Air Pollution, Affect, and Forecasting Bias: Evidence from Chinese Financial Analysts[*]

Rui Dong, Raymond Fisman, Yongxiang Wang, and Nianhang Xu[†]

Abstract

We document a negative relation between air pollution during corporate site visits by investment analysts and subsequent earnings forecasts. After accounting for analyst, weather , and firm characteristics, an extreme worsening of air quality from "good/excellent" to "severely polluted" is associated with a more than 1 percentage point lower profit forecast, relative to realized profits. We explore heterogeneity in the pollution-forecast relation to understand better the underlying mechanism. Pollution only affects forecasts that are announced in the weeks immediately following a visit, indicating that mood likely plays a role, and the effect of pollution is less pronounced when analysts from different brokerages visit on the same date, suggesting a debiasing effect of multiple perspectives. Finally, there is suggestive evidence of adaptability to environmental circumstances – forecasts from analysts based in high pollution cities are relatively unaffected by site visit pollution.

JEL classifications: D91; G41; Q5

Keywords: Pollution; Forecasting bias; Investment analysts; Adaptation

[†]Dong: Department of Finance, School of Business, Renmin University of China, Beijing, PR. China, 100872 (email: dongrui@ruc.edu.cn); Fisman: Economics Department, Boston University, Room 304A, Boston, MA 02215 (email:rfisman@bu.edu); Wang: Finance and Business Economics Department, Marshall School of Business, University of Southern California, HOH 716, Los Angeles, CA 90089 (email: yongxiaw@marshall.usc.edu), and China Academy of Financial Research; and Xu: Department of Finance, School of Business, Renmin University of China, Beijing, PR. China, 100872 (email: nhxu@ruc.edu.cn)

1. Introduction

We study the relation between air pollution during corporate site visits by investment analysts in China and earnings forecasts issued in the days that follow. This setting allows us to examine the effect of plausibly extraneous ambient circumstances on judgment for individuals who should have both the expertise and incentive to screen out such influences. Investment analysts are well-educated, well-trained, and well-motivated to make accurate assessments of corporate earnings (Beyer et al., 2010). Analysts themselves recognize site visits as a crucial input into profit projections (Brown et al., 2015), so it is a task for which they should be particularly attentive to objective determinants of profitability. [1]

At the same time, there exists a decades-old literature on the impact of environmental conditions on mood and the resultant effect on decision-making (for seminal contributions see Schwarz and Clore, 1983 and Cunningham, 1979). Finance scholars have extended this line of research to study the effect of weather on stock market prices and trading behavior, as mediated by weather's effect on mood, with the weight of the evidence indicating that better weather leads to more optimism and higher prices (see Saunders, 1993 and Hirshleifer and Shumway, 2003 for the original "sunshine effect" on stock prices, Kamstra et al., 2003 for the link between daylight and stock prices, and Goetzmann et al., 2015 for the effect of cloud cover on institutional investors' pessimism).

A more recent – and more closely related – body of work links pollution both to mood, and also trading behavior and stock prices, with mood posited as the mediating channel (see Vert et al., 2017 on the association between pollution and mood, Levy and Yagil, 2011 and Lepori, 2016 for the association between pollution and stock prices, and Huang et al., forthcoming and Li et al., forthcoming for the association between pollution and investor biases), further reinforcing the possibility that pollution during site visits may impact analyst forecasts.

China is a natural setting in which to study this link. First, since 2009, the Shenzhen Stock Exchange has required that all site visits be disclosed, so we may observe the timing of analysts' visits (in the U.S., for example, such disclosures are not required). These data allow us to identify 3,824 earnings forecasts made by 726 investment analysts in the weeks following

---

[1]For the impact of corporate site visits in the China setting, see Cheng et al. (2016) and Han et al. (2018) for the effect on forecast accuracy, and Cheng et al. (2018) for the effect on stock prices.

corporate site visits during 2009-2015. Second, pollution is very severe on average in China and highly variable both across geographies and across time, which provides variation in ambient circumstances that is of such magnitude as to plausibly have a causal impact on analyst affect. More specifically, the visits in our data set take place in 105 cities, spread across the country,[2] which, when combined with the random variation in pollution caused by differing meteorological conditions across analysts' visit dates, provides plausibly exogenous variation in pollution during site visits that we may exploit to explore the relation with subsequent forecasts. (The short-term randomness of local conditions also presents a ready placebo test, which we return to below.)

A natural conjecture, given the weather-mood relation documented in earlier work, is that higher air pollution will be associated with lower earnings forecasts. Consistent with higher pollution leading to increased pessimism, we find that a city's air quality index (AQI) on the date of a site visit is negatively correlated with the visiting analyst's subsequent earnings forecast, relative to realized earnings. Intriguingly, since analysts' forecasts are positively biased overall, pollution-induced pessimism brings forecasts closer to unbiasedness.[3]

We present several robustness checks and placebo tests which bolster our confidence in the AQI-pessimism relation: the pattern is robust to different functional forms and treatment of outliers, and survives the inclusion of analyst and city fixed effects. [4] Finally, we show that the correlation between pollution and pessimism is stronger for firms that do not themselves produce high emissions. This finding helps to rule out the possibility that a firm's own pollution causes a negative inference about its environmental risks or productivity (indeed, our results may suggest the opposite).

We further enrich our understanding of the channel through which pollution impacts forecasting bias by examining factors that accentuate (or attenuate) the relation between AQI and earnings forecasts. First, we show that the link between pollution and forecast bias dissipates with the time elapsed between visit and forecast, as would be expected if the link between pollution and forecast pessimism were driven by analyst mood during a visit. We also find that the negative

---

[2]More precisely, visits are spread across the eastern half of China. Visits in the western provinces of Tibet and Xinjiang are rare, comprising only 1% of our main sample.

[3]This fact does not necessarily imply that pollution leads to better forecasts. See Lim (2001), for a discussion of why analysts who utilize management information on profitability may optimally provide forecasts that are positively biased.

[4]We also present placebo tests using AQI figures 5 to 10 days before and after the site visit. These non-visit pollution readings are unrelated to forecast optimism once we control for visit-date AQI, and the correlation between visit-date AQI and forecast optimism is unaffected by the inclusion of these "placebo" pollution controls.

pollution-forecast relation is driven by longer-term forecasts, which involve more guesswork and speculation by the analyst.

We then explore how the pollution-pessimism relation is affected by characteristics of visiting analysts. Most notably, the pessimism associated with pollution disappears for cases in which analysts from different brokerage firms visit the same site on the same date (there is no direct effect of multiple analysts on forecast bias), possibly suggesting a debiasing effect of multiple perspectives. However, there is no significant difference in the relation between pollution and forecast bias across individual analyst attributes that reflect ability or experience.

Finally, we provide suggestive evidence that analysts acclimate to severe pollution, by exploiting variation in pollution in cities where analysts are based. We find that the difference between site visit pollution and home pollution is predictive of bias, and in particular our main results are driven exclusively by analysts visiting sites in regions with higher pollution than their own. While these results are only suggestive, they represent a new finding and possible insight on environmental influences and mood – we know of no prior work that looks at whether acclimation to environmental conditions limits their affective influence.

This result on analyst acclimation also provides indirect evidence that the relation between pollution and forecasts is driven by the effect on analysts, rather than the effect of pollution on others (for example, corporate CEOs and other senior managers who address questions from analysts) that might indirectly impact analyst forecast. Further bolstering this interpretation, we conduct a textual analysis of transcripts of CEOs and other top executives' comments during site visit Q&As, and do not find that pollution leads to more negative responses by CEOs and other top executives.

Our findings contribute most directly to the large literature in accounting and finance on the behavioral biases of investment analysts and their role in financial markets (see, for example, Hirshleifer et al., 2018; Hong and Kubik, 2003; Hong and Kacperczyk, 2010). Most directly related to our work, Dehaan et al. (2017) look at the relation between weather and response to earnings announcements. They show that bad weather negatively affects the speed with which U.S. analysts respond to earnings announcements in adjusting their recommendations and (in contrast to ancillary findings we report in our results) that bad weather also leads to more pessimistic EPS forecasts and target prices. We view our work as complementary to theirs, given our focus on

4

different shifts in environmental conditions (weather versus pollution), different outcomes (forecast bias versus delay), and a distinct input into analyst decision-making, which is enabled by the disclosure rules governing Chinese analyses. Furthermore, our heterogeneity analyses provide a new window into the conditions that can exacerbate, or mitigate, the bias induced by ambient circumstances. Our results suggest important roles both for acclimation/adaptation and also group decision-making; these are findings that, to our knowledge, are new to the literature.

