



**HAL**  
open science

## Policy change, mass media and air quality in China: new paths to face air pollution?

Sébastien Marchand, Damien Cubizol, Elda Nasho Ah-Pine, Huanxiu Guo

### ► To cite this version:

Sébastien Marchand, Damien Cubizol, Elda Nasho Ah-Pine, Huanxiu Guo. Policy change, mass media and air quality in China: new paths to face air pollution?. 2023. hal-03448375v2

**HAL Id: hal-03448375**

**<https://uca.hal.science/hal-03448375v2>**

Preprint submitted on 13 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



Centre d'Études  
et de Recherches  
sur le Développement  
International

## CERDI WORKING PAPERS

2023/6

---

### **Policy change, mass media and air quality in China: new paths to face air pollution?**

Sébastien Marchand  
Damien Cubizol  
Elda Nasho Ah-Pine  
Huanxiu Guo

---

To cite this working paper:

Marchand, S., Cubizol, D., Nasho Ah-Pine, E., Guo, H. "Policy change, mass media and air quality in China: new paths to face air pollution?", CERDI Working Papers, 2023/6, CERDI.

CERDI, Pôle tertiaire, 26 avenue Léon Blum, 63000 Clermont-Ferrand, France

## The authors

Sébastien Marchand

Associate professor, Université Clermont Auvergne,  
CNRS, IRD, CERDI, F-63000 Clermont-Ferrand, France

Email: [sebastien.marchand@uca.fr](mailto:sebastien.marchand@uca.fr)

Damien Cubizol

Associate professor, Université Clermont Auvergne,  
CNRS, IRD, CERDI, F-63000 Clermont-Ferrand, France

Email: [damien.cubizol@uca.fr](mailto:damien.cubizol@uca.fr)

Elda Nasho Ah-Pine

Associate Professor, ClerMA, ESC Clermont Business  
School, Clermont-Ferrand, France

Email: [elda.nasho@esc-clermont.fr](mailto:elda.nasho@esc-clermont.fr)

Huanxiu Guo

Assistant professor, Nanjing Audit University, Nanjing,  
Jiangsu, China

Email: [hxguo@nju.edu.cn](mailto:hxguo@nju.edu.cn)



This work was supported by the LABEX IDGM+ (ANR-10-LABX-0014) within the program “Investissements d’Avenir” operated by the French National Research Agency (ANR).

CERDI Working Papers are available online at: <https://tinyurl.com/2xwfw8s>

Director of Publication: Simone Bertoli

Editor: Catherine Araujo Bonjean

Publisher: Marie Dussol

ISSN: 2114 - 7957

**Disclaimer:** Working papers are not subject to peer review, they constitute research in progress. Responsibility for the contents and opinions expressed in the working papers rests solely with the authors. Comments and suggestions are welcome and should be addressed to the authors.

## **Abstract**

In response to the dramatic increase in air pollution-related illnesses and socioeconomic risks, China has undertaken a large-scale public policy to combat air pollution since 2013. The mass media has become ever since a major actor of this public policy through its important role in diffusing environmental news, and shaping people's perceptions and emotions regarding air pollution. In this study, we combine the GDELT media emotion database with daily air pollution monitoring data to investigate how the emotional tone in environmental news improves or worsens air quality. We find that more negative emotions in environmental news can decrease air pollution at both the city and county levels, and several robustness checks confirm this effect. Moreover, we discuss two potential mechanisms related to the context of China's war on air pollution in which the government relies on mass media as an informal regulator of air pollution. We demonstrate that firms reduce their pollution emissions after negative media emotions, and even more in cities that are more prone to environmental sanctions. Building on these results, we propose a reflection on the adoption of effective pro-environmental public policies in line with the SDGs beyond the Chinese context.

## **Keywords**

Air pollution; news emotional tone; mass media; informal regulation; China

## **JEL Codes**

D83; D9; L82; O53; Q53; Q58

# 1 Introduction

Decades of unprecedented economic growth from massive fossil fuel consumption and lax environmental regulations have led to huge degradation of air quality in China (Barwick et al., 2020), with a dramatic increase in air pollution-related illnesses and socio-economic risks. According to Yin et al. (2020), in 2017, particulate matter (PM) was the fourth leading risk factor for deaths and disability-adjusted life-years (DALYs) in China, and about 10% of annual deaths (1.03 million deaths) in China were attributable to ambient air pollution (PM 2.5 and ozone). Long-term exposure to ambient air pollution is associated with various chronic diseases, including stroke, ischaemic heart disease, chronic obstructive pulmonary disease (COPD), lung cancer (Yin et al., 2017; Yu et al., 2018), and even cause major insomnia disorders (Heyes and Zhu, 2019). Beyond the health costs, air pollution also disrupts social and economic stability by causing: the migration of skills to affect the performance of companies (Xue et al., 2021); a serious increase in student absenteeism (Chen et al., 2018); a decrease in firms' labor share (Zhou and Zhang, 2023); an increase in loan amounts for health and consumption (Chen et al., 2022); and even a drop in the number of people attending cultural places such as movie theaters (He et al., 2022).

Given these serious effects on public health and socioeconomic life, China has implemented vigorous public policies to combat air pollution since 2013. To this end, a stringent environmental policy has been applied by the Chinese government through the use of information-based instruments and the launching of a national program of real-time air quality monitoring and data streaming systems. This program was implemented to improve the accuracy of ambient air pollution data and facilitate public access to real-time pollution information on the Internet (Greenstone et al., 2020). In the meantime, mentions and reports of air pollution issues have become frequent and common, both in traditional newspapers and on the Internet (Wang and Jia, 2021), which has facilitated unprecedented diffusion of environmental news in China. In this context, the extent of public attention to the deterioration in air quality following the launch of the national program in 2013 has provoked a significant emotional mobilization in Chinese society (Tu et al., 2020). The uprising in environmental news could go hand in hand with this public emotionalism, through the emotional tone of the articles. Every online article on environmental news has a scale of emotion whose tone can be defined by a text analysis as negative (e.g. expressing fear, anxiety, aversion, sorrow, distress, terror), positive (e.g., expressing hope, joy, attraction, enthusiasm, pride), or neutral (Shapiro et al., 2020). Emotion is a content feature of a message and so a kind of stimulus property of the message (Zheng et al., 2021). This emotional tone has important implications for how people interpret, understand, and act on information and content. It can thus elicit responses from their readers and modify their perceptions of environmental risks. For instance, strong emotional responses to environmental fearsome risks are found to trigger overreaction and excessive preventative measures (Sunstein and Zeckhauser, 2011).

In this context, our study addresses a series of questions: does the surging environmental news, in particular, the emotional tone disseminated in the news improve or worsen the environmental quality in China? Put differently, can mass media amplify or attenuate public perceptions of air pollution risk according to how the information is presented? Moreover, how might the government rely on mass media as an informal regulator of polluters' behaviors at the local level? Despite the widespread policy interest in these questions, the answer from existing empirical studies is very limited. Yet a deeper understanding of these questions could shed light on the micro foundation of the mass media and generate broader public participation in the informal regulation of air pollution. To fill in the gaps in the literature, our study investigates the causal relationship between the emotions contained in mass media and air pollution levels in China. We also identify actors that adjust their pollutant emissions (firms) and potential mechanisms (anticipation of sanctions).

Our study is related to three strands of the literature. First, the literature on the importance of the use of information-based instruments in environmental regulation (Duflo et al., 2013) including studies on the impact of pollution information disclosure on polluters' behaviors in China (Barwick et al., 2020; Greenstone et al., 2020; Ito and Zhang, 2020; Zhang et al., 2022); second, the literature on the role of policy in media (Brady, 2009; Stockmann and Gallagher, 2011; Wang et al., 2014) and, third, the literature on the role of mass media and especially the Internet as a powerful tool for shaping public perceptions and awareness of environmental issues (Happer and Philo, 2016; D'Amato et al., 2019; Ha et al., 2019; Tu et al., 2020). Our study contributes to the literature in three ways. Firstly, we wish to demonstrate that public policy can act indirectly on economic agents through the media. In the Chinese context, the national real-time air quality disclosure program has enabled the media to take up the subject and relay the information, giving the government a large-scale relay. Secondly, we wish to demonstrate that the mass media can act as informal regulators of the behavior of firms and households generating air pollutants. This regulation is informal in the sense that it is not officially established by the government or competent authorities, even if it depends on them, in this case through the air quality disclosure program. Finally, we wish to reveal the importance of the emotional tone of press articles. Indeed, the regulating effect of the mass media is not only due to the quantity of information disseminated but also to its emotional content. In the Chinese context, the extent of awareness of the deterioration in air quality following the launch of the national program in 2013 has provoked a significant emotional charge in Chinese society (Tu et al., 2020). We would like to analyze this through the emotional tone of press articles.

From a methodological point of view, we collect data on the emotional tone in environmental news using the Global Database of Events, Language, and Tone (GDELT). The GDELT is an open-source database that provides daily information (e.g., theme, emotion, number of words, etc.) on news reports around the globe using the Natural Language Processing (NLP) technique<sup>1</sup>. We complement the GDELT with daily monitoring data of four air pollutants (Ni-

---

<sup>1</sup>The GDELT project was created by Kalev Leetaru, an internet entrepreneur, academic and senior fellow

trogen Dioxide (NO<sub>2</sub>), Sulfur Dioxide (SO<sub>2</sub>), Particulate Matter 2.5 (PM 2.5) and Particulate Matter 10 (PM<sub>10</sub>) at the city and the county level<sup>2</sup>. In the first step, we focus on these two levels because local governments are involved in the national “war on pollution”. They have to include air pollution prevention and control in their development planning and implement concrete measures to meet the national environmental quality standards (Greenstone et al., 2021). We compile these datasets, as well as meteorological data, to construct a comprehensive and high-frequency database to empirically investigate the impact of the emotional tone in environmental news on air pollution. In the second step, we use data on firms’ pollutant emissions and sanctions at the city level to identify the actors that react to the emotional tone in the mass media and the potential mechanisms.

We focus on daily news about air pollution reported by mass media in 2018 to tackle potential endogeneity problems of emotional tone in environmental news. More precisely, we rely on the cutoff created by the announcement of the second stage of the Chinese government’s “war on air pollution” to implement a Fuzzy Regression Discontinuity Design (FRDD)<sup>3</sup>. This strategy estimates the local impact of the emotional tone in environmental news around the announcement date. It allows us to separately test whether the action plan caused discontinuous changes in the emotional tone of media news and in the air pollution immediately before and after the announcement.

We find consistent results regarding the impact of emotional tone in environmental news on air pollution reduction (at both the city and county levels) within the context of China’s war on air pollution. We find a positive effect of the emotional tone in articles published up until three weeks before the day of observed air pollution (more negative (positive) emotions in environmental news can decrease (increase) air pollution). These results are consistent for all air pollutants (NO<sub>2</sub>, SO<sub>2</sub>, PM<sub>2.5</sub>, and PM<sub>10</sub>) at both the city and county levels, and are robust to the RDD strategy and several robustness tests. We attribute this positive effect of emotional tone to the context of China’s war on air pollution and the use of media as a tool for informal regulation (which strengthens the existing environmental regulations). More precisely, we discuss two potential mechanisms related to the Chinese context by exploiting data on corporate pollutant emissions (for about 4,500 firms located nationwide) and data on the number of sanctions after the violation of air pollution standards (at the city level). We find that corporate emissions are positively influenced by media emotions and that the impact of emotional tone on air pollution is stronger in cities that are more prone to sanctions. Finally,

---

at George Washington University, along with Philip Schrodt, a political scientist who developed the CAMEO event taxonomy used by the GDELT project. The current database includes the news from January 1, 1979, until today.

<sup>2</sup>Cities are the second level of the administrative hierarchy in China, also known as prefecture-level cities after the Provincial-level administrative divisions (the first level). Four municipalities (Beijing, Tianjin, Shanghai, and Chongqing) in our sample are considered as cities although they have the same rank as provinces. As such, these municipalities are simultaneously both a city and a province in its own right. Our sample is made up of 2,486 counties located in 286 cities and 31 provinces.

<sup>3</sup>This second stage takes the form of a Three-Year Action Plan implemented for Winning the Blue Sky Defense Battle. We detail this plan in the second section.

based on these empirical results and our conceptual considerations, we discuss other elements to understand public or private agents' attitudes that are most massively involved in air pollution. This discussion also contributes to the reflection on how to adopt effective pro-environmental public policies beyond the Chinese context, a main issue that needs to be addressed in every country as required within the SGDs goals.

The remainder of the paper is structured as follows. Section 2 presents the institutional background and the role of media as an informal regulator of air pollution. Section 3 describes the data. Section 4 discusses the econometric framework. Section 5 reports the main econometric results. Section 6 presents the robustness tests. Section 7 analyzes potential mechanisms and Section 8 discusses policy implications of our result beyond the Chinese contexts. Finally, this study is concluded in Section 9.

## 2 Background and literature

### 2.1 The key role of the information in air pollution control

A paradigm shift in air pollution control has occurred in China since 2013 insofar as the Chinese government made efforts to establish a real public policy on air pollution.

#### 2.1.1 Limited public awareness of air pollution in China prior to 2013

From the beginning of the 1980s to the early 2000s, environmental regulations in China aimed to reduce the emission of sulfur dioxide (SO<sub>2</sub>), considered the primary threat to air quality<sup>4</sup>. Since the early 2000s, fine particulate matter ("PM") has become the main source of air pollution. However, PM is produced from mixed sources (not only coal burning) such as automobile emissions, industrial activities, and rapid urbanization, especially due to China's accession to the World Trade Organization in 2001. Environmental regulation in China, therefore, shifted to the implementation of air quality standards at the city level. For instance, performance evaluations of local officials included the number of "blue-sky days" they resided over. However, these efforts have had little effect due to weak enforcement at the local level, given the competing incentives for economic growth (Wu and Cao, 2021) and the limited monitoring from the central government (Barwick et al., 2020). As a result, air pollution became severe and blue-sky days became rare in many urban centers, especially during winter. The failure of environmental policy to address air pollution was accompanied by a low level of public awareness prior to 2013. First, while the Air Pollution Index (API) for major cities was published by the Ministry of Environmental Protection (MEP) in 2000, this indicator did not include the quantity of

---

<sup>4</sup>Chinese environmental regulations on air pollution was launched in 1982 when China established its first national ambient air quality standards (NAAQs). These standards set limits for six air pollutants (total suspended particles (TSP), coarse particulate matter (PM<sub>10</sub>), sulfur dioxide (SO<sub>2</sub>), nitrogen oxides (NO<sub>x</sub>), carbon monoxide (CO), and ozone (O<sub>3</sub>)) and were amended in 1996, 2000 and 2012, respectively. To save space, we refer to Barwick et al. (2020) for a detailed presentation of Chinese environmental regulations prior to 2013.



PM 2.5, the main pollutant responsible for dense smog. In addition, the local government was responsible for publishing the API at the local level but with limited supervision from the central government, and manipulation of API data at the local level occurred ([Andrews, 2008](#); [Ghanem and Zhang, 2014](#)). Therefore, official newspapers and mass media rarely reported the API index and the public awareness of PM 2.5 was very limited prior to 2013. [Barwick et al. \(2020\)](#) refer to this situation as the public confusion of “fog-smog” and illustrate it with one example where the media characterized the dense smog that hit Beijing and northern cities during the winter of 2011 as fog.

### 2.1.2 New environmental policy and surging public awareness of air pollution

In 2012, the MEP revised the NAAQs for the fourth time, and for the first time, set national standards to regulate PM 2.5 emissions. In September 2013, the State Council of China issued the Air Pollution Prevention and Control Action Plan for 2013-2017 to achieve these new standards, with specific targets for PM 2.5 reduction and ten concrete policies<sup>5</sup>. Later, the 13th Five-Year Plan (2016-2020) set objectives for cities to reduce PM 2.5 concentrations by 18% between 2015 and 2020. To meet this objective, a second Three-year Action Plan to Win the Blue Sky Defense War (abbreviated as the Blue Sky Defense War) was launched at the end of June 2018 by the State Council. This action plan proposed new targets for SO<sub>2</sub>, NO<sub>2</sub>, and PM 2.5 levels, and promoted the green development plan for specific areas (e.g., “2+26” cities in the Beijing-Tianjin-Hebei region) and specific sectors such as the manufacturing industry (e.g., the control of production capacities of some manufacturing goods (steel, etc.)) and the energy and transportation sectors.

