421	878	
Post = 0	0.2979	0.2011	0.3638	
post = 1	0.3379	0.3080	0.3488	
<i>Diff</i>	-0.0400*	-0.1069***	0.0150	
Panel C Multivariate analyses				
Variables	(1)	(2)	(3)	(4)
	<i>TL</i>	<i>ACCURACY</i>	<i>BOLD</i>	<i>NEG_BOLD</i>
<i>POST</i>	0.0281 (1.43)	-0.0230 (-0.63)	0.0779** (2.07)	0.0482 (1.60)
<i>POLLUTED</i> × <i>POST</i>	-0.0463* (-1.68)	0.0520 (1.01)	-0.1229** (-2.47)	-0.0885** (-2.18)
<i>CLEAN</i> × <i>POST</i>	0.1090*** (3.05)	0.1573** (2.37)	0.0358 (0.64)	0.0747 (1.61)
Controls	YES	YES	YES	YES
Observations	5,524	2,615	2,613	2,613
Year	2010-2016	2010-2016	2010-2016	2010-2016
R-squared	0.193	0.248	0.269	0.367
Firm FE	YES	YES	YES	YES

This table presents the results of conducting difference-in-differences designs. We first identify treatment sample consisting of analyst-firm-quarter forecasts experiencing drastic changes (a difference of 40) in *AQI* between two adjacent quarters. Each pair of treatment observations represent the same analyst-firm forecasts with different *AQIs*. *POLLUTED* (*CLEAN*) equals 1 if there is an increase (a decrease) in *AQI* from the preceding to current quarter. For each treatment observation, we match one control observation where an analyst from a different city follows the same firm over the same periods. There should not be drastic changes in *AQI* of control observations between the two quarters. *POST* equals 1 (0) for the current (preceding) quarter. All standard errors are corrected for firm level correlation. ***, **, * represent statistical significance at the 1%, 5%, 10% level (two-tailed), respectively. See the appendix for variable definitions.

Table 7
Placebo tests

Coefficient of <i>LOG_AQI</i>	(1)	(2)	(3)	(4)
	<i>TL</i>	<i>ACCURACY</i>	<i>BOLD</i>	<i>NEG_BOLD</i>
Panel A Randomly assign <i>AQI</i> to an analyst-firm-announcement				
Estimated coefficients Table 4	-0.0868	-0.0216	-0.0169	-0.0226
Mean	-0.0015	-0.0039	-0.0069	-0.0063
Standard deviation	0.0043	0.0038	0.0036	0.0019
Minimum	-0.0114	-0.0153	-0.0158	-0.0150
P1	-0.0114	-0.0127	-0.0156	-0.0141
P5	-0.0080	-0.0096	-0.0138	-0.0120
P10	-0.0071	-0.0089	-0.0123	-0.0108
Max	0.0134	0.0075	0.0036	0.0019
Percentile of coefficient	0.00	0.00	0.00	0.00
Sample times	150	150	150	150
Avg Observations	42,806	60,046	59,299	59,299
Year	2010-2016	2010-2016	2010-2016	2010-2016
Avg R-squared	0.140	0.126	0.133	0.203
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Quarter FE	YES	YES	YES	YES
FPI FE		YES	YES	YES
Panel B: Randomly assign city to an individual analyst				
Estimated coefficients Table 4	-0.0868	-0.0216	-0.0169	-0.0226
Mean	-0.0357	-0.0043	-0.0021	-0.0030
Standard deviation	0.0075	0.0079	0.0085	0.0079
Minimum	-0.05460	-0.0260	-0.0209	-0.0224
P1	-0.05310	-0.0261	-0.0205	-0.0185
P5	-0.04830	-0.0180	-0.0170	-0.0158
P10	-0.0444	-0.0143	-0.0129	-0.0140
Max	-0.0174	0.0210	0.0240	0.0202
Percentile of coefficient	0.00	< 0.05	< 0.06	0.00
Sample times	150	150	150	150
Avg Observations	39,270	55,333	54,648	54,648
Year	2010-2016	2010-2016	2010-2016	2010-2016
Avg R-squared	0.144	0.129	0.136	0.207
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Quarter FE	YES	YES	YES	YES
FPI FE		YES	YES	YES

This table presents summary statistics for the coefficient of *AQI* and the fit statistics of models in two placebo tests. The aim of the two placebo tests is to ensure that the observed effects of air pollution are not driven by some time variant omitted variables. Panel A shows the results where we randomize the PM pollution measure within an analyst-firm level. Panel B shows results where we randomize the location information for one analyst. See the appendix for variable definitions.

estimate Equations (1) and (2), repeating the test 150 times to generate the distribution of coefficients of *LOG_AQI*.²⁵ Panel A column (1) reports results of estimating the test for model (1), where *TL* is the dependent variable. The mean coefficient of *LOG_AQI* is -0.0015, which is only 1.7% of the actual coefficient reported in [Table 3](#) column (3). Results of the placebo tests for estimating model (2) are similar. The actual coefficients reported in [Table 4](#) panel B for *ACCURACY*, [Table 5](#) for *BOLD*, and [Table 6](#) panel B for *NEG_BOLD* are below the minimum values of the distributions of coefficients from the placebo tests.