Our work also fits into the literature on how environmental conditions impact decision-making, discussed at the outset, and more broadly the literature on the extent to which decision-making in natural settings is afflicted by the biases and errors in judgment documented by behavioral economists and social psychologists, particularly among expert agents (see, for example, Harrison and List, 2008 on expertise and the winner's curse, and Haigh and List, 2005 on loss aversion among traders).

2. Background and data

Our data set is based on details gleaned from site visit disclosures for publicly traded Chinese firms, combined with analysts' reports issued in the 30 days following each visit. In the subsections that follow, we describe in greater detail the data sources and variable construction. In Appendix A, we describe the specifics of the final data set's construction.

2.1. Analyst site visits and forecasts

Since 2009, the Shenzhen Stock Exchange (SZSE) has mandated that all firms listed on the exchange must publicly disclose details about site visits, typically paid by stock analysts, mutual/hedge fund managers, reporters and individual investors, within two trading days of the visit, including all visitors' names, visit date, employers, and where the site visit took place.[5] (Firms listed on Shanghai Stock Exchange are not subject to this regulation.)

We limit our sample to cases in which the visitors' names are recorded, and the visitors are sell-side analysts from Chinese brokerage firms (87% of all visits).

These data are matched to analyst forecasts obtained from the Chinese Stock Market and

---

[5]When the site visit does not take place at the firm's headquarters, the record will generally list the exact location of the visit, which we use to match to our pollution and weather measures. For records that do not list a specific location, the site visit took place at the firm's headquarters.

Accounting Research (CSMAR) database, a commonly employed database available, for example, to North American researchers via Wharton Research Data Services. We look primarily at earnings forecasts issued in the 15 calendar days following a visit, to focus on assessments made as a result of information gathered on site. However, we will show patterns for samples of earnings reports with cutoffs as short as 5 calendar days and as long as 30 calendar days following the visit, to explore whether the effect of pollution dissipates with time.

Each earnings report may include multiple forecasts, for different time horizons. We control for time horizon in the analyses that follow, and maintain each forecast as a distinct (but non-independent) observation, as we will explore whether the relation between pollution and bias is affected by forecast horizon.

A natural concern with conditioning on the delay between site visits and earnings forecasts is that pollution may itself affect forecast timing. This possibility could in turn bias our estimates of the relation between pollution and forecast optimism. The direction of this bias is unclear – it depends on whether delayed forecasts tend to be more optimistic (which would induce a bias toward zero) or less optimistic (which would induce a negative bias). In Appendix B we show that the timing of earnings forecasts is in fact uncorrelated with site visit pollution, largely mitigating this concern.[6] A related concern is that analysts might time their visits to avoid high pollution days. In unreported analysis, however, we do not find that day-level pollution is correlated with site visit probability. Furthermore, even if pollution affected the choice of visit date, it implies no obvious relation between pollution and forecast bias.

Following Jackson (2005) and the vast literature in accounting on earnings forecasts, we define analysts' forecast optimism as follows:

$$Forecast\_Optimism_{ijt} = 100 * (FEPS_{ijt} - AEPS_{ijt})/P_j, \tag{1}$$

where $FEPS_{ijt}$ is analyst i's forecasted earnings per share (EPS) for firm j for year t, $AEPS_{ijt}$ is the realized EPS of firm j for year t, and P_j is firm j's stock price on the day prior to the earnings forecast. Following Huyghebaert and Xu (2016), we keep the EPS forecasts of all years in a report to explore whether pollution differentially affects analysts' forecast biases across various forecast

---

[6] While this finding may appear in tension with the findings of Dehaan et al. (2017), their emphasis is on processing time rather than affect. Furthermore, our measure of forecast delay is based on time elapsed following the site visit, during which time the analyst would have been working in their home city.

horizons.

2.2. Air quality and weather variables

For each city in China, we obtain the daily air quality index (AQI) from the official website of the Ministry of Environmental Protection of China (MEPC). These data are derived from daily air quality reports provided by province- and city-level environmental protection bureaus. The AQI is constructed based on the levels of six atmospheric pollutants: sulfur dioxide (SO_2), nitrogen dioxide (NO_2), suspended particulates smaller than 10 μm in aerodynamic diameter (PM10), suspended particulates smaller than 2.5 μm in aerodynamic diameter (PM2.5), carbon monoxide (CO), and ozone (O_3). Prior to 2014, the Chinese government monitored only SO_2, NO_2, and PM10, which was used to construct the air pollution index (API) that served as a summary measure of air quality in earlier years. While the API and AQI are not directly comparable, they are highly correlated (Zheng et al., 2014). For notational simplicity we refer to both as AQI in what follows. For a small fraction of city-day observations, the AQI readings are unavailable via the MEPC. We were able to fill in some of the missing data from the Qingyue Open Environment Data Center website, which obtains pollution data directly from local governments. [7]

The MEPC distinguishes among six categories of AQI: I-excellent (AQI≤50), II-good (50<AQI≤100), III-lightly polluted (100<AQI≤150), IV-moderately polluted (150<AQI≤200), V-heavily polluted (200<AQI≤300) and VI-severely polluted (AQI>300).[8]

Since an earlier literature suggests that weather can affect investors' moods and trading behavior, we collect weather data to match to analysts' site visits. Daily weather data are obtained from the 194 international meteorological stations in China, provided by the China Integrated Meteorological Information Service System. Variables include hours of sun, temperature, humidity, precipitation and wind speed. We match each city to the closest meteorological station based on straight line distance.

---

[7]The Qingyue Open Environment Data Center (https://data.epmap.org) is an organization which compiles environmental data from government sources and provides them freely to the public in standard data formats.

[8]The same six classifications were used both pre- and post-2014, though based on only three pollutants in the earlier period.

2.3. Firm and analyst characteristics

We control for basic firm attributes, including size ($log(Assets)$), market to book ratio, intangible asset ratio, stock price volatility, stock turnover, stock return, analyst attention, and industry (based on the China Securities Regulatory Commission (CSRC)'s 19 top-level industry categories). We also collected data on time-varying analyst characteristics, including the number of firms followed, and the number of forecasts made (we will include analyst fixed effects in our main specifications, which absorb the effects of any time-invariant analyst attribute). The analyst data were obtained from CSMAR and the firm controls from RESSET, a provider of Chinese financial research data.

Our main analysis sample is comprised of 3,824 earnings forecasts issued following 1,642 site visits (i.e., an average of 2.35 forecasts per visit). Extending the window to 30 calendar days, our longer sample includes 5,108 earnings forecasts, highlighting that the frequency of forecasts is considerably higher just following a site visit (the rate of drop-off is relatively rapid, with 2,756 of forecasts issued within 8 days).

We present summary statistics at the forecast-level in Panel A of Table 1, for the sample of visits for which the analyst provided a forecast within 15 calendar days. The sample mean and standard deviation of forecast optimism are 2.05 and 3.49, respectively, consistent with the prior literature which finds that sell-side analysts' earnings forecasts are generally higher than the realized values (e.g., Francis and Philbrick, 1993; Lim, 2001; Sedor, 2002). There is also considerable variation in analysts' excess optimism – the highest value is 63% and the lowest is -18 – though we will minimize the influence of these extreme errors by winsorizing the top and bottom 1% of observations (we will present the results without winsorizing to show that this step does not affect our conclusions). Panel B of Table 1 shows summary statistics for the firm-year variables.

3. Results

Our main analyses are based on specifications of the following form:

$$Forecast\_Optimism_{ijt} = \beta \times AQI_{ijt}/1000 + \boldsymbol{\gamma} \times \boldsymbol{X_{ijt}} + \epsilon_{ijt}, \tag{2}$$

where $\boldsymbol{X_{ijt}}$ is a vector of control variables including firm attributes, as well as industry, quarter, and analyst fixed effects. ϵ_{ijt} is the error term (clustered at the firm level). We divide AQI by 1,000 for ease of interpretation of the regression coefficients.