The new environmental policies were a success in terms of pollution reduction according to several studies ([Feng et al., 2019](#); [Huang et al., 2018](#); [Jiang et al., 2021](#); [Zhang et al., 2019, 2022](#)). For example, [Geng et al. \(2021\)](#), the first Action Plan reduced the annual population-weighted mean PM<sub>2.5</sub> from 62.5  $\mu\text{g}/\text{m}^3$  in 2013 to 44.4  $\mu\text{g}/\text{m}^3$  in 2017 while the second Action Plan further reduced the PM<sub>2.5</sub> concentrations to 33.1  $\mu\text{g}/\text{m}^3$  in 2020<sup>6</sup>.

Also, as a result of better air quality, there were substantial reductions in mortality and years of life lost over this period ([Huang et al., 2018](#); [Yin et al., 2020](#); [Zhang et al., 2019](#)). This success was made possible by the creation of a nationwide monitoring-and-disclosure program that aims to publicly disclose pollution data in real-time, avoid the manipulation of pollution data by local government and increase public awareness of air pollution. The monitoring system allows for continuous reading of firm-level pollution data in real-time through the installation of a comprehensive network of monitors. According to [Barwick et al. \(2020\)](#), by the end of the last wave of installations in November 2014, more than 1,400 monitoring stations in 337 cities covering an estimated 98% of the country’s population were built. The system also allows

---

<sup>5</sup>See [Yang \(2020\)](#) for a detailed presentation of this new Central government’s air pollution policy.

<sup>6</sup>However, the air pollution level remains high in comparison with the World Health Organization 2021 guidelines that state the air quality guideline levels at 5  $\mu\text{g}/\text{m}^3$  for PM<sub>2.5</sub> ([WHO, 2021](#)).

the dissemination of pollution data to the public. A real-time index called the Air Quality Index (AQI) was created and displayed in real-time on MEP's website. In contrast to the previous API reporting system, now the local government cannot manipulate pollution data because they are directly transmitted to MEP's information center in real time. Hourly and daily information on the AQI and concentrations of PM2.5, PM10, O3, CO, NO2, and SO2 are available at the individual station, and at the firm and city levels. As a result, many private websites and mobile phone applications have been developed to stream real-time air quality information from the official website to the public. All these new tools facilitate public access to pollution information. As a consequence, newspapers and other mass media broadcast more and more reports on air pollution.

## 2.2 The mass media as a key informal regulator of air pollution

In the context presented above, our study assumes that online mass media play an important role mainly in transmitting the central government's new environmental policy in its war on air pollution. As we will show in our empirical section, online mass media can act as a sort of informal regulator of polluters' behaviors at the local level through the information and emotions disseminated in the news reports. This relationship between government, mass media and air pollution relies on three strands of the literature to which our study is connected.

First of all, an interesting strand of the literature in the context of our study concerns the role of policy in media (Brady, 2009). In China, the media is mainly under government oversight and can be reliable means for controlling and incentivizing the implementation of public policies (Stockmann, 2013; Zhan, 2020). For instance, Stockmann and Gallagher (2011) finds that the Chinese media contribute to regime legitimacy by shaping citizens' attitudes and behavior toward the legal system in the context of labor problems and labor rights. From a content analysis of two Shanghai newspapers and one television show, the authors show that media representation of labor conflict often promotes the successful use of the legal system by a normal citizen, which enhances perceptions of the government's legitimacy through the rule of law. Another example is the study of Wang et al. (2014). The authors use content analysis to investigate the media coverage of nuclear power in China using two prominent Chinese newspapers, *The People's Daily* and *The Guangming Daily*, from 2004 to 2013. They show that almost all of the articles of both newspapers presented informational or pro-nuclear statements in accordance with the Chinese government's nuclear power policies, whereas little anti-nuclear opinion was expressed. We add to this literature by showing that since 2013 the Chinese government has relied on the media as an important transmission channel for its air pollution policy.

Second, our study is connected to the literature on the impact of pollution information disclosure in China. Since 2013 and the creation of a nationwide monitoring-and-disclosure program, online mass media has begun to become one of the primary channels of pollution

information for the general public. With this in mind, a growing body of literature studies the impact of pollution information disclosure in China. For example, [Barwick et al. \(2020\)](#) document the profound impacts of pollution information disclosure on household awareness of pollution issues and the triggering of behavioral changes in household activities. [Greenstone et al. \(2020\)](#) argue that the information disclosure program significantly improved the quality of China’s air pollution data and led to more avoidance behaviors in Chinese households. [Ito and Zhang \(2020\)](#) show that the willingness to pay estimates based on air purifier sales was larger after the pollution information disclosure. Finally, [Zhang et al. \(2022\)](#) find that China’s air quality monitoring and disclosure program has significantly increased the green innovation of firms, especially the number of green invention patent applications. We contribute to this literature by showing that information disclosure in terms of environmental news by online newspapers can also act on polluters’ behaviors.

Third, and in line with our previous developments, our study is also focused on the role of online mass media as a powerful instrument for shaping public perceptions and awareness about environmental issues. For instance, regarding Chinese households, [Tu et al. \(2020\)](#) demonstrate that the release of the environmental documentary “*Under the Dome*” in 2015 significantly increased public willingness to pay for better environmental quality and raised public awareness of environmental health risks. In Malaysia, [Ha et al. \(2019\)](#) found that mass media (TV, newspaper and the Internet) plays a role in consumers’ awareness, attitudes and knowledge of climate change as well as of environmentally friendly behavior (e.g., use of public transport, reduction of the amount of waste). [Happer and Philo \(2016\)](#) also argue that the mass media plays a powerful role in shaping public understanding of climate change. They find that people can change their awareness and attitude but often do so with unmet behavioral commitments, due to a lack of credibility and accuracy in mass media. From Eurobarometer surveys (EBs) on attitudes of European citizens towards the environment, [D’Amato et al. \(2019\)](#) show that environmental information on the Internet plays a significant role in stimulating pro-environmental behaviors (waste reduction, recycling and energy saving). Traditional media (tv, radio and newspapers) were found to mainly influence energy saving. The authors suggest that the explanation relies on the maturity of the topic of energy saving in traditional media compared to other measures such as recycling. Regarding firms, the media is also a channel through which environmental protection can become a key element of Responsible Business Conduct (RBC) in Chinese companies. According to the OECD, the concept of RBC for companies encompasses compliance with social expectations such as human rights, environmental protection, ethical governance, etc., but through channels other than laws<sup>7</sup>. These channels can be inter-governmental organizations, within the workplace, or via the press. The press cannot be neglected and falls within the scope of our study: in addition to environmental laws, firms may feel compelled to react and adjust their pollution emissions to meet the social expectations conveyed by the media.

---

<sup>7</sup>See the OCDE website for a more detailed presentation at <https://www.oecd.org/investment/toolkit/policyareas/responsiblebusinessconduct/>.

RBC has been promoted in China since the 2000s by the OECD (OECD, 2008). In addition to meeting social expectations, RBC in environmental protection also creates economic incentives for Chinese firms, particularly for their insertion in international value chains and their outward investments. By connecting the emotional tone of online news to air pollution, our study contributes to the investigation of the role played by mass media in shaping public perceptions and awareness of air pollution in China.

In summary, in the Chinese context, we assume that newspapers and other mass media play an important role in transmitting the national environmental policy against air pollution. They act as a sort of informal regulator of the polluters' behaviors through the information and emotions disseminated in the news reports. Is this informal regulation effective in reducing local air pollution? This remains an empirical issue, to which we now turn.

### 3 Data and summary statistics

#### 3.1 Media data and the GDELT project

We use the Global Database of Events, Language, and Tone (GDELT) project<sup>8</sup> to collect mass media information. The GDELT project extracts events, themes, and emotions from media articles and the persons/organizations involved and creates networks that tie them together. The database covers web, print and broadcast news media in over 100 languages from every country in the world and is updated every 15 minutes. With such vast coverage and high-frequency updating, it is considered as the most comprehensive dataset for study of global mass media.

In this study, we focus on articles in China (mostly in Chinese) related to air pollution in 2018. These articles are derived from popular mass media in China such as China daily (English and Chinese), China News Service, Caixin.com, Sina.com, Sohu.com, ifeng.com (Fenghuang news), China.com, Qianlong.com and Global News(Huanqiu), that provide daily news information to the public. It is noted that about 50% of the mass media are both digital and print newspapers, which covered a large national and foreign audience in China.

Using the BigQuery platform, We extract the following five variables for all articles and calculate for each of them the daily mean of emotional tone, positive emotional tone, negative emotional tone, polarity and number of words. The emotional tone is our main variable of interest. This variable is a combination of the positive and negative emotional connotations transmitted by an article through the words used. It is calculated for each article as the positive score minus the negative score of the article. Common values range between -10 and +10, with 0 indicating neutrality. In our database, the tone has a mean value of -2.7, a minimum value of -14.4 and a maximum value of 4.0 (see Table A1). These values suggest that articles tend to be negatively charged in terms of emotional tone. These results are relevant in the Chinese

---

<sup>8</sup>Source: <https://www.gdeltproject.org/>.

context, as explained in the previous section. Positive and negative emotional tone is calculated with positive and negative scores, respectively. These scores are the percentage of all words in the article that were found to have a positive or negative emotional connotation. They measure the degree to which the article contains words with positive or negative connotations. In our sample, the mean value of the negative score (5.0) is higher than the mean value of the positive score (2.6), reflecting the negative value of the emotional tone at the sample mean. According to the GDELT, the polarity of an article is “The percentage of words that had matches in the tonal dictionary as an indicator of how emotionally polarized or charged the text is. If the polarity is high, but the tone is neutral, this suggests the text was highly emotionally charged, but had roughly equivalent numbers of positively and negatively charged emotional words.” Note that an article can have a tone close to zero (the net emotional connotation of positive–negative words) but a high polarity (the gross emotional connotation of words, positive+negative), meaning that the article is emotionally charged with approximately the same quantity of positive and negative emotional connotations. At the sample mean, the polarity is about eight percent. Finally, we use the number of words in the article, which represents the volume of information in the article. The sample mean is 715 words per article.

For a deeper understanding of these media variables, we describe below three examples of articles. An article from *China Daily* (a major media outlet in China) published on April 7, 2018, describes the improvements in air quality due to the use of more natural gas and less coal<sup>9</sup>. The emotional tone is positive with a value of 1.1, a positive score of 3.3 and a negative score of 2.2. The positive emotional connotations of words (“improvement”, “enjoy”, “higher quality”, etc.) exceed the negative emotional connotations of words (“worse”, “hazardous”, etc.). The polarity (positive plus negative score) is thus above five, meaning that the article is emotionally charged, even if the tone is only equal to 1.1. An article from *South China Morning Post* (a Hong-Kong media outlet) published on April 11, 2018, alerts readers of ozone pollution that is worsening in Northern China<sup>10</sup>. The tone is strongly negative (-5.8) because of the high negative score (7.6) compared to the positive score (1.8). The article contains many words with negative connotations (“worsened”, “hazardous”, etc.), and the text is highly emotionally charged with a polarity score of 9.4. Another article from *South China Morning Post* published on October 2, 2018, explains that air pollution in China causes an average of 1.1 million premature deaths each year and costs more than 250 billion yuan a year<sup>11</sup>. Many words with negative emotional connotations increase the negative score (“killing”, “costing”, “damage”, etc.). Despite some words with positive emotional connotations, the tone remains negative (-3.9).

To check the relevancy of our media sample on environmental news, we also propose word clouds that emerge the most from important media in China (Figure 1). We include the

---

<sup>9</sup>Available at: <http://www.chinadaily.com.cn/a/201804/07/WS5ac805b2a3105cdcf65168a7.html>.

<sup>10</sup>Available at <http://www.scmp.com/news/china/policies-politics/article/2141259/harmful-ozone-pollution-worsening-northern-china-says>.

<sup>11</sup>Available at: <https://www.scmp.com/news/china/science/article/2166542/air-pollution-killing-1-million-people-and-costing-chinese>.

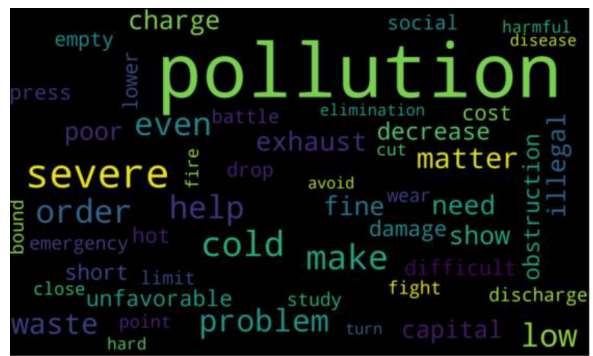


period 2018 with media in both English and Chinese, and we also split between words with negative and positive connotations (from General Inquirer Dictionary Harvard IV, used in the GDELT program). In the top-left cloud, the most prominent words with positive connotations are “quality”, “protection”, “health”, “improve”, “law”, “clean”, “health”, etc.; in the top-right cloud, the most prominent words with negative are “pollution”, “severe”, “charge”, “exhaust”, “problem”, “fight”, “discharge”, “fine”, “damage”, “battle”, “waste”, etc. Moreover, the bottom-left cloud reports words that emerge the most without positive or negative connotations. The words “pollution”, “air”, “emission”, “China”, “coal”, “environmental”, “Beijing”, and “protection” are the most prevalent. These words suggest both the relevancy of the articles used in our analysis and the presence of many words without emotional tone such as “China”, “emission” and “air”. The bottom-right cloud includes verbs and shows that “will” and “said” are strongly prevalent. This result highlights the various announcements of the central government’s plans to fight pollution.

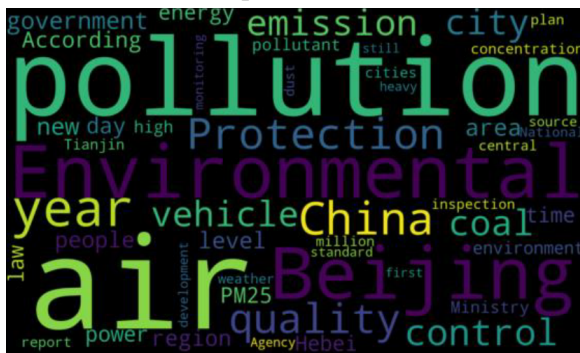
Figure 1: Word clouds of articles on air pollution in Chinese media articles published in 2018 (written in English and Chinese)



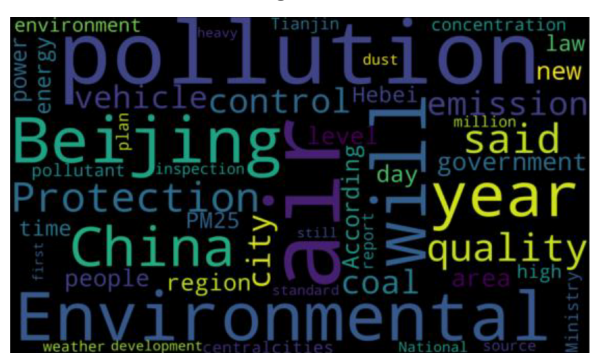
Words with positive connotation



Words with negative connotation



All words (without verbs)

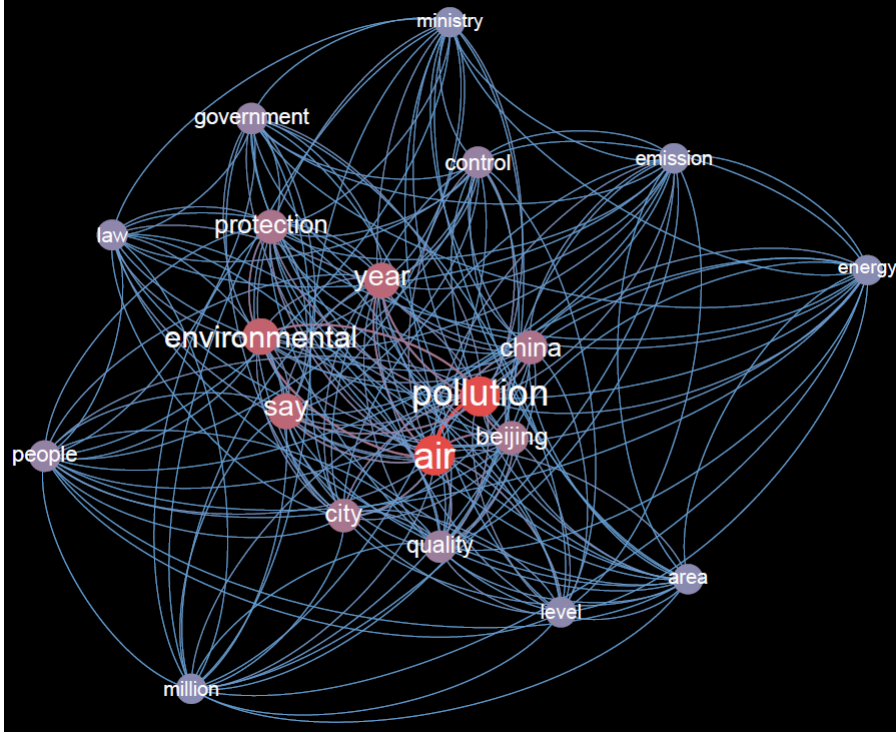


All words (with verbs)

Figure 2 is a network graph and connects the most quoted words of our media articles. When two words are often quoted together, the line that ties them together becomes red and thicker. This graph enables us to show the relevance and efficiency of the GDELT algorithm that selected our media articles on air pollution in China. The words “air”, “pollution”, “environmental”, “protection”, “China” and “Beijing” are often quoted together in the selected media

articles. Our dataset also contains articles with the words “government”, “environmental”, “protection”, “ministry”, “law”, “control”, etc., mentioned together to highlight the willingness of the government to fight air pollution in China.