Panel B [Table 7](#) reports results where we randomize analyst location information. For example, we randomly assign an analyst located in Beijing to another city in China and recalculate the *AQI* measure from that location. We then estimate our models and repeat this process 150 times. Although we expect that this randomized city *AQI* will not elicit the same results as our primary models, due to the large scale of air pollution in China on a given day, there is a likely a strong positive correlation in the PM pollution level between any two cities. This strong positive correlation implies that we may still obtain associations between the randomized *AQI* measure and our dependent variables. Column (1) indicates that the average coefficients linking randomized *AQI* to *TL* is -0.0357. Although not equal to zero, this value is only 41% of our original coefficient, which is in the 0th percentile of the distribution of false coefficients. In the remaining models, the actual coefficients of *ACCURACY*, *BOLD*, and *NEG_BOLD* reported in [Tables 5 and 6](#) are in the bottom 5th or better percentile of the distribution of placebo test results. Overall, the results from the both placebo tests provide confidence that our findings are not the result of some unobservable analyst-host city-time invariant characteristics.

²⁵ Another advantage of doing such placebo test is that there is no restriction on the (normal) distribution assumption of estimated coefficients.

5.5. Robustness tests

We conduct a series of robustness tests of our main models. To correct for biased standard errors introduced by city-level *AQI* correlations, we cluster standard errors by city. The t-values from this analysis are only marginally different from those reported in Tables 4 and 5. Our results are also robust to increasing the length of the short window after earnings announcements to [0,2] and [0,3]. Given that the seasonality may play a role in *AQI* variation, we expand our models by adding analyst-quarter paired fixed effects. Results of this analysis are qualitatively similar to those we originally reported. To investigate whether analyst revisions might be affected by the quality of pre-revision forecasts prior to earnings announcements, we compute accuracy, boldness and negative boldness measures for pre-revision forecasts in the same way that we develop the measures for models (1) and (2). Results fail to indicate a systematic association between pre-revision forecast attributes and *AQI*. Finally, rather than linear probability models, we estimate models (1) and (2) using probit models with fixed effects. Results continue to be supportive of our hypotheses.

5.6. What factors attenuate the negative effects of air pollution?

While PM pollution affects the timing and quality of analyst forecasts on average, there may be factors that help mitigate pollution's negative effects. For example, we conjecture that analysts with heavy workload located in high pollution cities are reluctant to work. If this conjecture holds, we should observe that there are stronger associations between our dependent variables and air pollution for analysts with higher expected workloads than for those with lower expected workload. We measure analyst workload of an analyst by counting the number of other earnings announcements made by firms that the analyst followed in the prior quarter within three days of the focal earnings announcement. Table 8, panel A provides results of this analysis. Results provided in panel A of Table 8, columns (1) and (2) indicate that the impact of pollution on forecast timeliness is lower for analysts with low workload versus those with a high workload ($p < 0.01$). Results reported in columns (3) through (8) indicate that the association between PM pollution and improvement in forecast accuracy, and negatively bold forecasts are primarily driven by analysts with higher workloads (at least $p < 0.05$).

Analysts facing greater labor market competition face additional pressure to perform well (Stickel, 1989; Zhang, 2008). If labor market competition motivates analysts to work hard regardless of environmental conditions, we should observe that competition mitigates the negative effects of air pollution. To examine this, we partition the sample based on competition among analysts, measured as the number of analysts following a firm. Panel B Table 8 provides results across this partition. Although competition does not appear to attenuate the negative association between air pollution and analysts' likelihood to issue timely forecasts (columns (1) and (2)), PM pollution has a greater effect on accuracy improvements and negative forecast boldness for analysts facing lower levels of competition (at least $p < 0.1$).

Finally, prior literature postulates that annual earnings announcements trigger higher trading volume relative to interim announcements (Griffin, 2003). High trading volume may reflect high information demand from investors, which could mitigate the negative effects of PM pollution on analyst information production. Table 8, panel C reports results of separating earnings announcements into annual or quarterly (*INTERIM*) announcements. While results reported in columns (1) and (2) indicate that PM pollution impacts the timeliness of forecasts following both annual and quarterly earnings announcements, the impact on quarterly earnings announcements is significantly greater ($p < 0.01$). The effect of PM pollution on forecast boldness (both *BOLD* and *NEG_BOLD*) are primarily related to quarterly earnings announcements (at least $p < 0.1$).

5.7. Market perception of pollution effects on analyst information production

Given the negative effects of air pollution on analyst information production, we investigate whether investors take the reduction in timeliness and lower quality of analyst forecasts into consideration in market valuations. We examine whether air pollution preceding issuance of analyst forecasts attenuates the association between share price reactions and forecast revisions. Following Gleason and Lee (2003), Bonner et al. (2007) and Harford et al. (2018), we estimate the following empirical model:²⁶

$$CAR = \gamma_0 + \gamma_1 REV + \gamma_2 LOG_AQI_PRE \times REV + X\beta + \varepsilon \quad (4)$$

where *CAR*, the cumulative abnormal returns, equals the cumulative daily raw returns over the two days centering on the forecast revision dates ([0, +1]) minus the cumulative daily value weighted market index returns over the same window with day 0 as the forecast revision date. If there are more than one forecast revisions on the same day, we divide *CAR* by the number of forecast revisions. By doing so, we essentially assume that investors do not price forecast revisions conditional upon air quality leading to those forecasts. As an alternative, to avoid the overlapping forecast revisions, we follow Gleason and Lee (2003) and re-estimate the model, keeping only standalone forecasts. *REV* is the difference between current analyst forecast for year *t* minus the prior forecast issued by the same analyst for year *t*, deflated by closing share price on day -1. *LOG_AQI_PRE* is the natural log of mean *AQI* preceding