We present these results in Table 2, with all variables winsorized to limit the influence of outliers (results using non-winsorized data are provided in Appendix C, and show very similar patterns). For conciseness, we do not report the coefficients on control variables, though we provide the full regression output in Appendix D. Column (1) shows the bivariate relation between forecast optimism and air pollution. The negative coefficient on AQI indicates that higher pollution during a site visit is associated with lower forecasts relative to realized earnings. Its value of -3.56 indicates that a 1 standard deviation increase in (winsorized) air pollution of 48 is associated with a reduction in earnings forecast of approximately 0.17 percentage points, or a little less than 10% of the average over-optimism of forecasts for the sample overall. The inclusion of day-of-week and year×quarter fixed effects in column (2) reduces the coefficient on AQI by about 40% , though when we add industry, analyst and city fixed effects (column (3)) and firm, analyst and weather controls (column (4)), the coefficient becomes more negative, taking on values of -4.21 and -3.77 respectively. Across all specifications, the coefficient on AQI is significant at least at the 10% level.[9]

While our focus is on the link between pollution and forecast bias, we show the coefficients on weather-related covariates in Appendix D, and observe that the coefficient on hours of sunshine is very small and does not approach statistical significance (nor do the coefficients on any other weather-related variables). The lack of a weather-bias relation warrants some discussion because of its contrast to the positive relation between sunshine and stock market optimism observed by Hirshleifer and Shumway (2003), and also the positive relation between good weather and stock analyst forecasts as reported by Dehaan et al. (2017). While it is outside of the scope of our paper to fully explore the possible reasons for our distinct weather-optimism result, one possibility is that, given the severity of (and high variance in) pollution in our setting, its effect dominates

---

[9]Two other natural outcomes to consider are target price and recommendations. Unfortunately, we have relatively few target prices (475) in our data set that we can link to site visits, and in the case of recommendations there is very little variation – no analyst issues a sell recommendation, and 98% of the 1659 recommendations in our data set are either "strong buy" or "buy." When we do employ target price optimism or recommendation optimism as outcome variables, we obtain a point estimate on *AQI* that is of the same sign as in our analyses in Table 2, but in neither case does any coefficient approach statistical significance, which we view as unsurprising given the lack of statistical power.

other possible ambient influences of analysts' moods.

In Table 3 we allow for greater flexibility in the relation between pollution and forecast optimism, replacing the linear form on the right-hand side of Eq. (2) with a dummy variable for each of the Chinese government's six categories of air pollution (category I, least polluted, is the omitted category). The results suggest that the linear specification fits the data well. In particular, in the full specification in column (4) the coefficients are monotonically decreasing in pollution severity, with roughly comparable decreases in the coefficients for each pollution level.

We next turn to probing the robustness of our results using a placebo test based on pollution in days surrounding the site visit. These results highlight the distinct relation between pollution on the site visit date and subsequent earnings forecasts. While there is, naturally, correlation across days in a given city in the extent of pollution, there is also residual variation as a result of changes in temperature, winds, and other factors. This short-run variation allows us to look at the effect of air pollution several days apart from the site visit date. In Table 4, we repeat our favored (saturated) specification from column (4) of Table 2, including air quality measures for the 5, 7, and 10 days prior to the analyst's visit, as well as the 5, 7, and 10 days following the visit. The coefficient on visit date air quality is stable across all six specifications while, after accounting for visit date pollution, air pollution on surrounding dates has no predictive power.

While we have emphasized the effect of pollution on analyst affect as the likely mechanism for our main result, it is also possible that analysts' negative profit outlooks could result from CEO and/or top management mood during the visit. While this explanation would still involve a relation between pollution and affect, it is an explanation that is quite distinct from the one we have put forth to this point. To assess the plausibility of this mechanism, in Appendix E we use the fraction of negative words used by firm CEOs during site visit Q&As as the outcome variable. To generate this measure, we follow Loughran and McDonald (2011) to classify words during visit Q&A sessions (transcripts obtained from WIND, a provider of Chinese financial research data) as positive, negative, or neutral. We find that there is no significant relation between pollution and top management negativity during a visit, and indeed the point estimates are generally of the "wrong" sign.

We conclude this section by examining whether a firm's own pollution might be responsible for the patterns we document in our main results. To do so, we define the indicator variable

HighPollution to denote firms in one of the 16 industries classified as high polluters by the Ministry of Ecology and Environment. These include sectors such as thermal power, pulp and paper industry, and fermentation; collectively these industries comprise 24.5% of our site visit observations. If we were to find that the negative relation between pollution and earnings forecasts were driven by this high pollution subsample, one may be concerned that pollution from the firm itself might lead visitors to infer that the company could face environmental enforcement actions in the future, for example. In Table 5, we present our main specification augmented by the interaction of *AQI* and *HighPollution*. In column (1), in the absence of any industry fixed effects, we may observe the direct effect of *HighPollution* on forecast optimism.[10] We observe no correlation. When we add *AQI * HighPollution* as a covariate in column (2), we find that the coefficient is positive and roughly the same magnitude as the direct effect of *AQI*. This finding argues against the firm's own pollution as the source of the negative relation with earnings forecasts. Indeed, the positive coefficient on the interaction term may reflect a (relatively) positive attribution from pollution for firms whose production is itself the source of emissions.

3.1. Factors influencing the relation between pollution and forecasting bias

In this section we explore several dimensions of heterogeneity in the relation between pollution and forecasting bias. We do so with the aim of enriching our understanding the underlying mechanisms behind the effect of pollution on earnings forecasts, and of the factors that exacerbate or mitigate this relation.

We begin by examining two time-based dimensions of heterogeneity: the time elapsed between site visits and earnings reports, and the time horizon of forecasts in a given report. We then look at heterogeneity based on several characteristics of the visiting analysts. First, we explore whether pollution in an analyst's city of employment moderates the impact of site visit pollution on forecasting. We then examine heterogeneity based on the number of analysts visiting on a particular date, and also whether the analysts are from the same brokerage firm or different ones. And finally we examine whether individual analyst attributes that reflect ability or experience are associated with a stronger or weaker effect of pollution on forecasts.

---

[10]We can identify this relation despite the inclusion of industry fixed effects because the high pollution flag has some within-industry variation. For example, the CSRC industry classification for power includes both wind power and thermal power, whereas only the latter is classified as high pollution. If we include the more detailed industry fixed effects, the coefficient on the *AQI * HighPollution* interaction is largely unaffected.

Each of these analyses is motivated by a distinct intuition and prior research on circumstances that might be expected to amplify (or attenuate) the impact of pollution on analyst pessimism. We first look at the time elapsed because, to the extent that the negative relation between pollution and forecasts is driven by analyst affect, this effect might dissipate after departing from the (polluted) visit site. (Alternatively, if forecasts are calculated on-site and only reported later, we would expect no effect of delay on the pollution-forecast relation.) We are motivated to look at heterogeneity by forecast horizon based on earlier research in accounting, which finds that analysts' forecasts over longer horizons have less precision and are more prone to bias (Kang et al., 1994). If longer-run forecasts are based more on speculation (rather than hard data) we argue they are potentially more swayed by analysts' moods.

Our analysis of whether pollution in an analyst's work city mitigates the impact of site visit pollution is motivated by the literature on affective forecasting and adjustment (e.g., Wilson and Gilbert, 2003), which finds that individuals adjust relatively quickly to adverse circumstances. We are motivated to examine individual and group attributes of analysts to explore whether experience and ability – whether collective or individual – affect how ambient circumstances influence judgments.

3.1.1. Forecast delay

In Fig. 1, we illustrate how our estimates of the relation between air pollution and forecast optimism are affected by the inclusion of forecasts that are further removed in time from the site visit. In the graph, we present a series of point estimates of β from specification (2), allowing for a range of forecast windows (and using the fully saturated specification), ranging from 1 to 5 dates following the visit, to a [1,30] calendar day window. Interestingly, while the negative relation holds for all samples, it is sharpest for relatively short windows, and becomes insignificant for the longer windows in the figure. This finding provides suggestive evidence that the affective impact of air pollution (which, recall, is uncorrelated with the delay in providing subsequent forecasts) may dissipate with time. Naturally, there are alternative interpretations. For example, it is possible that visits which uncover little relevant information do not lead to earnings forecasts in the days that follow, so that the visit is irrelevant to forecasts generated some weeks later. It is for this reason that we treat our interpretation of these findings with caution.

3.1.2. Forecast horizon

We next explore whether pollution differentially affects forecasts over longer time horizons. To do so, we add the interaction term $AQI * log(Horizon)$ to specification 2, where $Horizon$ denotes the days elapsed between the forecast date and the corresponding date of the actual earnings announcement. To facilitate interpretation of the direct effects in this specification, we demean both AQI and $log(Horizon)$. We present the findings in Table 6, in specifications that parallel the presentation of our main results in Table 2. Focusing first on the direct effect of pollution and forecast horizon, we observe a modest negative association between pollution and forecast bias at the mean forecast horizon. Consistent with Kang et al. (1994), we see a much greater (positive) bias in forecasts over long horizons. Our main interest in this table is in the interaction of these two variables, which is consistently negative and significant at least at the 1% level across all columns, indicating a much stronger effect of pollution on longer-term forecasts. In the final column, we include an extra specification which includes analyst visit fixed effects. In this final column, all covariates are effectively absorbed by the 1,642 visit fixed effects, but we can still identify the forecast horizon term and its interaction with AQI, which vary within a site visit. Even in this saturated specification, the interaction term is negative and significant at the 1% level.