Figure 2: Network graph (most quoted words in media articles dataset)



### 3.2 Air pollution data

We use air pollution databases at county-level and city-level<sup>12</sup>. At the city level, we collect daily monitoring data on the main air pollutants (the AQI index, PM2.5, PM10, SO2, NO2) from the China National Environmental Monitoring Center (2018) for about 300 cities. At the county level, we collect daily high-resolution data on the main air pollutants (PM2.5, PM10, SO2, NO2, CO and O3) from the Chinese Air Quality Reanalysis dataset (Tang et al., 2020b) for about 2,400 counties.

We focus on pollutants and major contributors of air pollution that are common to all databases: fine particle pollution with diameters that are generally 2.5 micrometers and smaller (PM2.5), coarse particle pollution with diameters that are generally 10 micrometers and smaller (PM10), nitrogen dioxide (NO2) and sulfur dioxide (SO2). These pollutants are mainly emitted by the burning of fossil fuels (coal and petroleum) and also by the smelting of iron ores for SO2 from power plants, industry and automobiles. Each pollutant is measured in  $\mu\text{g}$  per  $\text{m}^3$ . From

<sup>12</sup>A further description of the databases is available in A.1 of the Appendix.

A1 in the Appendix, the sample mean at the city level is 29  $\mu\text{g}$  of NO2 per m<sup>3</sup>, 14  $\mu\text{g}$  of SO2 per m<sup>3</sup>, 41  $\mu\text{g}$  of PM25 per m<sup>3</sup> and 77  $\mu\text{g}$  of PM10 per m<sup>3</sup><sup>13</sup>.

### 3.3 Weather data

Air pollution is very sensitive to daily weather conditions. Therefore, we supplemented pollution data with weather data from the China Meteorological Data Sharing Service System (<http://cdc.cma.gov.cn/>). Meteorological data include daily readings from more than 800 weather stations in China. The meteorological variables were then aggregated to the county or city level by averaging the daily readings of all the weather stations within a county or city. These data measure for each day the highest temperature, the lowest temperature, the wind speed and the type of weather (e.g., sunny, rainy, etc.). See A1 in the Appendix for the summary statistics of weather data at both the city and county levels.

## 4 Estimation strategy

### 4.1 Baseline model

We implement the following econometric model on our panel datasets:

$$E_{i,t}^p = \alpha + \beta_1 \text{Tone}_{t-n} + \beta_2 \text{Polarity}_{t-n} + \beta_3 \text{Count}_{t-n} + \sum_{k=1}^4 \gamma_k \text{weather}_{i,t}^k + mFE + \varepsilon_{i,t}, \quad (1)$$

where  $t$  is the day,  $i$  is either city or county,  $p$  is the pollutant and  $n = 1, 7, 14, 21, 28, 35$ .  $E_{i,p,t}$  is the daily emissions of individual  $i$  at day  $t$  by pollutant  $p$  (SO2, NO2, PM2.5 and PM10). Our variable of interest is the daily media emotional tone (*Tone*) that we control by the emotional charge of the article (*Polarity*) and the number of words of the article (*Count*). For these three media variables, we consider the mean of their values during the previous periods to capture all the information during the last days/weeks. Thus, in our model  $\text{Var}_{t-n}$  is not the value of the variable at  $t - n$  but rather the mean of the variable values from  $t = -1$  to  $t - n$  ( $\text{Mean}(\text{Var}_{i,p,t-1}; \dots; \text{Var}_{i,p,t-n})$ ,  $n = 1, 7, 14, 21, 28, 35$ ). Note that the media variables are defined at a daily basis but are common to all  $i$ .

We control for the daily weather at the county and city levels with four variables: highest and lowest temperature, type of weather (rainy, sunny, etc.) measured by an index variable and wind speed (see Section 3 for data sources). *mFE* are multi-level fixed effects (ID, weekday and upper-level ID\*month) that enable us to control for many characteristics at different levels (for example in the county database, upper-level ID\*month can control for intensive use of heating

---

<sup>13</sup>Annual national standard are 40  $\mu\text{g}$  of NO2 per m<sup>3</sup>, 20  $\mu\text{g}$  of SO2 per m<sup>3</sup>, 15  $\mu\text{g}$  of PM25 per m<sup>3</sup> and 40  $\mu\text{g}$  of PM10 per m<sup>3</sup>.



in a city during a winter month)<sup>14</sup>. Variables are in levels or logarithms (no 1st-difference because variables are stationary). For the calculation of standard errors, we cluster both at the ID and the upper-level ID (province for city, and city for county) in order to take into account serial and spatial correlations.

We analyze the potential impact of the emotional tone of environmental news on air pollution in two steps. First, our baseline model is to estimate the model with a one-week average lag ( $n = 7$ ) for the tone variable and other media variables at the city and country levels. We assume that a one-week average lag for media variables is a convenient and relevant time span. Second, still at the county and city levels, we use different lags on media variables in addition to the one-week average lag. More precisely, we use the following lags:  $n = 1, 14, 21, 28, 35$ . For  $n = 1$ , this is the value of the media variable one day before the observed pollution (so in  $t = -1$ ). For  $n = 14, 21, 28, 35$ , it is equivalent to gradually adding older news to the average of media variables for  $n = 7$  while keeping the previous values until five weeks ( $n = 35$ ). More precisely, for instance, for  $n = 14$ , we use the mean of the variable values from  $t = -1$  to  $t = -14$ . By so doing we try to capture a memory effect of emotional tone in environmental news on air pollution.

To verify the robustness of these baseline regressions, as described in Section 6 a series of robustness tests are conducted: the addition of new media variables, trends, economic/financial variables, and a different measure of lagged variables.

## 4.2 Positive and negative emotions

Positive and negative emotions do not shape the actions of individuals in the same way. In the Chinese context, we expect that individuals will be more sensitive to negative emotions transmitted by online newspapers. Regarding the stringency of China’s environmental policy in terms of restrictions and controls to tackle air pollution, as well as the government’s huge effort to develop awareness of the detrimental impacts of air pollution on human health, we expect individuals to react more to negative than to positive emotions.

To investigate this differential effect of positive versus negative emotions, we replicate the specifications of Model 1 by alternatively replacing the emotional tone with the positive and negative scores. These scores give the percentage of all words in an article with a positive or negative emotional connotation.

## 4.3 RD estimation

In the baseline model, we regress air pollution on the emotional tone by considering the latter as exogenous while we can assume that the emotional tone is endogenous for several reasons.

---

<sup>14</sup>We use the *reghdfe* estimators on Stata to include multi-level fixed effects and take into account collinearity within the fixed effects. This estimator also supports different slopes per individual and is faster than *xtreg* (it is all the more important given the high number of observations in our study).

There may be a simultaneity bias even if we think it remains low because (i) we use lags (tone articles from  $t - 1$  to  $t - n$ ), (ii) we have few articles on current critical pollution levels (only a few observations can be concerned by a potential impact of air pollution on tone articles) and (iii) we use online articles published at national level (through the Internet) to explain pollution at local level. Moreover, there may be an omitted variables bias that could plague the estimation of the effect of the tone.

Therefore, to isolate a causal impact of the emotional tone of articles on levels of air pollution, we implement a regression discontinuity design (RDD) which has been developed as a quasi-experimental method to estimate the causal impacts in various fields, such as education, psychology and economics (Imbens and Lemieux, 2008). Specifically, we exploit the official announcement of the government’s “Three-year Action Plan” on June 27, 2018, as a cutoff to separately test whether there were discontinuous changes in emotional tone in news media and air pollution immediately before and after the announcement.

As mentioned above, this plan was announced on the State Council website on June 27, 2018. It proposed new targets in terms of indicators for SO<sub>2</sub>, NO<sub>2</sub> and PM as well as the scope of the plan’s key target areas (i.e., Beijing-Tianjin-Hebei and surrounding areas; the Yangtze River Delta region, including Shanghai and Jiangsu; the Fenwei plain with the cities of Jinzhong, Yuncheng, etc.). The plan also promoted the green development of industry (e.g., the control of production capacities of some manufactured goods (steel, etc.)), and of the energy and transportation systems<sup>15</sup>.

The identification strategy hypothesis is that the announcement is exogenous information with respect to pollution at the local level, impacting it only through the media in the short term. This assumption implies three other underlying assumptions. First, we assume that online media react immediately to the announcement of the new action plan (which is relevant because of the Chinese context marked by strong proximity between politics and media) which, in turn, may have modified air pollution through media treatment. Second, as a corollary, we assume that it generally takes longer for the local government to formulate and implement national action plan decisions in the local context. Thus, the action plan announcement may not quickly translate into local decisions that could have influenced pollution through a channel other than the media. Regarding this assumption, we exclude cities and counties targeted by the announcement of the central government’s Three-Year Action Plan mentioned above. In these areas, the action plan could have been known in advance of the announcement by the local authorities given the “target” nature of these areas. Third, we assume that this identification strategy is true only in the short term, which leads us to use regression on discontinuity (RD). In summary, if air pollution is impacted by the announcement of the action plan in the following days, it is only through the media channel.

This strategy allows us to test separately whether the announcement of the action plan

---

<sup>15</sup>For more information, see the official announcement through the release of the notice on the State Council’s official website: [http://www.gov.cn/zhengce/content/2018-07/03/content\\_5303158.htm](http://www.gov.cn/zhengce/content/2018-07/03/content_5303158.htm).

caused discontinuous changes in the emotional tone of media news and in air pollution immediately before and after the day of its announcement.

In practice, we estimate the following equations with non-parametric local linear regression:

$$\begin{aligned} Tone_t &= \alpha_0 + \alpha_1 Announce_t + \alpha_2 Day_t + \iota + \omega + \iota \times \tau + u_{it} \\ \text{s.t. } & -h \leq Day_t \leq h \end{aligned} \quad (2)$$

$$\begin{aligned} E_{ipt} &= \alpha_0 + \alpha_1 Announce_t + \alpha_2 Day_t + \iota + \omega + \iota \times \tau + u_{it} \\ \text{s.t. } & -h \leq Day_t \leq h \end{aligned} \quad (3)$$

Where  $Announce_t$  is an indicator variable of 1 for the period after the announcement of the Three-year Action Plan.  $Day_t$  is the distance of days within  $h$  days around the announcement (the running variable). We also include in the model an interaction term of the  $Announce_t$  and  $Day_t$  to make the model flexible, and finally, we control for a set of fixed effects, including city/county, weekday, and city/county  $\times$  month fixed effects.

A central issue surrounding the non-parametric estimation in RD design is choosing for analysis an appropriate bandwidth ( $-h$  to  $h$ ), i.e., the range of observations around the cutoff point. Selecting a narrow bandwidth has the advantage that the treatment effect is estimated with limited bias because the underlying relationship between the outcome and the running variable tends to be linear. However, by reducing the sample size, a narrow bandwidth may also lessen the precision of impact estimates. Also, a narrow bandwidth makes the extrapolation of estimates more difficult as the estimated treatment effect is a local treatment effect (LATE) that generalizes only to units in the range around the cutoff. The choice of bandwidth, therefore, entails a trade-off between estimation bias and statistical power. To select the optimal bandwidth  $h$ , we rely on the robust CCT selector as proposed by [Calonico et al. \(2014\)](#).

Under the critical assumption that any unobserved factors change smoothly over days around the announcement, the estimation of the two equations will provide unbiased estimates of the impact of the announcement on news media emotion and air pollution. More importantly, we exclude cities and counties in the key areas targeted by the action plan to ensure that air pollution is influenced only by the announcement's impact on news media emotion. As such the two equations can be combined to develop a 2SLS estimate of the impact of news media emotion on air pollution as follows:

$$\begin{aligned} E_{ipt} &= \beta_0 + \beta_1 \widehat{Tone}_t + \beta_2 Day_t + \iota + \omega + \iota \times \tau + \varepsilon_{it} \\ \text{s.t. } & -h \leq Day_t \leq h \end{aligned} \quad (4)$$

Where  $\widehat{Tone}_t$  represents the fitted values from equation (1). In the non-parametric setup, the Local Average Treatment Effect (LATE)  $\beta_1$  can be estimated by taking the ratio of the estimated discontinuity in air pollution to the estimated discontinuity in media news emotion. We use a ‘‘fuzzy RD’’ approach, considering that the probability of treatment does not change from 0 to 1 at the cutoff.

## 5 Results

### 5.1 Effect of news emotion on air pollution

We first propose to investigate the impact of the emotional tone of news media at the county and city levels. To do so, we begin by focusing on a one-week average lag ( $n = 7$ ).

Tables 1 and 2 show estimation results at the county and city levels, respectively. In each table, we report both results for SO<sub>2</sub>, NO<sub>2</sub>, PM<sub>2.5</sub> and PM<sub>10</sub>. In columns 1, 3, 5 and 7, we estimate only the effect of the media variable (tone, polarity and count). In columns 2, 4, 6 and 8, we add weather variables. In all specifications, we control for county or city fixed effects, weekday fixed effects and upper-level ID×months fixed effects.

We find a significant and positive effect of the tone of online news related to air pollution on the emissions of the four pollutants either at the city or county level (except for PM 2.5 at the city level). More precisely, at the city level (county level), a one-percent decrease in the tone reduces air pollutants between 0.028% (0.027%) and 0.045% (0.049%). Put differently, a more negative tone in the news will reduce air pollution in Chinese cities and counties.