²⁶ Our model differs from the market reaction model in DeHaan et al. (2017). DeHaan et al. (2017) examine how weather affects investors' pricing of firm earnings announcements. In their test, they implicitly assume that investors cluster in New York and are affected by the weather in New York when firms announce earnings. In our test, we do not have this assumption. Instead, we build on our previous findings that air pollution, in analysts' locations, negatively affects analyst information production and employ model (4) to investigate whether investors perceive the inferior forecasts issued by analysts exposed to air pollution.

the forecast revision $[-1,0]$. X comprises a series of both firm- and analyst-level variables: 1) *SIZE* measured as the natural log of market capitalization on day -2; 2) *MB* equal to market value of equity on day -2 divided by book value of equity from the most recent quarterly report before an analyst forecast; 3) *BETA* equal to the coefficient of market index returns from a regression of individual stock returns on market index returns over $[-120,10]$; 4) *LOG_N_FOLLOW* measured as the natural log of one plus the number of analysts following a firm-year; 5) *PAST_RETURN* measured as the cumulative abnormal monthly returns over the six month before forecast revision month; 6) *STAR* equal to 1 if an analyst became a star in the previous year; 7) *LOG_EXP* equal to the natural log of one plus the number of years since an analyst's first analyst report recorded by the database; 8) *LOG_FIRM_EXP* equal to the natural log of one plus the number of years since an analyst's first analyst report for the firm covered recorded by the database; 9) *LOG_N_IND* equal to the natural log of one plus the number of industries followed by an analyst in the prior one year; 10) *LOG_DAYS* measured as the natural log of one plus the number of days between an analyst forecast date and the corresponding earnings announcement date; 11) *LOG_DIST* measured as the natural log of one plus the number of days between the current and previous analyst forecast dates. In

Table 8

What factors attenuate the effects of air pollution?

Panel A The effects of information supply-high vs. low workload								
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>TL</i>	<i>TL</i>	<i>ACCURACY</i>	<i>ACCURACY</i>	<i>BOLD</i>	<i>BOLD</i>	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>
	<i>LOW_LOAD</i>	<i>HIGH_LOAD</i>	<i>LOW_LOAD</i>	<i>HIGH_LOAD</i>	<i>LOW_LOAD</i>	<i>HIGH_LOAD</i>	<i>LOW_LOAD</i>	<i>HIGH_LOAD</i>
<i>LOG_AQI</i> (α_1)	-0.0627*** (-7.05)	-0.1034*** (-12.34)	-0.0122 (-1.41)	-0.0436*** (-3.24)	-0.0108 (-1.11)	-0.0289*** (-2.68)	-0.0119 (-1.27)	-0.0406*** (-4.07)
CONTROLS	YES	YES	YES	YES	YES	YES	YES	YES
α_1 between groups	(1)-(2) = -0.0407***		(3)-(4) = -0.0314**		(5)-(6) = -0.0181		(7)-(8) = -0.0287**	
Observations	22,109	20,172	31,566	27,741	31,566	27,741	31,566	27,741
Year	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016
R-squared	0.176	0.182	0.156	0.175	0.157	0.183	0.236	0.256
Firm FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Quarter FE	YES	YES	YES	YES	YES	YES	YES	YES
FPI FE			YES	YES	YES	YES	YES	YES

Panel B The effects of information supply-high vs. low labor market competition								
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>TL</i>	<i>TL</i>	<i>ACCURACY</i>	<i>ACCURACY</i>	<i>BOLD</i>	<i>BOLD</i>	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>
	High	Low	High	Low	High	Low	High	Low
<i>LOG_AQI</i>	-0.0894*** (-9.66)	-0.0833*** (-8.88)	-0.0055 (-0.60)	-0.0373*** (-3.65)	-0.0101 (-1.04)	-0.0193* (-1.86)	-0.0129 (-1.38)	-0.0301*** (-3.09)
CONTROLS	YES	YES	YES	YES	YES	YES	YES	YES
α_1 between groups	(1)-(2) = -0.0061		(3)-(4) = 0.0318**		(5)-(6) = 0.0092		(7)-(8) = -0.0172*	
Observations	21,741	20,540	31,327	27,980	31,327	27,980	31,327	27,980
Year	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016
R-squared	0.134	0.192	0.115	0.185	0.124	0.186	0.218	0.246
Firm FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Quarter FE	YES	YES	YES	YES	YES	YES	YES	YES
FPI FE			YES	YES	YES	YES	YES	YES

Panel C The effects of information demand-annual vs. interim earnings announcements								
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>TL</i>	<i>TL</i>	<i>ACCURACY</i>	<i>ACCURACY</i>	<i>BOLD</i>	<i>BOLD</i>	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>
	ANNUAL	INTERIM	ANNUAL	INTERIM	ANNUAL	INTERIM	ANNUAL	INTERIM
<i>LOG_AQI</i>	-0.0367** (-2.28)	-0.0985*** (-13.76)	-0.0253 (-1.56)	-0.0166** (-2.21)	0.0136 (0.88)	-0.0238*** (-3.08)	0.0024 (0.15)	-0.0254*** (-3.45)
CONTROLS	YES	YES	YES	YES	YES	YES	YES	YES
α_1 between groups	(1)-(2) = 0.0618***		(3)-(4) = -0.0087		(5)-(6) = 0.0374**		(7)-(8) = 0.0278*	
Observations	7,891	34,390	10,920	48,387	10,920	48,387	10,920	48,387
Year	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016