3.1.3. Analyst adaptation and the effects of pollution

We next turn to the adaptation hypothesis, which we emphasize is, to our knowledge, new to the analyst forecasting literature specifically, and a novel finding on forecasting bias more generally. We do so by examining whether the negative relation between pollution and earnings forecasts is driven by analysts based in less polluted cities. (Implicit in our examination of this question is the presumption that pollution's effect is asymmetric – exposure to pollution that is worse than one's usual experiences has a negative impact on affect, relative to the positive impact of experiencing relatively low pollution.)

In Table 7 we explore the "adaptability" hypothesis in a regression framework, in which we replace site visit AQI with a site visit spline with a kink at home-city AQI (i.e., the slope change will vary across analyst visits, with an analyst-specific knot in the spline, specifically captured

13

by the terms $\Delta AQI(When\ \Delta AQI < 0)$ and $\Delta AQI(When\ \Delta AQI \geq 0))$, where ΔAQI equals to $AQI - AQI\_home$ and AQI_home is the median AQI in the analyst's home city during the month preceding the site visit. We reprise the analyses of Table 2 with this substitution. Across all columns, the negative relation between AQI and forecast optimism is driven by analyst visits to sites that are more polluted than their home base. We note, however, that the negative portion of the spline is imprecisely measured so that we cannot reject equality of the two spline coefficients. As such, these results may be seen as merely suggestive. [11]

3.1.4. Individual analyst ability, experience, and forecast bias

We next turn to examining individual analyst attributes that could plausibly mitigate the effects of pollution on forecasting (and possibly reduce forecasting bias in general). Specifically, we consider the role of experience, as captured by (the log of) the number of quarters since the analyst's first forecast appeared, and two proxies for ability. The first is $Star$, an indicator variable denoting that the analyst is ranked as a star analyst by the New Fortune Magazine at the beginning of the visit year, and the second measures analyst forecast accuracy. To provide roughly comparable measures of accuracy for analysts with different experience levels, we focus on annual earnings forecasts made in the year prior to the site visit. $Accuracy$ is then defined as:

$$Accuracy = -\frac{1}{n}\Sigma_{i=1}^{n}\frac{|\widehat{EPS_i} - EPS_i|}{|EPS_i|}, \tag{3}$$

where n is the number of forecasts in the prior year, EPS_i is the realized earnings per share, and $\widehat{EPS_i}$ is the analyst's EPS forecast.

In the first four columns of Table 8, in which we look at the direct effect of analyst characteristics. Neither star status (column (1)) nor past accuracy (column (3)) is robustly associated with forecast optimism. However, experience is negatively associated with optimism, whether included on its own (column (2)) or together with other analyst attributes (column (4)). Given the optimism bias present on average, this result implies a higher level of accuracy among more experienced analysts. Of more direct relevance for our paper, we add the interaction of each variable with AQI in columns (5) – (7), and include all interactions in column (8). For the case of

---

[11]It is also natural to ask whether our spline specification is simply picking up on a non-linear or non-monotonic relation between site visit AQI and earnings forecasts. We observe, however, that a spline at the median of site visit AQI and a quadratic specification provide a poor fit for the data.

Star the coefficient suggests that star status may mitigate pollution-induced pessimism, though even in this instance the interaction is not statistically significant even at the 10% level. Overall, while we observe no evidence that the effects of pollution are mitigated by experience or ability, we cannot draw strong conclusions from these analyses given the imprecision of our estimates.

3.1.5. Group visits and forecast bias

In our final analyses we consider whether forecast bias is correlated with the presence of other analysts during the visit. We define two "group visit" variables. The first captures whether there is at least one other analyst from the same brokerage firm present (*GroupVisit_Same*), while the second measures whether there is at least one other analyst from another brokerage present (*GroupVisit_Other*). We are agnostic ex ante on the role of multiple visitors. On the one hand, "groupthink" can lead to magnification of individual biases (see, e.g., Janis, 1972 for a classic reference). The "wisdom of crowds" argues for the opposite – the aggregation of beliefs may help to erase individual errors. We distinguish between within-brokerage and cross-brokerage groups because one might, ex ante, expect the strength of these effects to differ between the two. In particular, we conjecture that analysts from the same brokerage will be more subject to the forces of social conformity, which is more apt to occur in groups with greater homogeneity in culture or attitudes (see Ishii and Xuan, 2014 for a discussion in a finance-focused setting).

We present results that show the direct effect of group visits (columns (1) – (3)) as well as their interactions with *AQI* (columns (4) – (6)) in Table 9. Neither type of group visit is a direct predictor of forecast optimism. When we include the interaction terms, we find a positive coefficient on $AQI * GroupVisit\_Other$, with a magnitude that is roughly equal to that of the direct effect of *AQI* (significant at the 5% level).[12] The interaction $AQI * GroupVisit\_Same$ is negative, though only marginally significant (p-value=0.100). The difference between the coefficients on the two interactions is significant at the 1% level.

---

[12]We consider whether the benefit of having analysts from other brokerages present may stem directly from the presence of other, less biased analysts. A natural approach to exploring this possibility is to examine whether the *GroupVisit_Other* finding is related to the adaptation results described in 3.1.3. That is, does the presence of other analysts help because it potentially adds the perspective of a visitor who is adapted to high pollution. To implement an empirical test of this idea, we take the adaptation specification, and ask whether the presence of others from high pollution cities (and hence adapted to pollution) mitigates the effect of pollution on pessimism, particularly for analysts that are themselves from low pollution locales (i.e., we interact $\Delta AQI(When\ \Delta AQI < 0)$ with a set of dummy variables denoting whether or not there is a "high adaptation" (high pollution) analyst visiting on the same date. In these specifications, none of the coefficients approaches significance, which we suggest may result in large part because of the inclusion of many highly correlated covariates.

Overall, these results suggest that the "wisdom of the crowds" effect may dominate for analysts from different (competing) brokerages, while groupthink dominates for visitors from the same brokerage. Naturally, these results and their interpretation should be treated as speculative – we have not attempted to model fully the decision to make site visits, let alone modeling whether visits are conducted by one or multiple analysts. We nonetheless believe these results – and our heterogeneity results more generally – to be provocative findings that may prompt further work in this area.

4. Conclusion

In this paper we study how environmental conditions impact sell-side analyst forecasts. We show that forecast optimism is lower following site visits on heavily polluted days, consistent with a negative impact of pollution on analyst affect. We further show that this effect is driven by the relation between pollution and forecasts issued soon after the site visit, suggesting that pollution's impact on affect dissipates with time. We also present suggestive evidence that the effect of pollution is weaker for analysts who themselves are based in highly polluted cities, consistent with analysts adjusting to the effects of poor air quality, and evidence that the effect of pollution is also weakened by the presence of analysts from other brokerage firms, suggesting that the "wisdom of the crowds" may mitigate the biases in individuals' judgments.

Our findings indicate that even expert agents may be influenced by apparently irrelevant environmental conditions, and furthermore, this takes place even in a high stakes setting. While finance scholars have focused on the impact of weather and pollution on stock prices and trading, it may be fruitful to extend this line of research to consider whether and how decisions of experts in other domains are impacted by environmental conditions: For example, are more bank loans rejected, or do economic forecasters issue more pessimistic macro predictions, on cloudy or polluted days? We may also delve more deeply into the conditions that lessen the influence of environmental factors, perhaps via required delays between environmental exposure and decision-making, or via a simple information treatment which informs decision-makers about the relation between environmental conditions and mood. We leave these avenues of inquiry for future research.

16

Appendix A: Data Set Construction

We begin our sample construction by hand collecting disclosures on site visits to all firms traded on the Shenzhen Stock Exchange. We obtained 22,200 such releases, covering 1,481 firms (and 67,443 visitors, including stock analysts, individual investors, mutual/hedge fund managers, and also reporters), over the period of 2009-2015. Based on this initial dataset, we use the following seven steps to assemble our final dataset which is used for our empirical analyses.