Table 1: The impact of tone of environmental news on city air quality levels

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	lnno2	lnno2	lnso2	lnso2	lnpm25	lnpm25	lnpm10	lnpm10
Tone <sub>7</sub>	0.0454*** (0.00659)	0.0416*** (0.00504)	0.0377*** (0.00948)	0.0328*** (0.00718)	0.00920 (0.0149)	0.0215 (0.0147)	0.0234 (0.0156)	0.0282* (0.0150)
Polarity <sub>7</sub>	0.116*** (0.0175)	0.0865*** (0.0138)	0.150*** (0.0140)	0.118*** (0.00951)	0.0395 (0.0452)	0.0539 (0.0397)	0.102*** (0.0355)	0.0955*** (0.0307)
Count <sub>7</sub>	0.0696*** (0.0100)	0.0746*** (0.00877)	-0.00637 (0.0138)	-0.000899 (0.0114)	0.167*** (0.0267)	0.162*** (0.0265)	0.143*** (0.0245)	0.143*** (0.0231)
Htmpt		0.0401*** (0.00204)		0.0520*** (0.00363)		0.0460*** (0.00471)		0.0649*** (0.00359)
Ltmpt		-0.0264*** (0.00243)		-0.0392*** (0.00339)		-0.0144** (0.00523)		-0.0388*** (0.00368)
Wind		-0.0925*** (0.00484)		-0.0500*** (0.00418)		-0.0466*** (0.00883)		-0.0157 (0.0117)
Weather		0.000221 (0.000255)		-0.000730*** (0.000220)		-0.00172*** (0.000463)		-0.00178*** (0.000447)
Constant	2.615*** (0.0730)	2.374*** (0.0609)	2.114*** (0.0960)	1.663*** (0.0978)	2.331*** (0.131)	1.746*** (0.139)	3.036*** (0.138)	2.240*** (0.140)
Observations	98,596	98,596	98,596	98,596	98,596	98,596	98,596	98,596
R-squared	0.606	0.677	0.624	0.670	0.483	0.541	0.470	0.562
Control of weather	N	Y	N	Y	N	Y	N	Y
F-statistic	38.57	138	43.28	53.53	27.56	36.06	25.50	123.4
Number of city	286	286	286	286	286	286	286	286

Note: The dependent variable is air pollution. Tone, polarity and count are lagged (mean over the week before the observation of air pollution). Htmpt (highest temperature), Ltmpt (lowest temperature), wind and weather are meteorological variables (not lagged). Robust standard errors are in parentheses. City fixed effects, weekday fixed effects and province-month fixed effects are controlled for in all columns. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 2: The impact of tone of environmental news on county air quality levels

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	lnno2	lnno2	lnso2	lnso2	lnpm25	lnpm25	lnpm10	lnpm10
Tone <sub>7</sub>	0.0427*** (0.00423)	0.0428*** (0.00394)	0.0200*** (0.00618)	0.0279*** (0.00623)	0.0350*** (0.00640)	0.0463*** (0.00634)	0.0409*** (0.00645)	0.0495*** (0.00664)
Polarity <sub>7</sub>	0.165*** (0.0101)	0.166*** (0.00993)	0.0398*** (0.0135)	0.0704*** (0.0140)	0.0111 (0.0148)	0.0555*** (0.0143)	0.102*** (0.0125)	0.138*** (0.0131)
Count <sub>7</sub>	0.0501*** (0.00438)	0.0658*** (0.00480)	0.0251*** (0.00879)	0.0350*** (0.00946)	-0.0105 (0.00807)	0.00904 (0.00870)	0.0798*** (0.00946)	0.100*** (0.0104)
Htmpt		0.0228*** (0.00165)		0.00652** (0.00282)		0.0168*** (0.00211)		0.0265*** (0.00248)
Ltmpt		-0.00792*** (0.00159)		0.0128*** (0.00350)		0.0162*** (0.00226)		0.00546* (0.00287)
Wind		-0.0310*** (0.00207)		-0.000445 (0.00300)		-0.0112*** (0.00292)		-0.000380 (0.00311)
Weather		-0.000142 (0.000147)		-0.000372* (0.000222)		-0.00139*** (0.000221)		-0.00120*** (0.000204)
Constant	1.780*** (0.0346)	1.409*** (0.0550)	1.753*** (0.0659)	1.410*** (0.0901)	3.381*** (0.0441)	2.779*** (0.0647)	3.045*** (0.0598)	2.323*** (0.0848)
Observations	644,617	644,542	644,617	644,542	644,617	644,542	644,617	644,542
R-squared	0.897	0.902	0.813	0.817	0.613	0.634	0.731	0.748
Control of weather	N	Y	N	Y	N	Y	N	Y
F-statistic	123.5	93.38	12.85	34.07	15.34	160.8	81.10	116.4
Number of county	2408	2,408	2,408	2,408	2,408	2,408	2,408	2,408

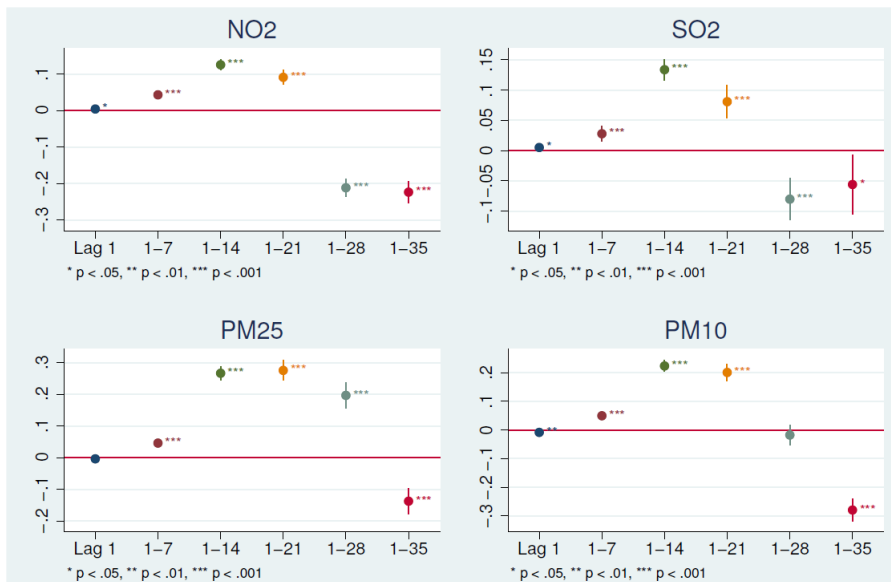
Note: The Dependent variable is air pollution. Tone, polarity and count are lagged (mean over the week before the observation of air pollution). Htmpt (highest temperature), Ltmpt (lowest temperature), wind and weather are meteorological variables (not lagged). Robust standard errors are in parentheses. County fixed effects, weekday fixed effects and city-month fixed effects are controlled for in all columns. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

## 5.2 Previous news emotion and the memory effect

Secondly, at the county and city levels, we introduce different lags on media variables in addition to a one-week average lag in step 1 ( $n = 1, 7, 14, 21, 28, 35$ ). The goal is to investigate if there is a memory effect on the tone of the news. To do so, we replicate the specifications of columns 2, 4, 6 and 8 with different lags of media variables ( $n = 1, 7, 14, 21, 28, 35$ ).

Figures 3 and 4 show the estimation results after plotting the coefficient of the tone variable according to its lag ( $n = 1, 7, 14, 21, 28, 35$ ) at the city level and county level respectively<sup>16</sup>. Also, we plot the estimation results according to the pollutant (upper graphs for NO2 and SO2, lower graphs for PM2.5 and PM10)<sup>17</sup>.

Figure 3: Dynamic effect of news coverage tone at the city level

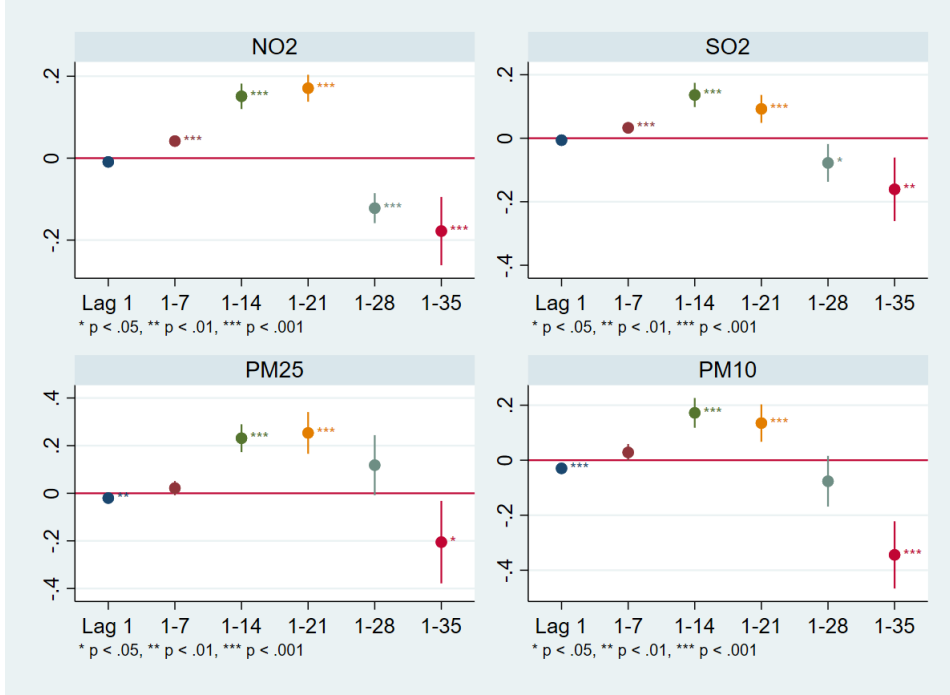


Three results are worth noting. First, the effect of  $tone_{t=-1}$  is not robust. It is often insignificant and is either positive or negative. Second, we find a positive and significant effect of tone until  $n = 21$  (that is to say until three weeks before the day of observed air pollution). We find the strongest effect of tone either when  $n = 14$  or  $n = 21$ . More particularly, at city level, a one-percent increase of tone leads to an increase of 0.17% ( $n = 21$ ), 0.14% ( $n = 14$ ), 0.25% ( $n = 21$ ) and 0.17% ( $n = 21$ ) for NO2, SO2, PM2.5 and PM10 respectively. At the county level, we find that a one-percent increase in tone leads to an increase of 0.13% ( $n = 14$ ), 0.14% ( $n = 14$ ), 0.28% ( $n = 21$ ) and 0.22% ( $n = 21$ ) for NO2, SO2, PM2.5 and PM10 respectively. This result means that the tone of older articles has an impact on polluters' behaviors. However, with regards to the measurement of the lags, this result can also suggest the presence of a media

<sup>16</sup>Complete estimation results are available in the supplementary materials.

<sup>17</sup>For  $n = 7$ , the plotted coefficient is the coefficient of the tone reported in Tables 1 and 2.

Figure 4: Dynamic effect of news coverage tone at the county level



“quantity” effect more than a media “quality” effect. Indeed, while we want to capture the effect of tone as a measure of the “quality” of media information through the emotional connotations of words used, the tone of all articles published during the previous weeks can also capture the fact that there are more articles taken into account. Then the positive impact of tone should just mean that more information on air pollution influences polluters’ behaviors. In the robustness part in Section 6, we propose new control variables to deal with this issue. Third, the effect of the oldest news ( $n = 28$  or  $35$ ) is less robust. The sign of coefficient estimates oscillate between negative and positive, the effect is sometimes significant, sometimes insignificant, the confidence intervals are much larger than for more recent weeks and the results change between the two panels.

### 5.3 Positive and negative emotions

To investigate the differential effect of positive versus negative emotions, we replicate the specifications of Figures 4 and 3 by alternatively replacing the emotional tone with the positive and negative scores, which calculate the percentage of all words in the article with a positive or negative emotional connotation. As explained before, we assume that in the Chinese context, individuals will be more sensitive to negative emotions transmitted by online newspapers.

Figures A1 and A3 in the Appendix show the estimation results for the four pollutants after plotting the coefficient of the positive score according to its lag ( $n = 1, 7, 14, 21, 28, 35$ ) while



Figures A2 and A4 in the Appendix report the same estimation results for the negative score<sup>18</sup>.

Two main results are worth noting. Firstly, we do find the same pattern than for the tone. For  $n = 1, 28$  and  $35$ , the effect of both the positive and negative scores is not robust. For  $n = 7, 14$  and  $21$ , we find that the two scores impact air pollution in the expected ways. An increase of the positive score significantly pushes up air pollution both at the city level (Figure A1) and county level (Figure A2), while an increase of the negative score significantly reduces air pollution both at the city level (Figure A3) and county level (Figure A4). Thus, these results are in the same direction as those found for the emotional tone but suggest that the two parts (positive and negative) of the tone act on air pollution. Individuals react both to the positive and negative emotions transmitted by online newspaper articles.

Secondly, while the two types of emotions have an influence on air pollution, the magnitude of their effects is not the same. In all regressions, negative emotions lead to a higher variation of air pollution than positive emotions. The effect of the positive score ranges between 0.19 (for NO<sub>2</sub> and SO<sub>2</sub> with  $n = 14$ ) and 0.36 (PM<sub>2.5</sub> with  $n = 14$ ) at the city level and between 0.15 (for NO<sub>2</sub> with  $n = 14$ ) and 0.41 (PM<sub>2.5</sub> with  $n = 14$ ) at the county level. In contrast, the effect of the negative score fluctuates between -0.42 (for SO<sub>2</sub> with  $n = 14$ ) and -0.74 (PM<sub>2.5</sub> with  $n = 14$ ) at the city level and between -0.36 (for NO<sub>2</sub> with  $n = 14$ ) and -0.86 (PM<sub>2.5</sub> with  $n = 14$ ) at the county level. As explained, we attribute these results to the specific Chinese context made up of both stringent environmental policy and a strong awareness of air pollution by the population.

## 5.4 Identification of a causal effect: the RDD strategy

### 5.4.1 The graphical evidence

We begin to present the graphical evidence from the fuzzy RD design to provide a straightforward visualization of the impact of news media tone on air pollution. We focus on the observations in a small window around the cutoff point. More precisely, we adopt an arbitrary bandwidth  $h$  of 50 days before and after June 27th, 2018. The optimal bandwidths in the following econometric analysis are narrower but we keep a bit larger bandwidth here to improve the visualization<sup>19</sup>. Moreover, we exclude cities and counties targeted by the announcement of the central government’s Three-Year Action Plan. These areas are Beijing-Tianjin-Hebei and surrounding areas, the Yangtze River Delta region (including Shanghai and Jiangsu) and the Fenwei plain with Shanxi and Shaanxi provinces as well as parts of Henan province.

We implement several data-driven regression-discontinuity plots with binned sample means. More precisely, we divide the window into evenly spaced “bins” of air pollution or tone and plot the mean value of each “bin” against the running variable (the distance of days within 50 days

---

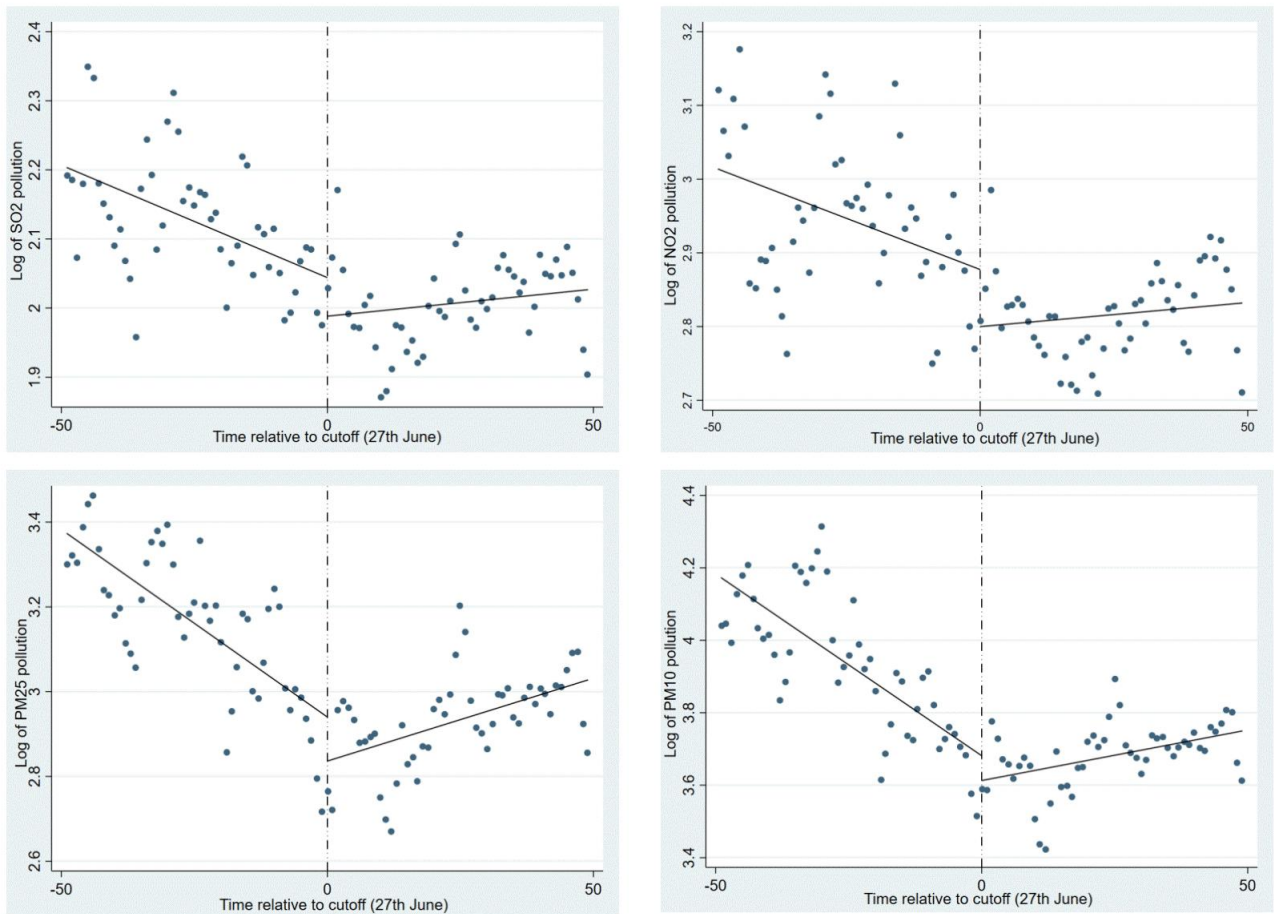
<sup>18</sup>Complete estimation results are available in the supplementary materials.