(continued on next page)

Table 8 (continued)

Panel C The effects of information demand-annual vs. interim earnings announcements								
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<i>TL</i>	<i>TL</i>	<i>ACCURACY</i>	<i>ACCURACY</i>	<i>BOLD</i>	<i>BOLD</i>	<i>NEG_BOLD</i>	<i>NEG_BOLD</i>
	ANNUAL	INTERIM	ANNUAL	INTERIM	ANNUAL	INTERIM	ANNUAL	INTERIM
R-squared	0.219	0.148	0.204	0.135	0.198	0.140	0.347	0.218
Firm FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Quarter FE	YES	YES	YES	YES	YES	YES	YES	YES
FPI FE			YES	YES	YES	YES	YES	YES

This table presents OLS regression results where we partition our sample based upon several analyst and firm level variables which may alter the negative association between *AQI* and dependent variables. The partitioning variable in Panel A, Panel B, and Panel C is the workload of an analyst, the competition among analysts following a firm in the previous year, and investor demand for analyst report. The workload is constructed as follows: given one particular analyst-firm-quarter announcement, we count the number of concurrent earnings announcements made by other firms followed by the analyst in the prior quarter. For more than half of analyst-firm-quarter announcements, there is no or only one concurrent earnings announcement. Accordingly, we define *LOW_LOAD* as those observations with no or one concurrent earnings announcement, with remaining observations as *HIGH_LOAD*. The competition among analysts is measured as the number of analysts following the firm-quarter in the previous year. The investor demand for analyst reports is measured as a dummy variable indicating whether an earnings announcement is annual or interim. In all regression models, we log transform *AQI*, and all continuous analyst level control variables, and add firm, year and quarter fixed effects. We further add the forecast period fixed effects (*FPI*) when taking *ACCURACY*, *BOLD* or *NEG_BOLD* as the dependent variable. All standard errors are clustered by firm. The sample period spans 2010-2016. ***, **, * represent statistical significance at the 1%, 5%, 10% level (two-tailed), respectively. See the appendix for variable definitions.

addition, we include firm and year fixed effects in model (4). We cluster the standard errors by forecast revision date to correct for the cross-correlation in market returns on the same date.

Panel A Table 9 summarizes the sampling process. We first extract the latest forecasts of each analyst issued for year t during 2009-2016 (87,178 observations). We delete those without preceding forecasts for year t (33,624 observations). To avoid the confounding effects of earnings announcements on stock price reactions, we drop those forecasts with quarterly earnings announcements falling in the window of $[-2, +2]$ surrounding forecast dates (19,771 observations). We drop 12,805 and 4,615 observations without analyst location or air quality information. Missing variables further reduces the sample by 2,912. The ultimate sample includes 13,451 analyst-firm-year observations.

Panel B Table 9 tabulates the descriptive statistics of variables in model (4). The mean *REV* is -0.0008, consistent with optimism in analyst forecasts released early in the year (Gleason and Lee, 2003). The mean value of *CAR* is positive, suggesting that there might be a forecast revision premium.

Panel C Table 9 reports results of estimating model (4). Columns (1)-(3) present results estimated using the full sample and columns (4)-(6) present results estimated using standalone forecast revisions. In columns (3) and (6), we expand the model by fully interacting *REV* with each of the control variables. Across all models, the coefficients of $REV \times LOG_AQI_PRE$ are significantly negative at least at the 5% level, implying that investors consider the impairment of information provided by analysts due to air pollution.²⁷

6. Conclusion

Many countries experience increasing episodes of severe air pollution, especially PM pollution. Prior studies provide evidence that air pollution reduces human capital formation and decreases workers' marginal productivity. Our study extends the literature by investigating the extent to which PM pollution affects financial analysts' forecasting behavior.

We find that PM pollution is negatively associated with analysts' propensity to issue forecasts within a short window following earnings announcements. Even when they issue forecasts within the short window, analysts exposed to air pollution are unlikely to improve the accuracy of these forecasts. Further tests indicate that PM pollution reduces the ability of analysts to synthesize information from earnings announcements. Using forecast boldness to capture analyst information production activities, we find that analysts forecasting during polluted periods are less likely to make bold, especially, negatively bold forecasts. Several factors appear to affect the impact of pollution on forecasts. Analysts with high workloads are particularly impacted by pollution. Labor market competition is also important—analysts who face higher levels of labor market competition appear to compensate for the negative effects of pollution, whereas analysts who do not face competition do not compensate. High information demand during annual earnings announcement periods also mitigates the negative effects of air pollution to a certain extent. Finally, we find that market

²⁷ After controlling for the two-way interaction between *REV* and *BOLD*, there is a slight decrease in the coefficient of the two-way interaction between *REV* and *LOG_AQI_PRE* (to 0.0797 with a t-statistic of 1.87) in column (3). Our results show that after controlling for analyst revision attributes, air pollution still attenuates the market response to analyst revisions. Further (untabled) analysis suggests that forecasts released by analysts located in high-pollution periods only trigger about 2/3 of the market response to forecasts released by other analysts.