Step 1: Since we are primarily interested in sell-side analysts who provide earnings-per-share (EPS) forecasts, we only keep observations in which sell-side analysts released at least one forecast report within 30 days after the visit, leaving us with 5,004 firm-visit $\times$ analyst level observations.

Step 2: We then merge in site-date level AQI and weather information into the master dataset. For 486 out of 5,004 observations, we do not have corresponding AQI information, leaving us with 4,518 analyst site visits.

Step 3: Each analyst report potentially covers multiple forecasts for different horizons (current year, next year, EPS in two years, and so forth). Because we wish to test the relationship between forecast horizon and pollution-induced bias, we treat each forecast as a distinct (though non-independent) observation, leading to a total of 10,068 visit $\times$ analyst $\times$ EPS forecast level observations. Since we need to calculate forecast optimism using the realized EPS data, we drop 2 observations for which the forecast fiscal year is later than 2016, the final year of our data.

Step 4: We merge in financial information in year $t-1$ for the listed firms in our sample. 448 observations (4.5 percent) do not have matched pre-visit year financial data, leaving us with 9,618 observations. Among these matched observations, 843 observations have missing financial information on total assets, market/book value, intangible assets, stock turnover, annual stock return and daily volatility (all in year $t-1$), leaving us with 8,775 observations.

Step 5: We then merge in analyst-specific information, including the number of firms the analyst follows, and the number of forecast reports generated by the analyst, in year t. 1,613 (18.4 percent) observations do not have matched analyst-level information at all, leaving us with 7,162 observations.

Step 6: To control for the influence of weather, we then merge in weather information on the site visit date, including hours of sun, temperature, humidity, precipitation, and wind speed.

We also further dropped 47 observations with missing values for weather variables (which are recorded as missing by the meteorological station, and attributed to equipment malfunction or human error). This filter leaves us with 7,115 observations.

Step 7: Finally, since we merge in information on each analyst's city of employment during the three months prior to the site visit. This filter further reduced the sample by 2,007 observations, leaving us with 5,108 observations. In our main analysis, we restrict our sample to EPS forecasts released within 15 days of the site visit, giving us a final sample of 3,824 for our main analysis.

Appendix B. The Relation Between Air Pollution and Forecast Delay
Numbers in parentheses are standard errors clustered by firm. The sample covers the period from 2009 to 2015. The sample in columns 1 - 5 is confined to the set of earnings forecasts issued within 30 days of a site visit (i.e., $Delay \leq 30$), in column 6 the sample is limited to foreacsts issued within 15 days. The dependent variable in columns 1-4 and in column 6 is $Delay$, which denotes the number of days between the site visit and the issuance of the forecast. The dependent variable in column 5 is $log(Delay)$. AQI denotes the Air Quality Index of the site visit city on the visit date, scaled by 1,000. Controls include $log(Horizon)$, $Hours\_of\_Sun$, $Temperature$, $Humidity$, $Precipitation$, $Wind\_Speed$, $log(Assets)$, $Market\_to\_Book$, $Intangible\_Asset$, $Volatility$, $Turnover$, $Return$, $Analyst\_Attention$, $Follow\_Co\_Num$ and $Forecast\_Num$, with output suppressed to conserve space. See the notes to Table 1 for detailed definitions of the control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable			*Delay*		*log(Delay)*	*Delay*
AQI	-4.320	-1.186	6.515	4.681	0.224	-5.457
	(3.998)	(4.155)	(7.180)	(7.213)	(0.873)	(4.066)
Year-Quarter FEs		Yes	Yes	Yes	Yes	Yes
Day of Week FEs		Yes	Yes	Yes	Yes	Yes
Industry FEs			Yes	Yes	Yes	Yes
City FEs			Yes	Yes	Yes	Yes
Analyst FEs			Yes	Yes	Yes	Yes
Controls				Yes	Yes	Yes
Delay $\leq$	30	30	30	30	30	15
Observations	5108	5108	5108	5108	5108	3824
R-Squared	0.001	0.025	0.687	0.690	0.674	0.755

Appendix C. Robustness Tests for Main Regressions Without Winsorizing

Numbers in parentheses are standard errors clustered by firm. This table presents the results from Table 2, without winsorizing any of the continuous variables. The sample covers the period from 2009 to 2015. The dependent variable in all columns is *Forecast_Optimism*, which denotes the difference between annual EPS forecast issued within calendar days [1,15] of the site visit and realized EPS, scaled by price as of the trading day prior to the forecast, multiplied by 100. *AQI* denotes the Air Quality Index of the visit city on the visit day, scaled by 1,000. Controls include *log(Horizon)*, *Hours_of_Sun*, *Temperature*, *Humidity*, *Precipitation*, *Wind_Speed*, log(*Assets*), *Market_to_Book*, *Intangible_Asset*, *Volatility*, *Turnover*, *Return*, *Analyst_Attention*, *Follow_Co_Num* and *Forecast_Num*, with output suppressed to conserve space. See the notes to Table 1 for detailed definitions of the control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

	(1)	(2)	(3)	(4)
Dependent Variable		*Forecast_Optimism*		
AQI	-3.199***	-1.967*	-4.291***	-4.515***
	(1.110)	(1.126)	(1.429)	(1.653)
Year-Quarter FEs		Yes	Yes	Yes
Day of Week FEs		Yes	Yes	Yes
Industry FEs			Yes	Yes
City FEs			Yes	Yes
Analyst FEs			Yes	Yes
Controls				Yes
Observations	3824	3824	3824	3824
R-Squared	0.002	0.046	0.425	0.543

Appendix D. The Relation Between Air Pollution and Analyst Forecast Optimism
Numbers in parentheses are standard errors clustered by firm. The sample covers the period from 2009 to 2015. The dependent variable in all columns is *Forecast_Optimism*, which denotes the difference between annual EPS forecast issued within calendar days [1,15] of the site visit and realized EPS, scaled by price as of the trading day prior to the forecast, multiplied by 100. *AQI* denotes the Air Quality Index of the visit city on the visit day, scaled by 1,000. See the notes to Table 1 for detailed definitions of the control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

	(1)	(2)	(3)	(4)
Dependent Variable		*Forecast_Optimism*		
AQI	-3.558***	-2.129*	-4.206***	-3.769***
	(1.072)	(1.104)	(1.322)	(1.420)
log(*Horizon*)				1.596***
				(0.095)
Hours_of_Sun				-0.000
				(0.002)
Temperature				-0.002
				(0.001)
Humidity				-0.005
				(0.006)
Precipitation				0.000
				(0.001)
Wind_Speed				-0.008
				(0.008)
log(*Assets*)				0.212
				(0.132)
Market_to_Book				0.105
				(0.065)
Intangible_Asset				2.627
				(2.600)
Volatility				-1.912
				(18.856)
Turnover				-0.046
				(0.054)
Return				0.095
				(0.162)
Analyst_Attention				-0.313**
				(0.133)
Follow_Co_Num				-0.159
				(0.331)
Forecast_Num				0.214
				(0.245)
Year-Quarter FEs		Yes	Yes	Yes
Day of Week FEs		Yes	Yes	Yes
Industry FEs			Yes	Yes
City FEs			Yes	Yes
Analyst FEs			Yes	Yes
Observations	3824	3824	3824	3824
R-Squared	0.004	0.065	0.443	0.608

Appendix E. The Relation Between Air Pollution and Management Negativity

Numbers in parentheses are standard errors clustered by firm. The sample covers the period from 2009 to 2015. The dependent variable in all columns is $Management\_Negativity$, which denotes the number of negative words divided by total words of management answers during the Q & A session. AQI denotes the Air Quality Index of the site visit city on the visit date, scaled by 1,000. Controls include $log(Horizon)$, $Hours\_of\_Sun$, $Temperature$, $Humidity$, $Precipitation$, $Wind\_Speed$, $log(Assets)$, $Market\_to\_Book$, $Intangible\_Asset$, $Volatility$, $Turnover$, $Return$, $Analyst\_Attention$, $Follow\_Co\_Num$ and $Forecast\_Num$, with output suppressed to conserve space. See the notes to Table 1 for detailed definitions of the control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

	(1)	(2)	(3)	(4)
Dependent Variable		$Management\_Negativity$		
AQI	-0.011	-0.004	-0.016	-0.020
	(0.007)	(0.007)	(0.011)	(0.013)
Year-Quarter FEs		Yes	Yes	Yes
Day of Week FEs		Yes	Yes	Yes
Industry FEs			Yes	Yes
City FEs			Yes	Yes
Analyst FEs			Yes	Yes
Controls				Yes
Observations	3086	3086	3086	3086
R-Squared	0.003	0.038	0.754	0.758

References

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

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

Cheng, Q., Du, F., Wang, X., Wang, Y., 2016. Seeing is believing: analysts' corporate site visits. Review of Accounting Studies 21, 1245–1286.