<sup>19</sup>The underlying discontinuity is still found even if we reduce the bandwidth. Results are available upon request.

of the date of the announcement). The fitted line is constructed using a flexible local linear regression, enabling us to obtain a clear picture of the data distribution.

Figure 5 shows the pattern over time in the four air pollutants. We can see a “jump” (discontinuity) for the four pollutants at the cutoff point that indicates discontinuous changes in air pollution immediately before and after the central government’s announcement.

Figure 5: The discontinuity of the outcomes (air pollution) at the cutoff (city-level)



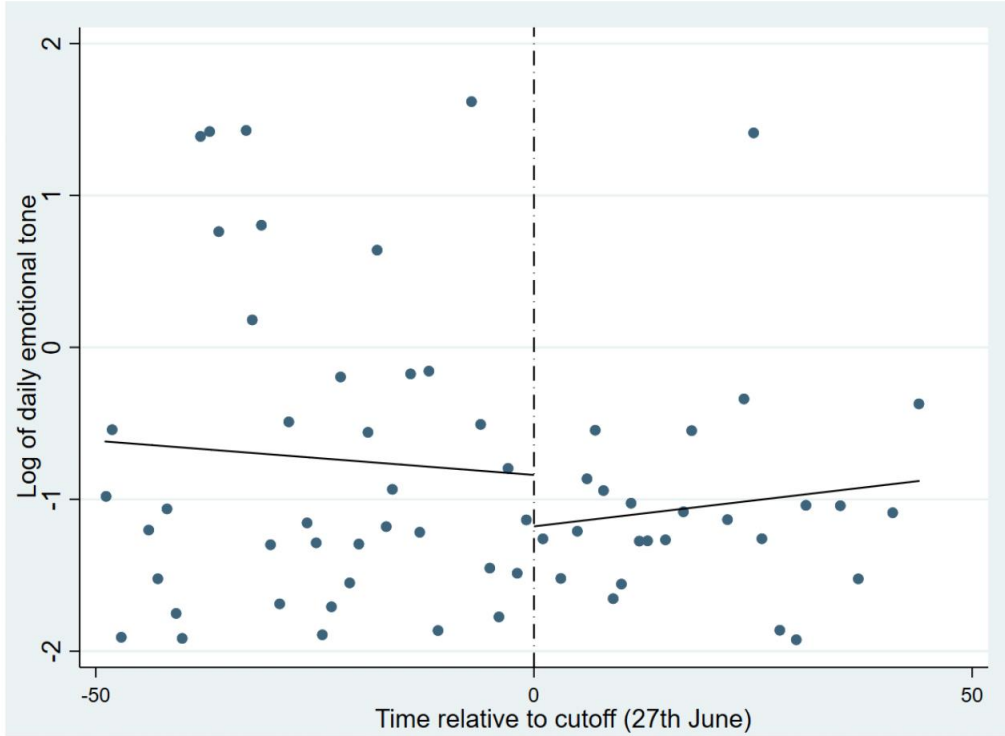
Note: RD plots with binned sample means mimicking the underlying variability of the data. The vertical dashed line designates the cutoff point (June 27) and the fitted lines are constructed with local linear regression method.

In the same vein, Figure 6 presents the pattern of emotional tone over time. We can also see a discontinuity at the date of the government announcement. This “jump” may also indicate discontinuous changes in the emotional tone of the news media immediately before and after the announcement.

These RD graphs clarify the internal validity of the RD analysis but the magnitude and statistical significance of the impact of tone on air pollution must be quantified. The following regression analysis within a small subset around the cutoff will allow us to estimate the LATE

impact of tone using non-parametric estimation.

Figure 6: The discontinuity of news media tone at the cutoff (city-level)



Note: RD plots with binned sample means mimicking the underlying variability of the data. The vertical dashed line designates the cutoff point (June 27) and the fitted lines are constructed with local linear regression method.

#### 5.4.2 The impact of news media tone on air pollution

Table 3 displays the same LATE impact of news media tone on air quality after controlling for several fixed effects (FE): city FE in panel A, county FE in panel B, weekday FE (panels A and B), province $\times$ month FE (panel A) and city $\times$ month (panel B). The LATE impact is estimated by using the cutoff created by the central government’s announcement on June 27th, 2018 regarding the triennial strategy in terms of the war on pollution. The optimal bandwidth is selected using the approach of [Calonico et al. \(2014\)](#), which relies on the coverage error rate (CER). More precisely, we use a different bandwidth before and after the cutoff (two CER) in columns 1 to 4 and a common bandwidth before and after the cutoff (common CER) in columns 5 to 8. The estimation sample excludes cities and counties targeted by the Three-year Action Plan (i.e., Beijing-Tianjin-Hebei and surrounding areas, the Yangtze River Delta region (including Shanghai and Jiangsu) and the Fenwei plain with Shanxi and Shaanxi provinces as well as parts of Henan province). Except for SO<sub>2</sub> at the county level, we still find a positive

impact of the media tone on air pollution with the two criteria used to select the optimal bandwidth. Regarding the magnitude of the coefficients, they range from 0.38 (for PM10 and common CER criteria) to 9.19 (for PM2.5 and two CER criteria) at the city level, and 0.16 (for PM10 and common CER criteria) to 1.16 (for PM2.5 and two CER criteria) at the county level. Overall, the magnitude of the estimate results is much stronger than in the baseline estimates reported in Figures 4 and 3. We can attribute this result to the RD strategy. Indeed, we report a LATE impact around a cutoff where we assume to find a strong effect of tone on air pollution given the Chinese context. In contrast, in the baseline estimates, we report the average effect of tone overall in 2018.

Table 3: Non-parametric RD estimates of the (LATE) impact of the tone of news media on air quality at the city and county levels – with fixed effects

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidth selection	lnno2	lnso2	lnpm25	lnpm10	lnno2	lnso2	lnpm25	lnpm10
	Two CER			Common CER				
	Panel A - City level							
Ln(tone <sub>t-1</sub> )	0.453*** (0.127)	3.364*** (1.302)	9.185*** (1.925)	0.538*** (0.0799)	0.672** (0.279)	0.513** (0.200)	0.466*** (0.107)	0.381*** (0.0670)
Observations	44,532	44,532	44,532	44,532	44,532	44,532	44,532	44,532
Nb of cities	206	206	206	206	206	206	206	206
Cutoff - c	June 27	June 27	June 27	June 27	June 27	June 27	June 27	June 27
Bandwidth below c	12.65	15.44	15.68	11.68	13.63	12.69	11.49	11.12
Bandwidth above c	20.56	15.10	21.20	13.37	13.63	12.69	11.49	11.12
	Panel B - County level							
Ln(tone <sub>t-1</sub> )	0.403*** (0.0503)	-0.00892 (0.0249)	1.105*** (0.284)	1.709*** (0.149)	0.429*** (0.150)	-0.00847 (0.0278)	0.0357 (0.0242)	0.160*** (0.0228)
Observations	275,598	275,598	275,598	275,598	275,598	275,598	275,598	275,598
Nb of counties	2,408	2,408	2,408	2,408	2,408	2,408	2,408	2,408
Cutoff - c	June 27	June 27	June 27	June 27	June 27	June 27	June 27	June 27
Bandwidth below c	6.012	8.550	4.325	3.964	14.41	8.240	8.910	9.975
Bandwidth above c	18.71	6.009	21.63	22.89	14.41	8.240	8.910	9.975

Note: The estimation sample excludes cities and counties targeted by the government's June 27 announcement, whose date is used as the cutoff. RD estimates with non-parametric local linear regressions. Conventional local linear regression robust standard errors are clustered at the city (panel A)/county (panel B) levels are reported in parentheses. CER = coverage error-rate. The optimal bandwidth is selected using the approach of [Calonico et al. \(2014\)](#). Fixed effects (FE) controlled for are: ID FE (city in panel A and county in panel B), weekday FE (both panel), province\*month FE (panel A) and city\*month (panel B). \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

To sum up, we attribute the positive effect of news media tone to the context of China's so-called war on air pollution. The government makes use of the environmental news emotion as an informal regulator of pollution. The announcement of the Three-year Action Plan is a set of measures related to command-and-control measures to improve the green development of industry, transportation and the energy system. Also, the plan settles on stringent thresholds in terms of air pollution reduction. On the whole, the emotional tone disseminated in this plan and transmitted by newspapers is negative. As a consequence, the positive effect suggests that a reduction of the tone (i.e., the words used are more negatively connoted in terms of emotions) strengthens air pollution. Put differently, the war on air pollution is relayed by mass media

through the emotional tone, which influences air pollution. Lastly, while we attribute the effect of news media tone on air pollution to the announcement of the action plan, remember that the effect is only a local average treatment effect (LATE). In other words, this is the effect of the tone around the date of the announcement. While the disadvantage is to estimate only a LATE of tone, the advantage is to estimate a causal impact thanks to the context of the government’s war on pollution.

## 6 Robustness tests

This section demonstrates that our main baseline estimate is robust to an alternative model with different control variables, a new measure of lagged variables and a new cohort of articles (RD placebo test).

### 6.1 Adding new control variables

We first examine the sensitivity of our baseline estimate with more control variables to reduce the potential omitted variable bias. We redo the baseline estimates (for all pollutants and at both city and county levels) by adding several new variables<sup>20</sup>.

First, we add two new control media variables from the GDELT database. We introduce the number of articles on air pollution published per day to capture a “quantity” effect from media. As explained before, we want to estimate the impact of tone as a new measure of information access to people. Tone by measuring the emotional connotation of words used in the article can be defined as a measure of the quality of the media treatment. However, given the calculation of the lags, we may attribute the effect of tone to more articles taken into account. To deal with this issue, we add the number of articles published daily. We also add the share of articles published by official media (ex. *China Daily* and *China News Service*) per day to control the relationship between media and policy in China. At mean, there are 50% of articles published by official media in our database. Note that we use the lag of these two variables. The lagging is made over the same period as the other media variables including tone depending on the specification chosen (i.e.,  $n = 1$  to 35). Second, we add lagged weather variables. More precisely, the four weather variables of the baseline model are lagged over the same period as media variables depending on the specification chosen (i.e.,  $n = 1$  to 35). Third, we also introduce a trend and its squared term to capture the U-shaped trend in air pollution over a calendar year. Finally, we add more fixed effects (FE) by introducing  $ID \times month$  FE and  $ID \times weekday$  FE (where ID is either city or county).

In Table A2, we report the estimate results for six panels. Panels A.1, A.2 and A.3 concern the city level while panels B.1, B.2 and B.3 focus on the county level. In panels A.1 and B.1, we report the coefficients of tone while panels A.2 and B.2 (A.3 and B.3) present the coefficients of

---

<sup>20</sup>The descriptive statistics of these variables are in Table A1 in the Appendix.

positive score (negative score). To save space, we focus on three time periods ( $n = 7, 14, 21$ ) and we do not report the coefficients of control variables<sup>21</sup>. Both the magnitude and significance of the three variables are close to the baseline estimates. Tone and positive score have a positive effect on all pollutants while negative score has the expected negative effect.

## 6.2 Adding daily financial variables

While the previous model with more control variables helps to reduce the potential omitted variable bias, it remains a concern. Daily media variables, and more particularly the tone variable, can capture the economic environment. When the economic conditions are improving, the emotional tone can also increase. Journalists may use words with positive connotations given the positive economic context. We use the model with all control variables used in the previous subsection and add four daily financial variables to control for economic conditions. These variables are the SSE Composite Index (a stock market index of all stocks that are traded at the Shanghai Stock Exchange), the energy price index, the exchange rate USD/RMB and the Shanghai Interbank Offered Rate (SHIBOR)<sup>22</sup>.

Tables A3 and A4 in the Appendix present the results at the city level and county level respectively. In each Table, we report the results for tone (panel A.1), positive score (panel A.2) and negative score (panel A.3). To save space, we focus on the same three time periods ( $n = 7, 14, 21$ ) than before and we only report the coefficients of tone, positive score, negative score and the four financial variables<sup>23</sup>. While the magnitude of the coefficients of tone, positive score and negative score is lower, the direction of their effects on air pollution remains the same as for the baseline estimates.

## 6.3 Changing the calculation of lags

We second investigate the sensitivity of our baseline estimate with a different definition of the lags for both media and weather variables. We now define the lags by separating weeks. More precisely, we still use lags from  $t = 1$  to  $t = 35$  but we now add separately different one-week averages. Then,  $n = 7$  is the one-week average between  $t = -1$  and  $t = -7$ ,  $n = 14$  is the one-week average between  $t = -8$  and  $t = -14$ ,  $n = 21$  is the one-week average between  $t = -15$  and  $t = -21$ ,  $n = 28$  is the one-week average between  $t = -22$  and  $t = -28$  and  $n = 35$  is the one-week average between  $t = -29$  and  $t = -35$ . With this new specification, we aim at testing if there is a differential effect between newer and older news.

Table A5 in the Appendix shows the estimation results of the effect of the new lagged tone variable respectively at the city level (panel A) and county level (panel B) for the four pollutants. In all specifications, we add the new controls defined in the subsection 6.1. To save

<sup>21</sup>Complete estimation results are available upon request.

<sup>22</sup>The data is derived from the database of Wind (<https://www.wind.com.cn/>), which is a significant provider of financial data in China.

<sup>23</sup>Complete estimation results are available upon request.



space, we focus only on the coefficient estimates of tone and we do not report the coefficients of control variables<sup>24</sup>. Tables A6 and A7 in the Appendix report the same estimation results for positive score and negative score respectively.

Two results are worth noting. First, these results confirm the baseline estimate. Tone and positive score have a significant and positive effect while a negative score impacts negatively air pollution. Second, these effects are robust for  $t = -1$  to  $-7$  and  $t = -8$  to  $-14$ . For older weeks (i.e.  $t = -15$  to  $-21$ ,  $t = -22$  to  $-28$  and  $t = -29$  to  $-35$ ), the coefficient estimates are less robust. The sign of coefficient estimates oscillate between negative and positive, the effect is sometimes significant, sometimes insignificant, and the results change between the two panels. Taken together, these results confirm the baseline estimate but also suggest that news published in the two previous weeks influence more polluters' behavior than older news.