Table 9

Market perception of the negative effect of air pollution on analyst information production

Panel A Sampling process							
	# Obs						
The latest analyst forecasts issued for the current year from 2009 to 2016	87,178						
Less: those without preceding forecasts for a firm-year issued by the same analyst	(33,624)						
Less: forecasts with quarterly earnings announcements falling in the window of [-2, +2] surrounding a forecast revision date	(19,771)						
Less: forecasts without analyst location information	(12,805)						
Less: forecasts without air pollution information	(4,615)						
Less: observation with missing variables	(2,912)						
Overall sample	13,451						
Panel B Descriptive statistics							
Variables	Mean	Sd	P25	P50	P75	N	
CAR	0.0081	0.0347	-0.0105	0.0040	0.0223	13,451	
REV	-0.0008	0.0095	-0.0003	0.0000	0.0000	13,451	
LOG_AQI_PRE	4.3480	0.4670	4.0160	4.2840	4.6200	13,451	
SIZE	23.2100	1.0020	22.5000	23.0900	23.8200	13,451	
MB	3.6210	3.7870	1.2510	2.5190	4.5420	13,451	
PAST_RET	0.0997	0.2790	-0.0834	0.0596	0.2350	13,451	
STAR	0.2850	0.4520	0.0000	0.0000	1.0000	13,451	
BETA	1.1190	0.3190	0.9040	1.1330	1.3370	13,451	
LOG_N_FOLLOW	1.9700	0.6440	1.6090	1.9460	2.4850	13,451	
LOG_EXP	1.5940	0.5540	1.3860	1.6090	1.9460	13,451	
LOG_FIRM_EXP	0.7860	0.6260	0.0000	0.6930	1.0990	13,451	
LOG_N_IND	1.6590	0.4800	1.3860	1.6090	1.9460	13,451	
LOG_DAYS	4.3430	0.7100	4.0070	4.5220	4.8980	13,451	
LOG_DIST	3.5910	0.9940	2.9960	3.7610	4.3570	13,451	
Panel C Multivariate regression							
Variables	(1)	(2)	(3)	(4)	(5)	(6)	
	CAR	CAR	CAR	CAR	CAR	CAR	
REV	0.6362** (3.19)	0.5777*** (2.73)	0.7216 (0.69)	0.7753** (2.52)	0.7386** (2.16)	0.6441 (0.54)	
LOG_AQI_PRE	0.0009 (0.87)	0.0009 (0.97)	0.0009 (0.92)	0.0006 (0.40)	0.0005 (0.36)	0.0004 (0.28)	
REV × LOG_AQI_PRE	-0.1125** (-2.49)	-0.0965** (-2.02)	-0.1247** (-2.42)	-0.1513** (-2.16)	-0.1401* (-1.84)	-0.1772** (-2.17)	
SIZE		-0.0115*** (-5.96)	-0.0116*** (-6.03)		-0.0124*** (-5.03)	-0.0125*** (-5.06)	
MB		-0.0000 (-0.06)	-0.0000 (-0.02)		-0.0002 (-0.50)	-0.0001 (-0.46)	
BETA		0.0004 (0.20)	0.0004 (0.24)		-0.0005 (-0.28)	-0.0005 (-0.27)	
LOG_N_FOLLOW		-0.0000 (-0.01)	-0.0001 (-0.10)		0.0005 (0.28)	0.0003 (0.19)	
PAST_RET		-0.0010 (-0.40)	-0.0010 (-0.38)		-0.0017 (-0.55)	-0.0017 (-0.52)	
STAR		0.0004 (0.55)	0.0005 (0.61)		0.0010 (0.93)	0.0010 (0.96)	
LOG_EXP		0.0011 (1.46)	0.0011 (1.42)		0.0018 (1.58)	0.0019 (1.56)	
LOG_FIRM_EXP		0.0001 (0.28)	0.0002 (0.31)		-0.0002 (-0.30)	-0.0002 (-0.25)	
LOG_N_IND		-0.0014** (-2.66)	-0.0013** (-2.45)		-0.0016* (-1.94)	-0.0015* (-1.75)	
LOG_DAYS		-0.0009** (-3.16)	-0.0010*** (-3.01)		-0.0019*** (-3.71)	-0.0019*** (-3.62)	
LOG_DIST		-0.0004** (-2.01)	-0.0004* (-1.80)		-0.0005 (-1.21)	-0.0005 (-1.13)	
REV Interacted with controls?		NO	YES		NO	YES	
Observations	13,451	13,451	13,451	9,066	9,054	9,054	
R-squared	0.199	0.209	0.209	0.233	0.243	0.244	

(continued on next page)

Table 9 (continued)

Panel C Multivariate regression						
Variables	(1)	(2)	(3)	(4)	(5)	(6)
	CAR	CAR	CAR	CAR	CAR	CAR
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

This table presents the results of estimating model (4). Panel A filters the sample used to estimate model (4). Panel B tabulates summary statistics of all variables in model (4) and Panel C presents the estimation results. We compute *CAR* as the market adjusted cumulative market returns within [0,1] of analyst forecast date (day 0 is the forecast date) to measure how investors value analyst revision. *REV* is the difference between current and the most recent preceding forecast, deflated by the closing market price on day -1. *LOG_AQI_PRE* is the natural log of mean *AQI* over [-1,0]. In the regression model, we log transform all continuous analyst level control variables, and add firm and year fixed effects. All standard errors are clustered by forecast date to correct for the cross-correlation in market returns on the same date. The sample period spans 2010-2016. ***, **, * represent statistical significance at the 1%, 5%, 10% level (two-tailed), respectively. See the appendix for variable definitions.

response to forecast revisions are consistent with investors understanding the detrimental effects of air pollution on analyst information production. Overall, these findings are consistent with the notion that air pollution represents a hidden cost to the capital market.