Cheng, Q., Du, F., Wang, Y., Wang, X., 2018. Do corporate site visits impact stock prices? Contemporary Accounting Research 36, 359–388.

Cunningham, M. R., 1979. Weather, mood, and helping behavior: Quasi experiments with the sunshine samaritan. Journal of Personality and Social Psychology 37, 1947.

Dehaan, E., Madsen, J., Piotroski, J. D., 2017. Do weather-induced moods affect the processing of earnings news? Journal of Accounting Research 55, 509–550.

Francis, J., Philbrick, D., 1993. Analysts' decisions as products of a multi-task environment. Journal of Accounting Research 31, 216–230.

Goetzmann, W., Kim, D., Kumar, A., Wang, Q., 2015. Weather-induced mood, institutional investors, and stock returns. Review of Financial Studies 28, 73–111.

Haigh, M. S., List, J. A., 2005. Do professional traders exhibit myopic loss aversion? an experimental analysis. The Journal of Finance 60, 523–534.

Han, B., Kong, D., Liu, S., 2018. Do analysts gain an informational advantage by visiting listed companies? Contemporary Accounting Research 35, 1843–1867.

Harrison, G. W., List, J. A., 2008. Naturally occurring markets and exogenous laboratory experiments: A case study of the winner's curse. The Economic Journal 118, 822–843.

Hirshleifer, D., Levi, Y., Lourie, B., Teoh, S. H., 2018. Decision fatigue and heuristic analyst forecasts. Journal of Financial Economics 133, 83–98.

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

Hong, H., Kacperczyk, M., 2010. Competition and bias. The Quarterly Journal of Economics 125, 1683–1725.

Hong, H., Kubik, J. D., 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. The Journal of Finance 58, 313–351.

Huang, J., Xu, N., Yu, H., Forthcoming. Pollution and performance: Do investors make worse trades on hazy days? Management Science .

Huyghebaert, N., Xu, W., 2016. Bias in the post-ipo earnings forecasts of affiliated analysts: Evidence from a chinese natural experiment. Journal of Accounting and Economics 61, 486–505.

Ishii, J., Xuan, Y., 2014. Acquirer-target social ties and merger outcomes. Journal of Financial Economics 112, 344–363.

Jackson, A. R., 2005. Trade generation, reputation, and sell-side analysts. The Journal of Finance 60, 673–717.

Janis, I. L., 1972. Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes. Oxford, England: Houghton Mifflin .

Kamstra, M. J., Kramer, L. A., Levi, M. D., 2003. Winter blues: A sad stock market cycle. American Economic Review 93, 324–343.

Kang, S. H., O'Brien, J., Sivaramakrishnan, K., 1994. Analysts' interim earnings forecasts: Evidence on the forecasting process. Journal of Accounting Research 32, 103–112.

Lepori, G. M., 2016. Air pollution and stock returns: Evidence from a natural experiment. Journal of Empirical Finance 35, 25–42.

Levy, T., Yagil, J., 2011. Air pollution and stock returns in the us. Journal of Economic Psychology 32, 374–383.

24

Li, J. J., Massa, M., Zhang, H., Zhang, J., Forthcoming. Behavioral bias in haze: Evidence from air pollution and the disposition effect in china. Journal of Financial Economics .

Lim, T., 2001. Rationality and analysts' forecast bias. The Journal of Finance 56, 369–385.

Loughran, T., McDonald, B., 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. The Journal of Finance 66, 35–65.

Saunders, E. M., 1993. Stock prices and wall street weather. The American Economic Review 83, 1337–1345.

Schwarz, N., Clore, G. L., 1983. Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. Journal of personality and social psychology 45, 513.

Sedor, L. M., 2002. An explanation for unintentional optimism in analysts' earnings forecasts. The Accounting Review 77, 731–753.

Vert, C., Sánchez-Benavides, G., Martínez, D., Gotsens, X., Gramunt, N., Cirach, M., Molinuevo, J. L., Sunyer, J., Nieuwenhuijsen, M. J., Crous-Bou, M., et al., 2017. Effect of long-term exposure to air pollution on anxiety and depression in adults: A cross-sectional study. International journal of hygiene and environmental health 220, 1074–1080.

Wilson, T. D., Gilbert, D. T., 2003. Affective forecasting. Advances in experimental social psychology 35, 345–411.

Zheng, S., Cao, C.-X., Singh, R. P., 2014. Comparison of ground based indices (api and aqi) with satellite based aerosol products. Science of the Total Environment 488, 398–412.

Table 1: Summary Statistics

Forecast_Optimism denotes the difference between annual EPS forecast issued within calendar days [1,15] of the site visit and realized EPS, scaled by price as of the trading day prior to the forecast, multiplied by 100. *AQI* denotes the Air Quality Index of the site visit city on the visit date, scaled by 1,000. *log(Horizon)* denotes the natural logarithm of the days between the forecast date and the corresponding date of the actual earnings announcement. *Hours_of_Sun* denotes hours of sun of the site visit city on the visit date (0.1h). *Temperature* denotes the average temperature of the site visit city on the visit date (0.1℃). *Humidity* denotes the average humidity of the site visit city on the visit date (1%). *Precipitation* denotes the total precipitation of the site visit city on the visit date (0.1mm). *Wind_Speed* denotes the average wind speed of the site visit city on the visit date (0.1m/s). *log(Assets)* denotes the natural logarithm of total assets at the beginning of the year when the site visit took place (visit year). *Market_to_Book* denotes the ratio of market value of equity to book value of equity at the beginning of the visit year. *Intangible_Asset* denotes the ratio of intangible assets to total assets at the beginning of the visit year. *Volatility* denotes daily volatility of stock returns during the year prior to the visit year. *Turnover* denotes the daily turnover rate of the visit year. *Return* denotes annual stock returns of the year prior to the visit year. *Analyst_Attention* denotes the natural logarithm of the number of analysts following the firm during the visit year. *Follow_Co_Num* denotes the natural logarithm of the number of companies the analyst followed during the visit year. *Forecast_Num* denotes the natural logarithm of the number of reports issued by the analyst during the visit year. Panel A provides summary statistics based on the main sample of forecast × analyst visit observations. Panel B provides summary statistics collapsed to the firm-year level.

Panel A: Sample for Main Analysis

Variable Name	Mean	StdDev	Observations
Forecast_Optimism	2.051	3.486	3824
AQI	0.089	0.052	3824
log(Horizon)	5.920	0.828	3824
Hours_of_Sun	49.978	41.051	3824
Temperature	172.825	91.507	3824
Humidity	68.855	17.038	3824
Precipitation	37.127	109.890	3824
Wind_Speed	22.201	10.037	3824

Panel B: Firm-Year Aggregates

Variable Name	Mean	StdDev	Observations
log(Assets)	21.740	1.054	1046
Market_to_Book	3.124	1.743	1046
Intangible_Asset	0.045	0.050	1046
Volatility	0.028	0.006	1046
Turnover	2.787	2.163	1046
Return	0.253	0.615	1046
Analyst_Attention	2.428	0.755	1046
Follow_Co_Num	2.328	0.802	1046
Forecast_Num	2.867	1.049	1046

Table 2: The Relation Between Air Pollution and Analyst Forecast Optimism
Numbers in parentheses are standard errors clustered by firm. The sample covers the period from 2009 to 2015. The dependent variable in all columns is *Forecast_Optimism*, which denotes the difference between annual EPS forecast issued within calendar days [1,15] of the site visit and realized EPS, scaled by price as of the trading day prior to the forecast, multiplied by 100. *AQI* denotes the Air Quality Index of the visit city on the visit day, scaled by 1,000. Controls include $log(Horizon)$, $Hours\_of\_Sun$, $Temperature$, $Humidity$, $Precipitation$, $Wind\_Speed$, $log(Assets)$, $Market\_to\_Book$, $Intangible\_Asset$, $Volatility$, $Turnover$, $Return$, $Analyst\_Attention$, $Follow\_Co\_Num$ and $Forecast\_Num$, with output suppressed to conserve space. See the notes to Table 1 for detailed definitions of the control variables. Appendix D shows the results including point estimates for all control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