## 6.4 RDD falsification test

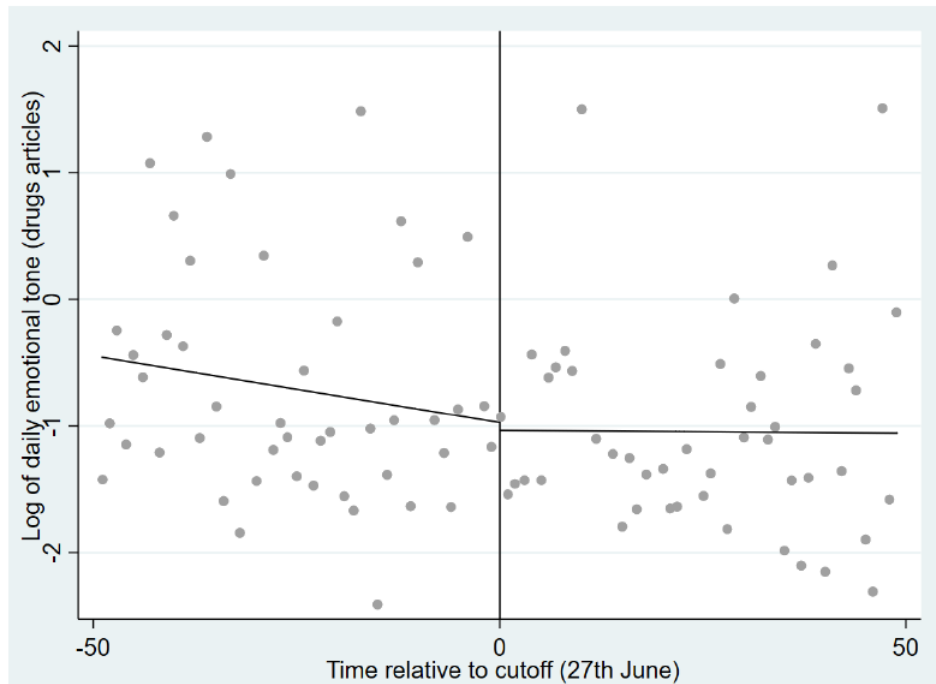
Our RD strategy above suggests a statistically significant effect of the tone on air pollution. However, it is important to test further whether we have found what seems likely to be the impact of the tone of articles on air pollution or just an artifact of the data. Regarding the design of our RD strategy and the use, as a cutoff, of the date of the central government's announcement of a new air pollution policy, only the tone of articles on air pollution should be effectively the driving force behind our results. Then we would expect to find no discontinuity with a different cohort of articles. In order to test this assumption, we thus use a different cohort of 1,500 articles related to drugs published in 2018, and we replicate the previous RD strategy.

First, we begin with the graphical evidence from the fuzzy RD design to provide a straightforward visualization of the absence of the impact of news media tone related to drugs on air pollution. Figure 7 presents the pattern of the emotional tone of articles on drugs over time. We see no discontinuity at the date of the government announcement. This absence of "jump" indicates no discontinuous changes in the emotional tone before and after the announcement. Moreover, we estimate the LATE impact of news media tone related to drugs on air quality (SO<sub>2</sub>, NO<sub>2</sub>, PM<sub>2.5</sub> and PM<sub>10</sub>) at the city level (panel A) and county level (panel B) with fixed effects in Table A8. We find no significant estimates of the tone of drug articles on air pollution suggesting that the positive tone effects that are observed in Table 3 occur only with the cohorts of articles on air pollution.

---

<sup>24</sup>Complete estimation results are available upon request.

Figure 7: The discontinuity of news media tone related to drugs at the cutoff (city-level)



Note: RD plots with binned sample means mimicking the underlying variability of the data. The vertical dashed line designates the cutoff point (June 27) and the fitted lines are constructed with local linear regression method.

## 7 Potential mechanisms

### 7.1 Adjustment of corporate pollutant emissions

To go further into the analysis of potential mechanisms, as a first step, we use data on corporate air pollutant emissions to identify the presence of the two first mechanisms. We thus aim at identifying firms as agents that adjust their pollutant emissions to environmental news emotions either by responsible behavior to meet social expectations or by anticipation of controls and sanctions.

To do so, we use data at the firm level from China’s continuous emission monitoring systems (CEMS) network (Tang et al., 2020a)<sup>25</sup>. We restrict our sample to the year 2018 in order to synchronize with previous estimations at the city and county levels. In addition to daily media and weather variables, we control for daily variations of economic activity with the Shanghai Stock Exchange (SSE) index, an energy price index, the exchange rate USD/RMB and the

<sup>25</sup>The real-time CEMS-monitored data was collected by the Institute of Public and Environmental Affairs (IPE) (<http://www.ipe.org.cn/>) and we aggregate the monitor level data to the firm level. CEMS is a system of sampling, conditioning, and analytical components and software that is designed to provide direct, real-time, continuous measurements of pollutant concentration by monitor. The descriptive statistics of these variables are in Table A1 in the Appendix.



Shibor index. All these indexes are both not lagged and lagged over the same time span as the lagged media variables. As in regressions at county and city levels, we use multiple fixed effects (*FE*) including firm FE, weekday FE, city-level×weekday FE and city-level×month FE that enable us to control for many unobservable characteristics at different levels (we also do clustering at firm and city level).

In Tables A9 and A10 in the Appendix, we observe a positive sign for the coefficient of emotional tone, particularly during the first week. This result confirms our previous results and suggests that firms are one type of actors adjusting their pollutant emissions (NOx and SO2). When emotions become more negative, some firms can be incited to reduce their production and thus their pollutant emissions, either by responsible behavior to meet social expectations (the second mechanism in the previous subsection) or by anticipation of controls and sanctions (the first mechanism in the previous subsection). In the next part, we further analyze the “sanctions” mechanism.

## 7.2 Anticipation of controls and sanctions

In the second step, we try to reveal the existence of the policy mechanism regarding the anticipation of controls and sanctions by firms. For this purpose, we use data on standard violations and sanctions at the city level concerning environmental cases. This data is derived from the dataset of law “Beida Faba” managed by the Peking University that collected comprehensive judgment documents across mainland China since 1949. We split the sample between cities with a number of sanctions higher than the median (sample 1), and those with a number of sanctions lower than the median (sample 2)<sup>26</sup>.

As a result, in Tables A11 and A12 in the Appendix, we find that there is a better significance of the tone coefficient in the sample with more sanctions (particularly for pm10 and pm25, and a more persistent effect along time for NO2 and SO2) and its magnitude is higher. We can explain this difference by firms’ behavior. Firms in cities with more sanctions in case of pollution standard violations are more sensitive to the emotional tone in media articles (potentially in anticipation of future sanctions). This reaction to media articles can be interpreted as a complement to cities’ environmental regulations. Media may act as an informal regulator of firms in cities that are particularly prone to sanctions. In other words, media are one factor that increases the compliance of firms with environmental regulation. This result completes the literature on compliance with environmental regulation (for example, Gangadharan, 2006; Kathuria, 2007; Duflo et al., 2013; Barwick et al., 2020).

---

<sup>26</sup>Sample 1 has a mean of 380.1, a maximum of 3,677 and a minimum of 70 (number of sanctions for a year at city level). Sample 2 has a mean of 25, a maximum of 67 and a minimum of 1.

## 8 Policy implications

### 8.1 The importance of adopting effective public policies based on a real-time pollution monitoring system and an efficient sanctioning system

Despite numerous international summits on environmental preservation, such as the recent COP26, public policies to combat environmental damages vary greatly from one country to another. The Chinese air pollution management example, even though largely conditioned by China's specific mode of governance based on a complete centralization of decision-making, may provide inspiration for other countries when dealing with urgent environmental issues such as climate change or air pollution. Indeed a system of daily monitoring of air pollution on a real-time basis, combined with a system of efficient sanctions in case of non-compliance, is likely to be one of the most effective strategies for achieving the objectives set out in the latest IPCC report of 4 April 2022. The report challenges the "apathy of policymakers" regarding the climate emergency, which therefore contrasts with the present case study on Chinese air pollution management. As well, the literature continues to demonstrate the need for the development of more robust public policies against air pollution in the wake of disease and in high-growth societies with excessive pollution levels (Jin et al., 2020). After all, if public policies in other countries are insufficiently ambitious and far-sighted in addressing pollution issues, they may experience the same negative health and socioeconomic impacts that China experienced prior to its 2013 regulations.

### 8.2 The use of media as a powerful instrument for informal regulation of air pollution issues

Whereas the media's role as a mean for raising people's awareness and fight against air pollution is well established in the literature (see the Introduction and the Background sections), there seems to be limited information about the most effective and concrete resources to be mobilized to this end. Our contribution to the use of the emotional tone by the media to achieve this goal may also serve as a source of inspiration beyond the Chinese context. It may be considered as an informal regulation instrument that is relatively effective in the fight against air pollution. Of course, the Chinese context where the media is mostly controlled by the government favors this continuity of public policy/informal media regulation. A pure and simple transplant of this model of governance in other countries committed to the independence of the media is not, however, conceivable.

Nevertheless, other countries' public policies would be depriving themselves of an efficient means to control pollution levels and to promote environmentally responsible behavior on the part of socioeconomic agents. This is why, in order to respond to the many challenges of the

climate emergency, a massive media mobilization, together with the legitimizing of a large use of emotional tones similar to the Chinese case, seems to be necessary in other national contexts.

Finally, despite these positive inspirations in the development of effective public policies for the management of pollution, several points of vigilance deserve to be raised. These concerns are related to the potentially divergent interests of the stakeholders. They need to be considered beforehand in order to avoid counterproductive effects in the fight against pollution.

### 8.3 Some areas of concern requiring future research

These points of vigilance concern, first of all, companies that obey different rationalities from those of public agents or households. These rationalities are bound up with performance targets. The literature informs us that companies that receive bad press due to their polluting behavior have a tendency to only invest in green acquisitions (clean production technologies and emission reduction technologies) but do not actually commit to the green transformation process (ranging from green investments to production process transformation (Wang and Zhang, 2021)). Yet, the green transformation process is the only way to guarantee a real commitment on the part of the company to become sustainable in a permanent way. In reality, to avoid such an expensive process in terms of time, resources and energy (which would negatively impact their performance), businesses are tempted to prefer green acquisitions, which will immediately draw media attention, which will express positive opinions about the efforts of these companies. The green transformation process, on the other hand, is less eye-catching because it is a long-term process that takes several years to deliver visible results. Companies will therefore try their best to seem pro-environment and get positive media coverage, which in turn may increase the chances for green-washing practices. In the end, this may prove counterproductive in the fight against pollution. So as to reduce such risks, it appears essential to concretely formalize the whole process of green acquisitions and green transformations that companies must undergo in the short, medium and long term within the perspective of public policies against pollution. In addition, media information should be as precise as possible about the progress that has been made by companies so far as well as on the progress that needs to be made in order for them to succeed in the entire green transition process.

## 9 Conclusion

The goal of this paper was to investigate the role of environmental online news as an informal regulator of air pollution in China. Using daily data on environmental news (GDELT database) and on air pollution monitoring at the county and city levels, we find consistent results regarding the impact of the emotional tone: more negative (positive) emotions in environmental news can decrease (increase) air pollution until three weeks before the day of observed air pollution. This result is robust to several air pollutants (e.g., NO<sub>2</sub>, SO<sub>2</sub>, PM 2.5 and PM<sub>10</sub>) both at the

city and county levels, and also to the addition of other media variables, economic/financial variables, time trends and a different measure of lagged variables.

We attribute this effect to the context of China’s “war on air pollution” in which the mass media act as an informal regulator of pollution. To confirm this hypothesis, we use an RDD to identify a causal and local impact of news emotions on air pollution levels. Specifically, we exploit the official announcement on June 27, 2018, of China’s “Three-year Action Plan” as a cut-off to separately test whether there were discontinuous changes in the emotional tone of news media articles and air pollution levels immediately before and after the announcement. The RDD strategy confirms that more negative (positive) emotions in environmental news can decrease (increase) air pollution; this result is robust to a placebo test.

Moreover, we identify potential mechanisms by exploiting two databases on corporate pollutant emissions and the number of sanctions after violation of air pollution standards by firms at the city level. We find that firms are more prone to reduce their pollutant emissions after negative media emotions and that air pollution is more influenced by the emotional tone in cities where the number of sanctions is important. This effect can be interpreted as a complement to cities’ environmental regulations. Media seem to act as an informal regulator on firms in cities that are particularly prone to sanctions, improving the compliance of firms with environmental regulations.

Finally, it is not clear enough whether socioeconomic agents undergo through deep and lasting changes in their environmental protection beliefs and responsible behavior, or if these changes are only short-term in order to escape sanctions or performance damage. More focused actor-specific research is therefore needed to better understand the effect of the media’s emotional tones on behavioral changes.

## Acknowledgements

We thank Camille Meneur and Alexis Repon for their excellent research assistance. We also thank Julien Ah-Pine, Simone Bertoli, Olexiy Kyrychenko, Sylvain Marsat, Anna Risch, Lu Xu, and Peng Zhang as well as seminar and conference participants at CEE-M (Montpellier University), CERDI (University Clermont Auvergne - CNRS), Clerma (ESC Clermont), BSE (Bordeaux University - CNRS), Nanjing Audit University, 8th FAERE Annual Conference in Grenoble, and Asian Economic Development Conference 2022 (Asian Development Bank). This work was supported by the Agence Nationale de la Recherche of the French government through the program “Investissements d’avenir” ANR-10-LABX-14-01. All remaining errors are ours.

## References

- Andrews, S. Q. (2008). Inconsistencies in air quality metrics: ‘blue sky’ days and pm10 concentrations in beijing. *Environmental Research Letters*, 3(3):034009.
- Barwick, P. J., Li, S., Lin, L., and Zou, E. (2020). From fog to smog: The value of pollution information. Technical report, National Bureau of Economic Research.
- Brady, A.-M. (2009). *Marketing dictatorship: Propaganda and thought work in contemporary China*. Rowman & Littlefield Publishers.
- Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.
- Chen, S., Guo, C., and Huang, X. (2018). Air pollution, student health, and school absences: Evidence from china. *Journal of Environmental Economics and Management*, 92:465–497.
- Chen, Z., Jin, M., and Chen, F. (2022). Air pollutions and loan decision bias. *Global Environmental Change*, 72:102420.
- D’Amato, A., Giaccherini, M., and Zoli, M. (2019). The role of information sources and providers in shaping green behaviors. evidence from europe. *Ecological economics*, 164:106292.
- Duffo, E., Greenstone, M., Pande, R., and Ryan, N. (2013). Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from india. *The Quarterly Journal of Economics*, 128(4):1499–1545.
- Feng, Y., Ning, M., Lei, Y., Sun, Y., Liu, W., and Wang, J. (2019). Defending blue sky in china: Effectiveness of the “air pollution prevention and control action plan” on air quality improvements from 2013 to 2017. *Journal of Environmental Management*, 252:109603.
- Gangadharan, L. (2006). Environmental compliance by firms in the manufacturing sector in mexico. *Ecological Economics*, 59(4):477–486.
- Geng, G., Xiao, Q., Liu, S., Liu, X., Cheng, J., Zheng, Y., Xue, T., Tong, D., Zheng, B., Peng, Y., et al. (2021). Tracking air pollution in china: Near real-time pm2.5 retrievals from multisource data fusion. *Environmental Science & Technology*, 55(17):12106–12115.
- Ghanem, D. and Zhang, J. (2014). ‘effortless perfection:’ do chinese cities manipulate air pollution data? *Journal of Environmental Economics and Management*, 68(2):203–225.
- Greenstone, M., He, G., Jia, R., and Liu, T. (2020). Can technology solve the principal-agent problem? evidence from china’s war on air pollution. Technical report, National Bureau of Economic Research.