Our study suggests several directions for future research. First, a more direct test could be conducted to examine how air pollution affects the work pattern of analysts. For instance, how do they allocate time to work and leisure during pollution days? Second, financial analysts are only one of the agents producing information for capital markets. Further studies can investigate how air pollution changes the behaviors of other market participants such as auditors and the media.

Acknowledgements

The authors thank valuable comments from Lili Dai, Ian Gow, Jesper Haga, Andrew Jackson, Like Jiang, Flora Kuang, Richard Morris, Matthew Pinnuck, Stefan Schantl, Baljit Sidhu, Rencheng Wang and seminar participants at the EAA Annual Congress 2018 Milan, UNSW Sydney, The University of Melbourne, Xi'an International Studies University, Xi'an Jiaotong University, Zhongnan University of Economics and Law, Xiamen University and Jinan University. Luo acknowledges financial support from the National Natural Science Foundation of China (Grant No. 71572160). An earlier version of the paper was entitled 'Particulate Matter Pollution and Analyst Information Production'.

Appendix A. Variable definitions

This table tabulates the definitions of all variables used in the study.

Variable name	Definition
<i>TL</i>	= An indicator variable that equals 1 if an analyst revises his forecasts within two days (i.e. day 0 and 1) of a quarterly earnings announcement of a focal firm, otherwise 0.
<i>ACCURACY</i>	= An indicator variable that equals 1 if one analyst's revised earnings forecast is more accurate than his prior earnings forecast, otherwise 0.
<i>BOLD</i>	= An indicator variable that equals 1 if an analyst revises his forecasts below or above both his own prior forecast and peer consensus forecasts, otherwise 0.
<i>NEG_BOLD</i>	= An indicator variable that equals 1 if the analyst's revised earnings forecasts are below his own prior forecast and peer consensus forecasts, otherwise 0.
<i>POS_BOLD</i>	= An indicator variable that equals 1 if the analyst's revised earnings forecasts are above his own prior forecast and peer consensus forecasts, otherwise 0.
<i>AQI</i>	= The average daily air quality index between the earning announcement date and the analyst forecast revision date if an earnings forecast is issued within two days of the earnings announcement date. If the earnings forecast is not issued within the above short window, we measure <i>AQI</i> over the two-day window. In our study, <i>AQI</i> is measured for each analyst according to the air pollution level in his own location (city). High levels of <i>AQI</i> correspond to high levels of air pollution.
<i>LOG_AQI</i>	= The natural logarithm of <i>AQI</i> .
<i>N_FIRM</i>	= The number of firms one analyst follows.
<i>N_IND</i>	= The number of unique industries one analyst follows.
<i>SIZE_BROKER</i>	= The number of analysts employed by a broker.
<i>EXP</i>	= The number of quarters for an analyst between his first forecast report date recorded in the database and the current earnings announcement date.
<i>N_FOLLOW</i>	= The number of analysts who issued at least one report for the firm.
<i>STAR</i>	= An indicator variable that equals 1 if one analyst was awarded the title of star analyst by the <i>NEW FORTUNE</i> magazine in the preceding calendar year, otherwise 0.
<i>DAYS</i>	= The number of days between the analyst forecast revision date and the forecast period end date.
<i>HOLDPERCT</i>	= The percentage of shares held by funds.
<i>ROA</i>	= Net profit divided by total assets.
<i>LEV</i>	= The total liabilities deflated by total assets.

<i>SALES_GROWTH</i>	=	Quarter-to-quarter sales revenue difference deflated by the sales revenue for the same quarter last year
<i>TRVOL</i>	=	The average trading volume over the [-1, +1] window of the three earnings announcement dates, scaled by the average trading volume in the non-earnings announcement days in the previous year.
<i>RETVAR</i>	=	The standard deviation of daily stock returns over the [-1, +1] window of the three earnings announcement dates minus the standard deviation of daily stock returns over other non-earnings announcement days in the previous year.
<i>MB</i>	=	Total market capitalization plus the book value of total debt deflated by the book value of total assets
<i>SUE</i>	=	The quarter-to-quarter net profit difference deflated by the market capitalization on the day before the earnings announcement date.
<i>SIZE</i>	=	The logarithm of total market capitalization on the day before the earnings announcement dates.
<i>ANNUAL</i>	=	An indicator variable that equals 1 if an earnings announcement is made for annual earnings, otherwise 0.
<i>INTERIM</i>	=	An indicator variable that equals 1 if an earnings announcement is made for interim earnings, otherwise 0.
<i>LOSS</i>	=	An indicator variable that equals 1 if the ROA is less than 0, otherwise 0.
<i>WIND</i>	=	The average wind speed in a calendar day.
<i>CLOUD_RAIN</i>	=	An indicator variable that equals 1 if the weather in a calendar day is cloudy/rainy, otherwise 0.
<i>VOLUME_O_REV</i>	=	Abnormal trading volume within [0,1] of analyst forecast date (day 0 is the forecast date), where abnormal trading volume is daily trading volume adjusted by the average trading volume over the 60 trading days before the analyst forecast date.
<i>TO_END_DAYS</i>	=	The number of days between an analyst forecast revision date and the forecasts period end.