	(1)	(2)	(3)	(4)
Dependent Variable		*Forecast_Optimism*		
AQI	-3.558***	-2.129*	-4.206***	-3.769***
	(1.072)	(1.104)	(1.322)	(1.420)
Year-Quarter FEs		Yes	Yes	Yes
Day of Week FEs		Yes	Yes	Yes
Industry FEs			Yes	Yes
City FEs			Yes	Yes
Analyst FEs			Yes	Yes
Controls				Yes
Observations	3824	3824	3824	3824
R-Squared	0.004	0.065	0.443	0.608

Table 3: The Effect of Different AQI Categories
Numbers in parentheses are standard errors clustered by firm. The sample covers the period from 2009 to 2015. The dependent variable in all columns is $Forecast\_Optimism$, which denotes the difference between annual EPS forecast issued within calendar days [1,15] of the site visit and realized EPS, scaled by price as of the trading day prior to the forecast, multiplied by 100. $AQI50 - 100$, $AQI100 - 150$, $AQI150 - 200$, $AQI200 - 300$, and $AQI300+$ are indicator variables that correspond to each of the government's air pollution categories ($AQI < 50$ is the omitted category). See the text for details. Controls include $log(Horizon)$, $Hours\_of\_Sun$, $Temperature$, $Humidity$, $Precipitation$, $Wind\_Speed$, $log(Assets)$, $Market\_to\_Book$, $Intangible\_Asset$, $Volatility$, $Turnover$, $Return$, $Analyst\_Attention$, $Follow\_Co\_Num$ and $Forecast\_Num$, with output suppressed to conserve space. See the notes to Table 1 for detailed definitions of the control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

	(1)	(2)	(3)	(4)
Dependent Variable		$Forecast\_Optimism$		
$AQI50 - 100$	-0.006	0.099	-0.232	-0.376
	(0.152)	(0.150)	(0.221)	(0.230)
$AQI100 - 150$	-0.342*	-0.161	-0.567**	-0.664**
	(0.181)	(0.191)	(0.256)	(0.262)
$AQI150 - 200$	-0.443**	-0.255	-0.779***	-0.856***
	(0.223)	(0.215)	(0.269)	(0.297)
$AQI200 - 300$	-0.567*	-0.296	-1.228***	-1.062***
	(0.293)	(0.287)	(0.366)	(0.371)
$AQI300+$	-1.522***	-0.988***	-1.057*	-1.190*
	(0.289)	(0.340)	(0.578)	(0.624)
Year-Quarter FEs		Yes	Yes	Yes
Day of Week FEs		Yes	Yes	Yes
Industry FEs			Yes	Yes
City FEs			Yes	Yes
Analyst FEs			Yes	Yes
Controls				Yes
Observations	3824	3824	3824	3824
R-Squared	0.006	0.067	0.444	0.609

Table 4: The Effect of Pollution Persistence on Forecast Optimism

Numbers in parentheses are standard errors clustered by firm. The sample covers the period from 2009 to 2015. The dependent variable in all columns is $Forecast\_Optimism$, which denotes the difference between annual EPS forecast issued within calendar days [1,15] of the site visit and realized EPS, scaled by price as of the trading day prior to the forecast, multiplied by 100. AQI denotes the Air Quality Index of the visit city on the visit day, scaled by 1,000. AQI_Past5, AQI_Past7, and AQI_Past10 denote AQI of the site visit city 5, 7, and 10 days prior to the visit date respectively, scaled by 1,000. $AQI\_Forward5$, $AQI\_Forward7$, and $AQI\_Forward10$ denote AQI of the site visit city 5, 7, and 10 days following the visit date respectively, scaled by 1,000. Controls include $log(Horizon)$, $Hours\_of\_Sun$, $Temperature$, $Humidity$, $Precipitation$, $Wind\_Speed$, $log(Assets)$, $Market\_to\_Book$, $Intangible\_Asset$, $Volatility$, $Turnover$, $Return$, $Analyst\_Attention$, $Follow\_Co\_Num$ and $Forecast\_Num$, with output suppressed to conserve space. See the notes to Table 1 for detailed definitions of the control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable			$Forecast\_Optimism$			
AQI	-3.548**	-3.813***	-3.772***	-3.782***	-3.841***	-3.540**
	(1.481)	(1.443)	(1.420)	(1.407)	(1.421)	(1.461)
AQI_Past5	-1.501					
	(1.649)					
AQI_Past7		0.378				
		(1.346)				
AQI_Past10			-0.360			
			(1.474)			
$AQI\_Forward5$				0.181		
				(1.764)		
$AQI\_Forward7$					0.501	
					(1.374)	
$AQI\_Forward10$						-1.410
						(1.142)
Year-Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes
Day of Week FEs	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FEs	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3824	3822	3822	3824	3824	3824
R-Squared	0.609	0.608	0.608	0.608	0.608	0.609

Table 5: The Effect of Firm Type

Numbers in parentheses are standard errors clustered by firm. The sample covers the period from 2009 to 2015. The dependent variable in all columns is $Forecast\_Optimism$, which denotes the difference between annual EPS forecast issued within calendar days [1,15] of the site visit and realized EPS, scaled by price as of the trading day prior to the forecast, multiplied by 100. AQI denotes the Air Quality Index of the visit city on the visit day, scaled by 1,000. $HighPollution$ is a dummy variable indicating that the visited firm belongs to one of the 16 high pollution industies defined by Ministry of Ecology and Environment of China (see text for details). Controls include $log(Horizon)$, $Hours\_of\_Sun$, $Temperature$, $Humidity$, $Precipitation$, $Wind\_Speed$, $log(Assets)$, $Market\_to\_Book$, $Intangible\_Asset$, $Volatility$, $Turnover$, $Return$, $Analyst\_Attention$, $Follow\_Co\_Num$ and $Forecast\_Num$, with output suppressed to conserve space. See the notes to Table 1 for detailed definitions of the control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

	(1)	(2)
Dependent Variable	$Forecast\_Optimism$	
$HighPollution$	0.344	-0.279
	(0.349)	(0.473)
AQI	-3.625**	-5.378***
	(1.461)	(1.604)
$AQI * HighPollution$		7.410**
		(2.992)
Year-Quarter FEs	Yes	Yes
Day of Week FEs	Yes	Yes
Industry FEs		
City FEs	Yes	Yes
Analyst FEs	Yes	Yes
Controls	Yes	Yes
Observations	3824	3824
R-Squared	0.601	0.602

Table 6: The Relation Between Air Pollution and Forecast Optimism for Different Forecast Horizons

Numbers in parentheses are standard errors clustered by firm. The sample covers the period from 2009 to 2015. The dependent variable in all columns is *Forecast_Optimism*, which denotes the difference between annual EPS forecast issued within calendar days [1,15] of the site visit and realized EPS, scaled by price as of the trading day prior to the forecast, multiplied by 100. *AQI* denotes the (demeaned) Air Quality Index of the visit city on the visit date, scaled by 1,000. *log(Horizon)* denotes the (demeaned) natural logarithm of the days elapsed between the forecast date and the corresponding date of the actual earnings announcement. Controls include *Hours_of_Sun*, *Temperature*, *Humidity*, *Precipitation*, *Wind_Speed*, *log(Assets)*, *Market_to_Book*, *Intangible_Asset*, *Volatility*, *Turnover*, *Return*, *Analyst_Attention*, *Follow_Co_Num* and *Forecast_Num*, with output suppressed to conserve space. See the notes to Table 1 for detailed definitions of the control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

	(1)	(2)	(3)	(4)	(5)
Dependent Variable			*Forecast_Optimism*		
AQI	-1.817*	-2.285**	-3.836***	-4.388***	
	(1.049)	(1.135)	(1.345)	(1.468)	
log(Horizon)	1.607***	1.603***	1.629***	1.630***	1.634***
	(0.075)	(0.075)	(0.093)	(0.094)	(0.106)
log(Horizon) * AQI	-3.711***	-3.676***	-4.407***	-4.337***	-4.282***
	(0.964)	(0.952)	(1.340)	(1.356)	(1.563)
Year-Quarter FEs		Yes	Yes	Yes	
Day of Week FEs		Yes	Yes	Yes	
Industry FEs			Yes	Yes	
City FEs			Yes	Yes	
Analyst FEs			Yes	Yes	
Controls				Yes	
Visit FEs					Yes
Observations	3824	3824	3824	3824	3824
R-Squared	0.213	0.236	0.609	0.612	0.693