- Greenstone, M., He, G., Li, S., and Zou, E. Y. (2021). China’s war on pollution: Evidence from the first 5 years. *Review of Environmental Economics and Policy*, 15(2):281–299.
- Ha, J., Akhtar, R., Masud, M. M., Rana, M. S., and Banna, H. (2019). The role of mass media in communicating climate science: An empirical evidence. *Journal of Cleaner Production*, 238:117934.
- Happer, C. and Philo, G. (2016). New approaches to understanding the role of the news media in the formation of public attitudes and behaviours on climate change. *European Journal of Communication*, 31(2):136–151.
- He, X., Luo, Z., and Zhang, J. (2022). The impact of air pollution on movie theater admissions. *Journal of Environmental Economics and Management*, 112:102626.
- Heyes, A. and Zhu, M. (2019). Air pollution as a cause of sleeplessness: Social media evidence from a panel of chinese cities. *Journal of Environmental Economics and Management*, 98:102247.
- Huang, J., Pan, X., Guo, X., and Li, G. (2018). Health impact of china’s air pollution prevention and control action plan: an analysis of national air quality monitoring and mortality data. *The Lancet Planetary Health*, 2(7):e313–e323.
- Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, 142(2):615–635.
- Ito, K. and Zhang, S. (2020). Willingness to pay for clean air: Evidence from air purifier markets in china. *Journal of Political Economy*, 128(5):1627–1672.
- Jiang, X., Li, G., and Fu, W. (2021). Government environmental governance, structural adjustment and air quality: A quasi-natural experiment based on the three-year action plan to win the blue sky defense war. *Journal of Environmental Management*, 277:111470.
- Jin, Y., Andersson, H., and Zhang, S. (2020). Do preferences to reduce health risks related to air pollution depend on illness type? evidence from a choice experiment in beijing, china. *Journal of Environmental Economics and Management*, 103:102355.
- Kathuria, V. (2007). Informal regulation of pollution in a developing country: Evidence from india. *Ecological Economics*, 63(2-3):403–417.
- OECD (2008). *OECD Investment Policy Reviews: China 2008 : Encouraging Responsible Business Conduct*. OECD Investment Policy Reviews, Éditions OCDE, Paris.
- Shapiro, A. H., Sudhof, M., and Wilson, D. J. (2020). Measuring news sentiment. *Journal of Econometrics*.

- Stockmann, D. (2013). *Media commercialization and authoritarian rule in China*. Cambridge University Press.
- Stockmann, D. and Gallagher, M. E. (2011). Remote control: How the media sustain authoritarian rule in china. *Comparative Political Studies*, 44(4):436–467.
- Sunstein, C. R. and Zeckhauser, R. (2011). Overreaction to fearsome risks. *Environmental and Resource Economics*, 48(3):435–449.
- Tang, L., Xue, X., Qu, J., Mi, Z., Bo, X., Chang, X., Wang, S., Li, S., Cui, W., and Dong, G. (2020a). Air pollution emissions from chinese power plants based on the continuous emission monitoring systems network. *Scientific Data*, 7(1):1–10.
- Tang, X., Kong, L., Zhu, J., Wang, Z., Li, J., Wu, H., Wu, Q., Chen, H., Zhu, L., Wang, W., et al. (2020b). A six-year long high-resolution air quality reanalysis dataset over china from 2013 to 2018 (monthly and annual version), v1, sci. data bank.
- Tu, M., Zhang, B., Xu, J., and Lu, F. (2020). Mass media, information and demand for environmental quality: Evidence from the “under the dome”. *Journal of Development Economics*, 143:102402.
- Wang, J. and Jia, Y. (2021). Social media’s influence on air quality improvement: Evidence from china. *Journal of Cleaner Production*, 298:126769.
- Wang, K. and Zhang, X. (2021). The effect of media coverage on disciplining firms’ pollution behaviors: Evidence from chinese heavy polluting listed companies. *Journal of Cleaner Production*, 280:123035.
- Wang, Y., Li, N., and Li, J. (2014). Media coverage and government policy of nuclear power in the people’s republic of china. *Progress in Nuclear Energy*, 77:214–223.
- WHO (2021). *WHO global air quality guidelines: particulate matter (PM<sub>2.5</sub> and PM<sub>10</sub>), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization.
- Wu, M. and Cao, X. (2021). Greening the career incentive structure for local officials in china: Does less pollution increase the chances of promotion for chinese local leaders? *Journal of Environmental Economics and Management*, 107:102440.
- Xue, S., Zhang, B., and Zhao, X. (2021). Brain drain: The impact of air pollution on firm performance. *Journal of Environmental Economics and Management*, 110:102546.
- Yang, C. (2020). Policies, regulatory framework and enforcement for air quality management: The case of china. *OECD Environment Working Papers*, No. 157.

- Yin, P., Brauer, M., Cohen, A., Burnett, R. T., Liu, J., Liu, Y., Liang, R., Wang, W., Qi, J., Wang, L., et al. (2017). Long-term fine particulate matter exposure and nonaccidental and cause-specific mortality in a large national cohort of chinese men. *Environmental health perspectives*, 125(11):117002.
- Yin, P., Brauer, M., Cohen, A. J., Wang, H., Li, J., Burnett, R. T., Stanaway, J. D., Causey, K., Larson, S., Godwin, W., et al. (2020). The effect of air pollution on deaths, disease burden, and life expectancy across china and its provinces, 1990–2017: an analysis for the global burden of disease study 2017. *The Lancet Planetary Health*, 4(9):e386–e398.
- Yu, K., Qiu, G., Chan, K.-H., Lam, K.-B. H., Kurmi, O. P., Bennett, D. A., Yu, C., Pan, A., Lv, J., Guo, Y., et al. (2018). Association of solid fuel use with risk of cardiovascular and all-cause mortality in rural china. *Jama*, 319(13):1351–1361.
- Zhan, J. (2020). Moral problems and ethical issues in china’s media. In *The Routledge Handbook of Mass Media Ethics*, pages 209–220. Routledge.
- Zhang, Q., Zheng, Y., Tong, D., Shao, M., Wang, S., Zhang, Y., Xu, X., Wang, J., He, H., Liu, W., et al. (2019). Drivers of improved pm<sub>2.5</sub> air quality in china from 2013 to 2017. *Proceedings of the National Academy of Sciences*, 116(49):24463–24469.
- Zhang, S., Zhang, M.-a., Qiao, Y., Li, X., and Li, S. (2022). Does improvement of environmental information transparency boost firms’ green innovation? evidence from the air quality monitoring and disclosure program in china. *Journal of Cleaner Production*, page 131921.
- Zheng, X., Lang, A., and Ewoldsen, D. R. (2021). The measurement of positive and negative affect in media research. In *Routledge International Handbook of Emotions and Media*, pages 48–66. Routledge.
- Zhou, T. and Zhang, N. (2023). Does air pollution decrease labor share? evidence from china. *Global Environmental Change*, 82:102706.



# A Appendix

Table A1: Summary statistics of data

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A-1: Pollution data at city level					
NO2 ( $\mu\text{g}/\text{m}^3$ )	29.1	16.3	1	163	63,968
SO2 ( $\mu\text{g}/\text{m}^3$ )	14.0	13.0	1	251	63,968
PM25 ( $\mu\text{g}/\text{m}^3$ )	41.3	32.1	2	375	63,968
PM10 ( $\mu\text{g}/\text{m}^3$ )	77.2	60.9	0	1953	63,968
Panel A-2: Pollution data at county level					
NO2 ( $\mu\text{g}/\text{m}^3$ )	18.9	15.9	.1	170	549,024
SO2 ( $\mu\text{g}/\text{m}^3$ )	10.7	8.8	0	286	549,024
PM25 ( $\mu\text{g}/\text{m}^3$ )	36.1	29.5	0	835	549,024
PM10 ( $\mu\text{g}/\text{m}^3$ )	59.8	47.5	0	857	549,024
Panel A-3: Pollution data at firm level					
NOx ( $\mu\text{g}/\text{m}^3$ )	96.5	127	0	18,346	883,036
SO2 ( $\mu\text{g}/\text{m}^3$ )	45.8	159.7	0	30,540	892,740
Panel B-1: Weather data at city level					
Highest temperature ( $^{\circ}\text{C}$ )	19.6	11.3	-27	40	63,637
Lowest temperature ( $^{\circ}\text{C}$ )	9.8	11.7	-40	30	63,637
Wind index	2.9	1.1	1	17.3	63,637
Weather index	19.0	8.8	1	45	63,637
Panel B-2: Weather data at county level					
Highest temperature ( $^{\circ}\text{C}$ )	18.7	11.6	-33	45	409,196
Lowest temperature ( $^{\circ}\text{C}$ )	8.2	12.0	-45	32	409,196
Wind index	3.0	1.5	1	19	409,196
Weather index	19.1	9.7	1	45	409,196
Panel C: Media data (common to city and county levels)					City sample/county sample
Tone	-2.7	2.8	-14.9	4.04	63,968/549,024
Positive score	2.6	1.1	0.7	7.1	63,968/549,024
Negative score	5.3	2.3	0.7	16.4	63,968/549,024
Polarity (%)	7.9	2.3	2.1	18.4	63,968/549,024
Count (number of words)	575.7	318.3	174	1,905	63,968/549,024
Articles from official media (%)	0.5	0.4	0	1	63,968/549,024
Nb of articles	2.2	1.8	1	13	63,968/549,024
Panel D: Financial data (common to city and county levels)					City sample/county sample
Energy price index	1094.7	69.9	884	1232.3	44,265/382,872
SSE index (stock market index)	2,996.5	295.3	2,493.9	3,559.5	44,265/382,872
Exchange rate USD/RMB	7.8	0.1	7.4	8	44,265/382,872
Shibor (interest rate)	2.5	0.2	1.5	2.9	45,693/382,872

Note: all variables are at daily level. The time span is the year 2018.

## A.1 Additional information - Data

We collect information on air pollution at county and city levels from two different databases.

### 1. County level air pollution emissions data

- The county-level air pollution data was derived from the Chinese Air Quality Re-analysis dataset ([Tang et al., 2020b](#)).
- The dataset was produced by the chemical data assimilation system (ChemDAS) developed by the Chinese Academy of Sciences (IAP, CAS), which assimilates over 1000 surface air quality monitoring sites from CNEMC based on the ensemble Kalman filter (EnKF) and the Nested Air Quality Prediction Modeling System (NAQPMS).
- This dataset provides the surface fields of six conventional air pollutants (i.e. PM2.5, PM10, SO2, NO2, CO and O3) in China in a high accuracy and high spatial (15km) and temporal (1 hour) resolutions.
- We aggregate the high-resolution data to the county level and at daily frequency for our analysis.

### 2. City level air pollution emissions data:

- Daily monitoring data from China National Environmental Monitoring Center (CNEMC).
- The city-level data was collected and aggregated by the Geographic Data Sharing Infrastructure, College of Urban and Environmental Science, Peking University.

## A.2 Estimation results

### A.2.1 Figures

Figure A1: Positive emotions and air pollution at city level

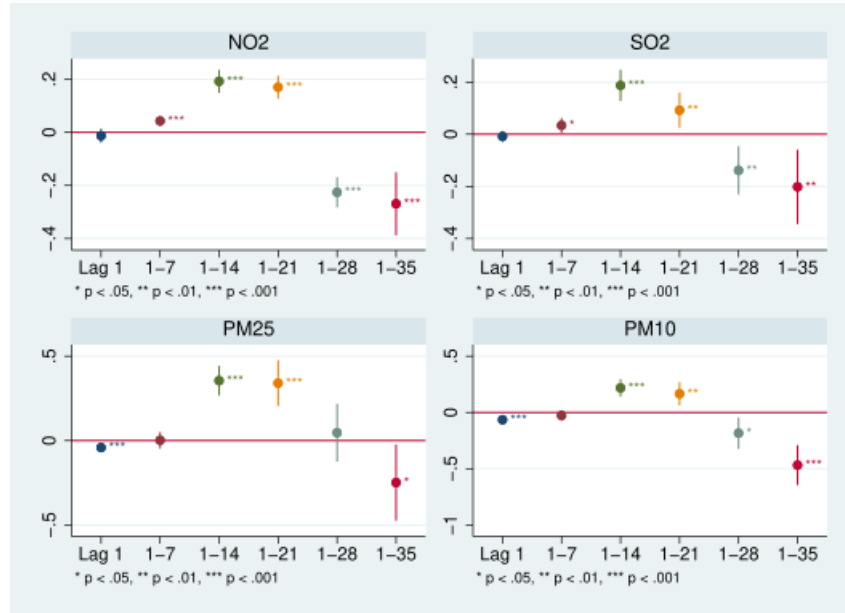


Figure A2: Negative emotions and air pollution at city level

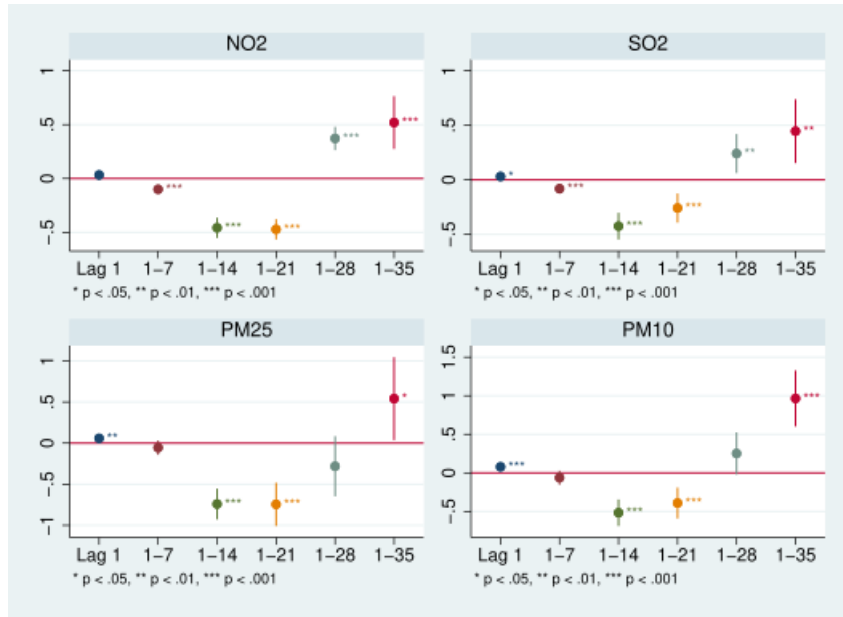


Figure A3: Positive emotions and air pollution at county level

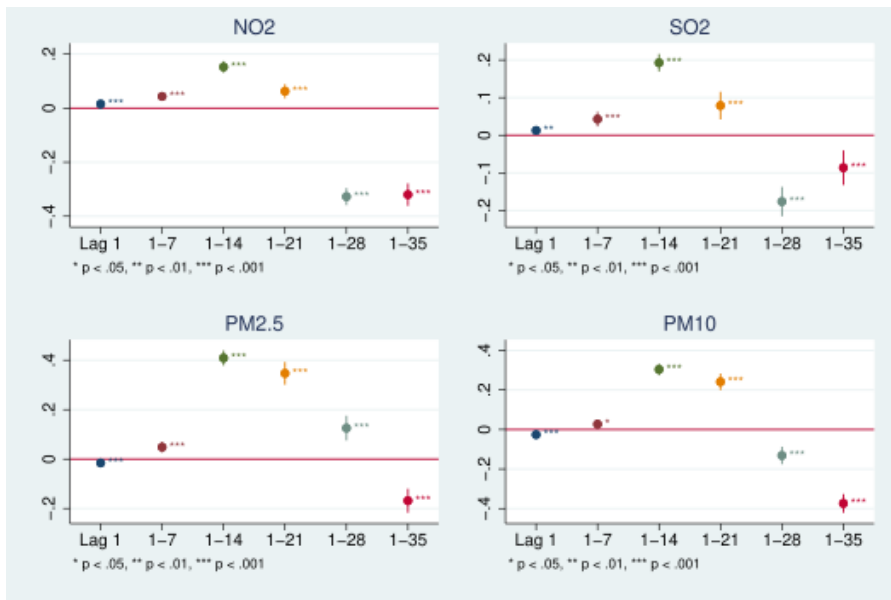
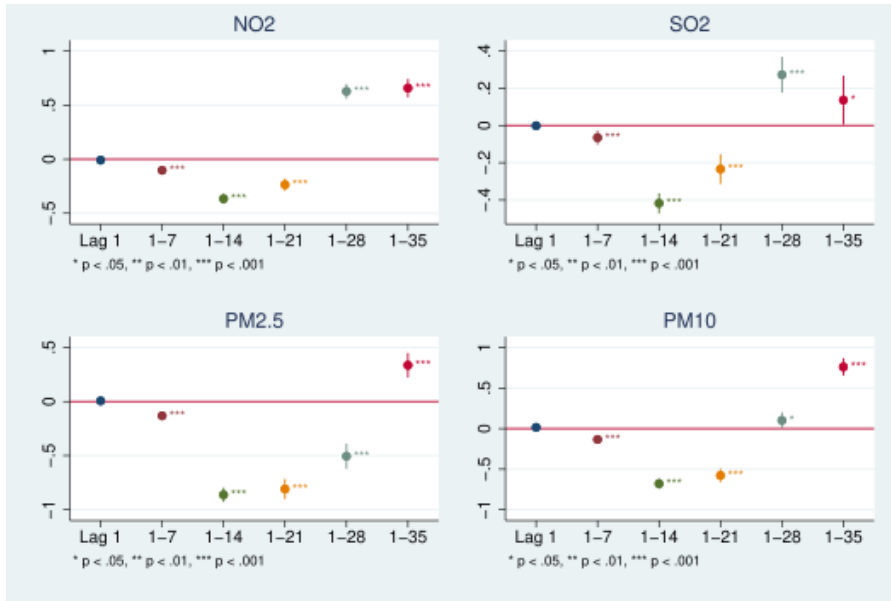