References

- Angrist, J.D., Pischke, J.S., 2008. Mostly harmless econometrics: an Empiricist's companion. Princeton university press.
- Archsmith, J., Heyes, A., Saberian, S., 2016. Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation. Working paper. University of California, California.
- Bakian, A.V., Huber, R.S., Coon, H., Gray, D., Wilson, P., McMahon, W.M., Renshaw, P.F., 2015. Acute air pollution exposure and risk of suicide completion. *American Journal of Epidemiology* 181 (5), 295–303.
- Ball, R., Brown, P., 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6 (2), 159–178.
- Bernard, V.L., Thomas, J.K., 1989. Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* 27 (1), 1–36.
- Beyer, A., Cohen, D.A., Lys, T.Z., Walthe, B.R., 2010. The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics* 50 (2–3), 296–343.
- Block, M.L., Elder, A., Auten, R.L., Bilbo, S.D., Chen, H., Chen, J., Cory-Slechta, D.A., Costa, D., Diaz-Sanchez, D., Dorman, D.C., 2012. The outdoor air pollution and brain health workshop. *Neurotoxicology* 33 (5), 972–984.
- Bonner, S.E., Hugon, A., Walther, B.R., 2007. Investor reaction to celebrity analysts: The case of earnings forecast revisions. *Journal of Accounting Research* 45 (3), 481–513.
- Calderon-Garciduenas, L., Calderon-Garciduenas, A., Torres-Jardon, R., Avila-Ramirez, J., Kulesza, R.J., Angiulli, A.D., 2015. Air pollution and your brain: What do you need to know right now. *Prim. Health Care Res. Dev.* 16 (4), 329–345.
- Chang, T., Zivin, J.S.G., Gross, T., Neidell, M.J., 2014. Particulate pollution and the productivity of pear packers. Working paper. National Bureau of Economic Research, Cambridge.
- Chang, T., Zivin, J.S.G., Gross, T., Neidell, M.J., 2016. The effect of pollution on worker productivity: Evidence from call-center workers in China. Working paper. National Bureau of Economic Research, Cambridge.
- Chen, C., Chen, Y., Podolski, E.J., Veeraraghavan, M., 2017. Managerial mood and earnings forecast bias: Evidence from sunshine exposure. Working paper. Monash University, Melbourne.
- Chiquiar, D., Hanson, G., 2005. International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *Journal of Political Economy* 113 (2), 239–281.
- Chiswick, B., 1999. Are immigrants favourably self-selected? *Am. Econ. Assoc. Papers Proc.* 89 (2), 181–185.
- Christmann, P., Taylor, G., 2001. Globalization and the environment: Determinants of firm self-regulation in China. *Journal of International Business Studies* 32 (3), 439–458.
- Clarke, J., Subramanian, A., 2006. Dynamic forecasting behavior by analysts: Theory and evidence. *Journal of Financial Economics* 80 (1), 81–113.
- Clement, M.B., Tse, S.Y., 2005. Financial analyst characteristics and herding behavior in forecasting. *The Journal of Finance* 60 (1), 307–341.
- Correia, S., 2015. Singletons, cluster-robust standard errors and fixed effects: a bad mix. Technical Note. Duke University.
- DeHaan, E., Madsen, J., Piotroski, J.D., 2017. Do weather-induced moods affect the processing of earnings news? *Journal of Accounting Research* 55 (3), 509–550.
- Dong, R., Fisman, R., Wang, Y., Xu, N., 2019. Air pollution, affect, and forecasting bias: evidence from Chinese financial analysts (July 15, 2019). *J. Fin. Econ.* Forthcom Available at SSRN. <https://ssrn.com/abstract=3419885>.
- EPA, 2006. National Ambient Air Quality Standards (NAAQS) for Particulate Matter (PM_{2.5}). Environmental Protection Agency, Washington DC cited 17 September 2019]. Available at: <https://www.epa.gov/pm-pollution/2006-national-ambient-air-quality-standards-naaqs-particulate-matter-pm25>.
- Gleason, C.A., Lee, C.M., 2003. Analyst forecast revisions and market price discovery. *The Accounting Review* 78 (1), 193–225.
- Greene, W., 2004. Fixed effects and bias due to the incidental parameters problem in the Tobit model. *Econometric reviews* 23 (2), 125–147.
- Griffin, P.A., 2003. Got information? Investor response to form 10-K and form 10-Q EDGAR filings. *Review of Accounting Studies* 8 (4), 433–460.
- Gul, F., Lundholm, R., 1995. Endogenous timing and the clustering of agents' decisions. *Journal of Political Economy* 103 (5), 1039–1066.
- Guttman, I., 2010. The timing of analysts' earnings forecasts. *The Accounting Review* 85 (2), 513–545.
- Hanlon, M., Hoopes, J.L., 2014. What do firms do when dividend tax rates change? an examination of alternative payout responses. *Journal of Financial Economics* 114 (1), 105–124.
- Harford, J., Jiang, F., Wang, R., Xie, F., 2018. Analyst career concerns, effort allocation, and firms' information environment. *The Review of Financial Studies* 32 (6), 2179–2224.
- Heyes, A., Neidell, M., Saberian, S., 2016. The effect of air pollution on investor behavior: Evidence from the S&P 500. Working paper. National Bureau of Economic Research, Cambridge.
- Hoberg, G., Phillips, G., Prabhala, N., 2014. Product market threats, payouts, and financial flexibility. *The Journal of Finance* 69 (1), 293–324.
- Hong, H., Kubik, J.D., 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance* 58 (1), 313–351.
- Hong, H., Kubik, J.D., Solomon, A., 2000. Security analysts' career concerns and herding of earnings forecasts. *The Rand Journal of Economics* 31 (1), 121–144.
- Huang, H.H., Kerstein, J., Wang, C., 2018. The impact of climate risk on firm performance and financing choices: An international comparison. *Journal of International Business Studies* 49 (5), 633–656.
- Ivkovic, Z., Jegadeesh, N., 2004. The timing and value of forecast and recommendation revisions. *Journal of Financial Economics* 73 (3), 433–463.
- Jacob, J., Lys, T.Z., Neale, M.A., 1999. Expertise in forecasting performance of security analysts. *Journal of Accounting and Economics* 28 (1), 51–82.
- Kadous, K., Krische, S.D., Sedor, L.M., 2006. Using counter-explanation to limit analysts' forecast optimism. *The Accounting Review* 81 (2), 377–397.
- Kadous, K., Mercer, M., Thayer, J., 2009. Is there safety in numbers? the effects of forecast accuracy and forecast boldness on financial analysts' credibility with investors. *Contemporary Accounting Research* 26 (3), 933–968.
- Ke, B., Yu, Y., 2006. The effect of issuing biased earnings forecasts on analysts' access to management and survival. *Journal of Accounting Research* 44 (5), 965–999.