Table 7: Air Pollution Adaption and Forecast Optimism

Numbers in parentheses are standard errors clustered by firm. The sample covers the period from 2009 to 2015. The dependent variable in all columns is $Forecast\_Optimism$, which denotes the difference between annual EPS forecast issued within calendar days [1,15] of the site visit and realized EPS, scaled by price as of the trading day prior to the forecast, multiplied by 100. ΔAQI equals to $AQI - AQI\_home$. AQI denotes the Air Quality Index of the site visit city on the visit date, scaled by 1,000. AQI_home is the median AQI in the analyst's home city during the month preceding the site visit, scaled by 1,000. Controls include $log(Horizon)$, $Hours\_of\_Sun$, $Temperature$, $Humidity$, $Precipitation$, $Wind\_Speed$, $log(Assets)$, $Market\_to\_Book$, $Intangible\_Asset$, $Volatility$, $Turnover$, $Return$, $Analyst\_Attention$, $Follow\_Co\_Num$ and $Forecast\_Num$, with output suppressed to conserve space. See the notes to Table 1 for detailed definitions of the control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

	(1)	(2)	(3)	(4)
Dependent Variable		$Forecast\_Optimism$		
$\Delta AQI (When \ \Delta AQI \leq 0)$	5.612**	3.072	0.381	-0.395
	(2.593)	(2.472)	(4.719)	(4.529)
$\Delta AQI (When \ \Delta AQI > 0)$	-5.035***	-3.679***	-4.993***	-3.885**
	(1.446)	(1.388)	(1.648)	(1.711)
Year-Quarter FEs		Yes	Yes	Yes
Day of Week FEs		Yes	Yes	Yes
Industry FEs			Yes	Yes
City FEs			Yes	Yes
Analyst FEs			Yes	Yes
Controls				Yes
Observations	3824	3824	3824	3824
R-Squared	0.004	0.066	0.443	0.608

Table 8: Pollution, Analyst Skills, and Forecasting Bias
Numbers in parentheses are standard errors clustered by firm. The sample covers the period from 2009 to 2015. The dependent variable in all columns is *Forecast_Optimism*, which denotes the difference between annual EPS forecast issued within calendar days [1,15] of the site visit and realized EPS, scaled by price as of the trading day prior to the forecast, multiplied by 100. *AQI* denotes the Air Quality Index of the visit city on the visit day, scaled by 1,000. *Star* is a dummy variable denoting whether the visiting analyst is ranked as a star by New Fortune magazine in the visit year. *Experience* is measured as the natural logarithm of the number of quarters since the analyst make his/her first forecast up to the end of the visit year. *Accuracy* is the average accuracy of the analyst's forecast within the past 1 year, with accuracy defined as the absolute difference between annual EPS forecast and realized EPS, scaled by realized EPS, multiplied by -1. Controls include $log(Horizon)$, *Hours_of_Sun*, *Temperature*, *Humidity*, *Precipitation*, *Wind_Speed*, $log(Assets)$, *Market_to_Book*, *Intangible_Asset*, *Volatility*, *Turnover*, *Return*, *Analyst_Attention*, *Follow_Co_Num* and *Forecast_Num*, with output suppressed to conserve space. See the notes to Table 1 for detailed definitions of the control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

|	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable				*Forecast_Optimism*				
AQI	-3.770***	-3.746***	-4.005***	-4.020***	-4.075***	-3.570	-3.063	-3.028
	(1.423)	(1.423)	(1.458)	(1.460)	(1.480)	(2.870)	(2.583)	(3.534)
Star	0.073			0.140	-0.230			-0.259
	(0.273)			(0.275)	(0.388)			(0.379)
Experience		-0.295**		-0.317**		-0.287		-0.309
		(0.125)		(0.129)		(0.201)		(0.205)
Accuracy			0.080	0.079			0.049	0.044
			(0.066)	(0.065)			(0.109)	(0.108)
AQI * Star					3.624			4.857
					(3.123)			(3.070)
AQI * Experience						-0.089		-0.149
						(1.252)		(1.243)
AQI * Accuracy							0.364	0.430
							(0.782)	(0.776)
Year-Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of Week FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3824	3824	3757	3757	3824	3824	3757	3757
R-Squared	0.608	0.609	0.608	0.608	0.609	0.609	0.608	0.608

Table 9: The Effect of Analyst Group Visit on Optimism

Numbers in parentheses are standard errors clustered by firm. The sample covers the period from 2009 to 2015. The dependent variable in all columns is $Forecast\_Optimism$, which denotes the difference between annual EPS forecast issued within calendar days [1,15] of the site visit and realized EPS, scaled by price as of the trading day prior to the forecast, multiplied by 100. AQI denotes the Air Quality Index of the visit city on the visit day, scaled by 1,000. $Group\_Visit\_Same$ is an indicator variable denoting that at least one other analyst from the same brokerage was present during the visit. $GroupVisit\_Other$ is an indicator variable denoting that at least one other analyst from a different brokerage was present during the visit. Controls include $log(Horizon)$, $Hours\_of\_Sun$, $Temperature$, $Humidity$, $Precipitation$, $Wind\_Speed$, $log(Assets)$, $Market\_to\_Book$, $Intangible\_Asset$, $Volatility$, $Turnover$, $Return$, $Analyst\_Attention$, $Follow\_Co\_Num$ and $Forecast\_Num$, with output suppressed to conserve space. See the notes to Table 1 for detailed definitions of the control variables. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable			$Forecast\_Optimism$			
AQI	-3.864***	-3.768***	-3.864***	-3.311**	-8.064***	-7.882***
	(1.432)	(1.420)	(1.432)	(1.525)	(2.447)	(2.459)
$GroupVisit\_Same$	-0.277		-0.277	0.039		0.087
	(0.222)		(0.222)	(0.349)		(0.343)
$GroupVisit\_Other$		0.019	0.017		-0.536*	-0.580*
		(0.166)	(0.165)		(0.315)	(0.316)
$AQI*GroupVisit\_Same$				-3.633		-4.684*
				(2.930)		(2.842)
$AQI*GroupVisit\_Other$					6.150**	6.752**
					(2.654)	(2.656)
Year-Quarter FEs	Yes	Yes	Yes	Yes	Yes	Yes
Day of Week FEs	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FEs	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3824	3824	3824	3824	3824	3824
R-Squared	0.609	0.608	0.609	0.609	0.609	0.610

Figure 1: The Attenuating Effect of Forecast Delay
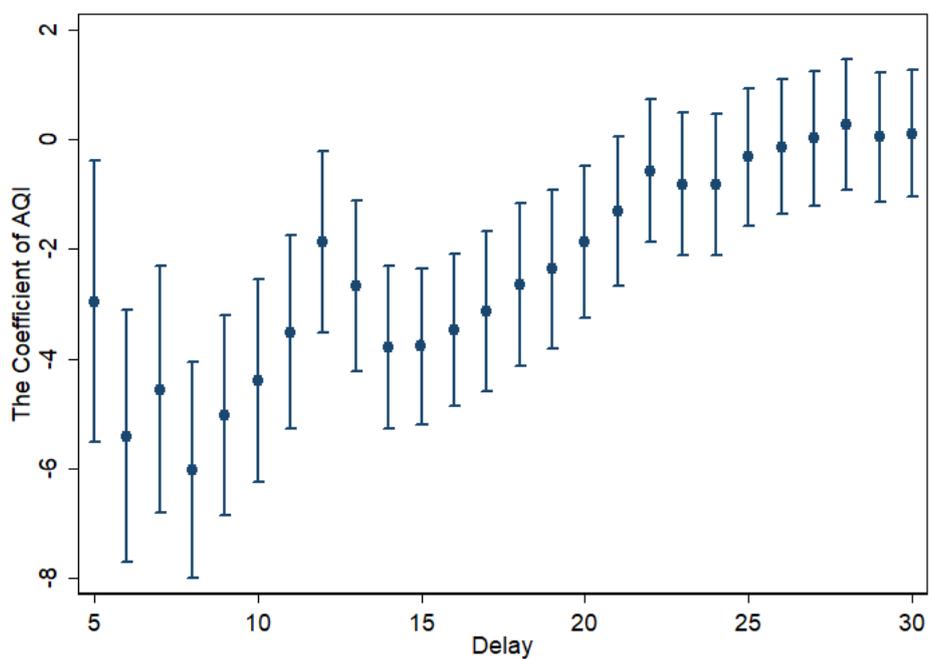
This figure shows how the coefficient estimates of AQI vary as a function of the number of days between analyst site visits and subsequent earnings forecasts. Each circle indicates the point estimate from Equation (1), including the full set of controls, and includes all forecasts issued up to and including d days after the site visit, where d ranges from 5 to 30. The whiskers show the 95 percent confidence interval of each coefficient estimate.