- Keskek, S., Tse, S., Tucker, J.W., 2014. Analyst information production and the timing of annual earnings forecasts. *Review of Accounting Studies* 19 (4), 1504–1531.
- Lavy, V., Ebenstein, A., Roth, S., 2014. The impact of short term exposure to ambient air pollution on cognitive performance and human capital formation. Working paper. National Bureau of Economic Research, Cambridge.
- Li, J.J., Massa, M., Zhang, H., Zhang, J., 2019. Behavioral bias in haze: Evidence from air pollution and the disposition effect in China. Working paper. INESEAD, Paris.
- Lim, Y., Kim, H., Kim, J.H., Bae, S., Park, H.Y., Hong, Y., 2012. Air pollution and symptoms of depression in elderly adults. *Environmental Health Perspectives* 120 (7), 1023.
- Michaely, R., Womack, K.L., 1999. Conflict of interest and the credibility of underwriter analyst recommendations. *The Review of Financial Studies* 12 (12), 653–686.
- O'Brien, P.C., McNichols, M.F., Hsiou-Wei, L., 2005. Analyst impartiality and investment banking relationships. *Journal of Accounting Research* 43 (4), 623–650.
- Petersen, M.A., 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies* 22 (1), 435–480.
- Piotroski, J.D., Roulstone, D.T., 2004. The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices. *The Accounting Review* 79 (4), 1119–1151.
- Power, M.C., Kioumourtoglou, M., Hart, J.E., Okereke, O.I., Laden, F., Weiskopf, M.G., 2015. The relation between past exposure to fine particulate air pollution and prevalent anxiety: Observational cohort study. *BMJ* 350, h1111.
- Ramnath, S., Rock, S., Shane, P., 2008. The financial analyst forecasting literature: A taxonomy with suggestions for further research. *International Journal of Forecasting* 24 (1), 34–75.
- Sedor, L.M., 2002. An explanation for unintentional optimism in analysts' earnings forecasts. *The Accounting Review* 77 (4), 731–753.
- Stickel, S.E., 1989. The timing of and incentives for annual earnings forecasts near interim earnings announcements. *Journal of Accounting and Economics* 11 (2), 275–292.
- Suglia, S.F., Gryparis, A., Wright, R.O., Schwartz, J., Wright, R.J., 2007. Association of black carbon with cognition among children in a prospective birth cohort study. *American Journal of Epidemiology* 167 (3), 280–286.
- Szyzkowicz, M., 2007. Air pollution and emergency department visits for depression in Edmonton, Canada. *International Journal of Occupational Medicine and Environmental Health* 20 (3), 241–245.
- Truman, B., 1994. Analyst forecasts and herding behavior. *Review of Financial Studies* 7, 97–124.
- Vette, A.F., Rea, A.W., Lawless, P.A., Rodes, C.E., Evans, G., Highsmith, V.R., Sheldon, L., 2001. Characterization of indoor-outdoor aerosol concentration relationships during the Fresno PM exposure studies. *Aerosol Science & Technology* 34 (1), 118–126.
- Vidal, J., 2016. Air pollution rising at an 'alarming rate' in world's cities. *The Guardian*(May 12). <https://www.theguardian.com/environment/2016/may/12/air-pollution-rising-at-an-alarming-rate-in-worlds-cities>.
- Walsh, J., 2016. Air pollution: how does it affect you? *The Guardian*(May 12). <https://www.theguardian.com/environment/2016/may/12/air-pollution-how-does-it-affect-you>.
- Yezeqel, A., 2015. Why do analysts revise their stock recommendations after earnings announcements? *Journal of Accounting and Economics* 59 (2-3), 163–181.
- Zhang, Y., 2008. Analyst responsiveness and the post-earnings-announcement drift. *Journal of Accounting and Economics* 46 (1), 201–215.
- Zivin, J.G., Neidell, M., 2012. The impact of pollution on worker productivity. *American Economic Review* 102 (7), 3652–3673.