Figure A4: Negative emotions and air pollution at county level



## A.2.2 Tables

Table A2: Dynamic effects of tone, positive score and negative score on air pollution with new control variables

Dep. variable	(1) lnno2 Lag 7	(2) lnno2 Lag 14	(3) lnno2 Lag 21	(4) lnso2 Lag 7	(5) lnso2 Lag 14	(6) lnso2 Lag 21	(7) lnpm25 Lag 7	(8) lnpm25 Lag 14	(9) lnpm25 Lag 21	(10) lnpm10 Lag 7	(11) lnpm10 Lag 14	(12) lnpm10 Lag 21
Panel A.1 - Tone at city level												
Tone	0.047*** (0.0057)	0.13*** (0.013)	0.16*** (0.017)	0.027*** (0.0068)	0.096*** (0.018)	0.062*** (0.020)	0.011 (0.014)	0.20*** (0.032)	0.23*** (0.040)	0.016 (0.015)	0.16*** (0.029)	0.15*** (0.033)
Observations	98,582	98,868	98,868	98,582	98,868	98,868	98,582	98,868	98,868	98,582	98,868	98,868
R-squared	0.753	0.742	0.735	0.770	0.763	0.759	0.621	0.613	0.605	0.638	0.630	0.621
F-statistic	206	216	402	53.1	42.1	48.3	160	196	82.7	149	176	132
Panel A.2 - Positive score at city level												
Positive score	0.037*** (0.0086)	0.14*** (0.016)	0.13*** (0.021)	0.011 (0.012)	0.11*** (0.026)	0.035 (0.026)	-0.020 (0.023)	0.29*** (0.047)	0.29*** (0.060)	-0.051** (0.023)	0.19*** (0.042)	0.15*** (0.048)
Observations	98,582	98,868	98,868	98,582	98,868	98,868	98,582	98,868	98,868	98,582	98,868	98,868
R-squared	0.752	0.741	0.734	0.770	0.763	0.759	0.621	0.612	0.605	0.638	0.629	0.620
F-statistic	206	206	268	48.5	41.8	57.6	158	223	84.9	150	193	128
Panel A.3 - Negative score at city level												
Negative score	-0.11*** (0.017)	-0.38*** (0.038)	-0.43*** (0.049)	-0.058*** (0.020)	-0.29*** (0.058)	-0.17*** (0.058)	-0.026 (0.043)	-0.64*** (0.10)	-0.68*** (0.12)	-0.030 (0.047)	-0.48*** (0.093)	-0.41*** (0.097)
Observations	98,582	98,868	98,868	98,582	98,868	98,868	98,582	98,868	98,868	98,582	98,868	98,868
R-squared	0.752	0.742	0.735	0.770	0.763	0.759	0.621	0.613	0.605	0.638	0.630	0.621
F-statistic	202	214	379	52.2	41.6	49.4	160	213	85.4	149	180	132
Number of cities (panels A)	286	286	286	286	286	286	286	286	286	286	286	286
Number of provinces (panels A)	31	31	31	31	31	31	31	31	31	31	31	31
Panel B.1 - Tone at county level												
Tone	0.051*** (0.0038)	0.12*** (0.0068)	0.11*** (0.0100)	0.033*** (0.0060)	0.11*** (0.0091)	0.068*** (0.014)	0.051*** (0.0038)	0.12*** (0.0068)	0.11*** (0.0100)	0.033*** (0.0060)	0.11*** (0.0091)	0.068*** (0.014)
Observations	644,427	644,446	644,447	644,427	644,446	644,447	644,427	644,446	644,447	644,427	644,446	644,447
R-squared	0.916	0.915	0.913	0.833	0.833	0.832	0.916	0.915	0.913	0.833	0.833	0.832
F-statistic	167	152	116	51.5	62.6	62.2	167	152	116	51.5	62.6	62.2
Panel B.2 - Positive score at county level												
Positive score	0.0080** (0.0034)	0.12*** (0.0091)	0.064*** (0.013)	0.047*** (0.0096)	0.14*** (0.012)	0.053*** (0.018)	0.050*** (0.010)	0.42*** (0.015)	0.37*** (0.023)	0.0041 (0.012)	0.29*** (0.015)	0.29*** (0.021)
Observations	405,614	644,446	644,447	644,427	644,446	644,447	644,427	644,446	644,447	644,427	644,446	644,447
R-squared	0.921	0.915	0.913	0.833	0.833	0.832	0.671	0.668	0.666	0.779	0.777	0.774
F-statistic	112	150	114	48.6	59.7	57.4	246	321	217	185	218	115
Panel B.3 - Negative score at county level												
Negative score	-0.0030 (0.0048)	-0.33*** (0.021)	-0.27*** (0.028)	-0.079*** (0.018)	-0.33*** (0.028)	-0.20*** (0.041)	-0.15*** (0.018)	-0.91*** (0.033)	-0.88*** (0.046)	-0.11*** (0.020)	-0.68*** (0.032)	-0.70*** (0.042)
Observations	405,614	644,446	644,447	644,427	644,446	644,447	644,427	644,446	644,447	644,427	644,446	644,447
R-squared	0.921	0.915	0.913	0.833	0.833	0.832	0.672	0.669	0.666	0.779	0.777	0.774
F-statistic	112	151	115	47.3	62.1	61.6	249	324	233	182	225	127
Number of counties (panels B)	1805	1805	1805	1805	1805	1805	1805	1805	1805	1805	1805	1805
Number of cities (panels B)	325	325	325	325	325	325	325	325	325	325	325	325

Note: In each specification, control variables not presented are: (1) polarity, count, number of articles and percentage of official articles (all media variables are lagged); (2) Htmpt (highest temperature), Ltmpt (lowest temperature), wind and weather (all weather variables are both not lagged and lagged); (3) a time trend and its squared term; (4) city fixed effects (panels A.1, A.2 and A.3), county fixed effects (panels B.1, B.2 and B.3), weekday fixed effects (all panels), city-weekday fixed effects (panels A.1, A.2 and A.3), county-weekday fixed effects (panels B.1, B.2 and B.3), city-month fixed effects (panels A.1, A.2 and A.3) and county-month fixed effects (panels B.1, B.2 and B.3). Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A3: Dynamic effects of tone, positive score and negative score on air pollution with financial control variables at city level

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	lnno2	lnno2	lnno2	lnso2	lnso2	lnso2	lnpm25	lnpm25	lnpm25	lnpm10	lnpm10	lnpm10
	Lag 1-7	Lag 1-14	Lag 1-21	Lag 1-7	Lag 1-14	Lag 1-21	Lag 1-7	Lag 1-14	Lag 1-21	Lag 1-7	Lag 1-14	Lag 1-21
Panel A.1 - Effect of tone at city level												
Tone (lag)	0.039*** (0.0057)	0.072*** (0.010)	0.076*** (0.017)	0.013* (0.0063)	0.053*** (0.018)	-0.019 (0.018)	0.010 (0.013)	0.14*** (0.032)	0.076* (0.043)	0.0052 (0.013)	0.097*** (0.028)	-0.019 (0.030)
SSE index (lag)	1.38*** (0.17)	2.05*** (0.19)	2.35*** (0.34)	0.90*** (0.25)	1.43*** (0.31)	2.29*** (0.43)	0.72* (0.40)	2.23*** (0.56)	4.37*** (0.90)	1.19*** (0.32)	2.34*** (0.45)	4.63*** (0.66)
Energy price (lag)	-0.32*** (0.11)	-0.13 (0.11)	0.30* (0.15)	-0.48*** (0.11)	-0.47*** (0.12)	-0.32** (0.12)	-0.76*** (0.17)	-0.45* (0.25)	-0.10 (0.23)	-1.74*** (0.18)	-1.05*** (0.25)	-0.80*** (0.19)
Exchange rate USD/RMB (lag)	0.061 (0.45)	-0.24 (0.60)	0.96 (0.66)	-2.03*** (0.67)	-0.58 (1.05)	1.83* (0.90)	0.45 (1.09)	0.37 (1.20)	3.08** (1.27)	-2.06** (0.99)	0.84 (1.29)	3.36*** (1.15)
Shibor (lag)	0.29*** (0.072)	0.20* (0.11)	0.69*** (0.21)	0.073 (0.090)	-0.12 (0.090)	0.13 (0.16)	0.49*** (0.12)	0.70*** (0.20)	0.77** (0.37)	0.19* (0.11)	0.32* (0.18)	-0.0045 (0.30)
Observations	97,152	98,576	98,576	97,152	98,576	98,576	97,152	98,576	98,576	97,152	98,576	98,576
R-squared	0.755	0.745	0.738	0.770	0.764	0.761	0.623	0.615	0.611	0.643	0.634	0.629
F-statistic	261	389	405	109	38.8	37.4	709	386	148	671	252	155
Panel A.2 - Effect of positive score at city level												
Positive score (lag)	0.022** (0.0085)	0.047*** (0.015)	0.031 (0.024)	-0.0063 (0.012)	0.054* (0.029)	-0.062** (0.029)	-0.018 (0.023)	0.20*** (0.052)	0.099 (0.073)	-0.057** (0.021)	0.088* (0.044)	-0.056 (0.052)
SHCOMP index (lag)	1.51*** (0.17)	2.31*** (0.22)	2.59*** (0.35)	0.98*** (0.25)	1.55*** (0.31)	2.38*** (0.44)	0.82* (0.41)	2.32*** (0.58)	4.42*** (0.92)	1.37*** (0.31)	2.59*** (0.46)	4.70*** (0.68)
Energy price (lag)	-0.32*** (0.11)	-0.17 (0.11)	0.28* (0.15)	-0.47*** (0.11)	-0.49*** (0.12)	-0.34** (0.13)	-0.74*** (0.17)	-0.50* (0.25)	-0.089 (0.23)	-1.69*** (0.18)	-1.09*** (0.25)	-0.81*** (0.19)
Exchange rate USD/RMB (lag)	-0.48 (0.43)	-0.63 (0.60)	0.89 (0.66)	-2.27*** (0.66)	-0.82 (1.05)	1.83* (0.91)	0.19 (1.13)	-0.071 (1.21)	3.03** (1.27)	-2.43** (1.04)	0.38 (1.29)	3.36*** (1.16)
Shibor (lag)	0.25*** (0.071)	0.16 (0.11)	0.66*** (0.21)	0.052 (0.092)	-0.13 (0.089)	0.12 (0.16)	0.46*** (0.12)	0.69*** (0.20)	0.76** (0.37)	0.14 (0.11)	0.28 (0.18)	-0.010 (0.30)
Observations	97,152	98,576	98,576	97,152	98,576	98,576	97,152	98,576	98,576	97,152	98,576	98,576
R-squared	0.754	0.744	0.738	0.770	0.764	0.761	0.623	0.615	0.611	0.643	0.634	0.629
F-statistic	268	418	388	105	39.7	46.4	752	386	149	532	252	155
Panel A.3 - Effect of negative score at city level												
Negative score (lag)	-0.089*** (0.017)	-0.19*** (0.030)	-0.19*** (0.049)	-0.017 (0.019)	-0.16** (0.059)	0.059 (0.056)	-0.032 (0.039)	-0.45*** (0.10)	-0.24* (0.13)	-0.0053 (0.042)	-0.27*** (0.090)	0.064 (0.093)
SHCOMP index (lag)	1.44*** (0.17)	2.13*** (0.20)	2.40*** (0.34)	0.94*** (0.25)	1.45*** (0.31)	2.29*** (0.43)	0.72* (0.41)	2.23*** (0.56)	4.36*** (0.91)	1.20*** (0.32)	2.41*** (0.45)	4.64*** (0.66)
Energy price (lag)	-0.32*** (0.11)	-0.15 (0.11)	0.30* (0.15)	-0.48*** (0.11)	-0.48*** (0.12)	-0.32** (0.12)	-0.76*** (0.17)	-0.48* (0.25)	-0.096 (0.23)	-1.74*** (0.18)	-1.08*** (0.25)	-0.80*** (0.19)
Exchange rate USD/RMB (lag)	-0.13 (0.45)	-0.39 (0.60)	0.94 (0.66)	-2.16*** (0.67)	-0.66 (1.05)	1.83* (0.90)	0.44 (1.11)	0.23 (1.20)	3.07** (1.27)	-2.12** (1.01)	0.66 (1.29)	3.36*** (1.16)
Shibor (lag)	0.27*** (0.072)	0.19* (0.11)	0.68*** (0.21)	0.064 (0.091)	-0.12 (0.089)	0.13 (0.16)	0.49*** (0.12)	0.71*** (0.20)	0.77** (0.37)	0.18 (0.11)	0.31* (0.18)	-0.0052 (0.30)
Observations	97,152	98,576	98,576	97,152	98,576	98,576	97,152	98,576	98,576	97,152	98,576	98,576
R-squared	0.755	0.745	0.738	0.770	0.764	0.761	0.623	0.615	0.611	0.643	0.634	0.629
F-statistic	261	397	401	114	39.6	39.4	725	391	147	606	255	155
Number of cities (all panels)	286	286	286	286	286	286	286	286	286	286	286	286
Number of provinces (all panels)	31	31	31	31	31	31	31	31	31	31	31	31

Note: In each specification, other control variables not presented are: (1) polarity, count, number of articles and percentage of official articles (all media variables are lagged); (2) Htmpt (highest temperature), Lttmp (lowest temperature), wind and weather (all weather variables are both not lagged and lagged); (3) a time trend and its squared term; (4) city fixed effects, weekday fixed effects, city-weekday fixed effects, city-month fixed effects. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A4: Dynamic effects of tone, positive score and negative score on air pollution with financial control variables at county level

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	lnno2	lnno2	lnno2	lnso2	lnso2	lnso2	lnpm25	lnpm25	lnpm25	lnpm10	lnpm10	lnpm10
	Lag 1-7	Lag 1-14	Lag 1-21	Lag 1-7	Lag 1-14	Lag 1-21	Lag 1-7	Lag 1-14	Lag 1-21	Lag 1-7	Lag 1-14	Lag 1-21
Panel A.1 - Effect of tone at county level												
Tone (lag)	0.044*** (0.0035)	0.080*** (0.0063)	0.058*** (0.0093)	0.018*** (0.0059)	0.060*** (0.0100)	-0.0061 (0.017)	0.051*** (0.0054)	0.25*** (0.011)	0.19*** (0.017)	0.024*** (0.0064)	0.15*** (0.012)	0.062*** (0.016)
SHCOMP index (lag)	1.05*** (0.099)	1.31*** (0.13)	1.63*** (0.17)	1.04*** (0.15)	1.47*** (0.16)	2.09*** (0.20)	0.22 (0.17)	1.46*** (0.21)	2.91*** (0.32)	1.76*** (0.16)	2.74*** (0.21)	4.74*** (0.26)
Energy price (lag)	0.15** (0.078)	0.44*** (0.091)	0.87*** (0.10)	0.26*** (0.095)	0.20* (0.10)	0.28** (0.11)	0.29*** (0.092)	0.31*** (0.10)	0.48*** (0.10)	-1.07*** (0.096)	-0.80*** (0.11)	-0.51*** (0.11)
Exchange rate USD/RMB (lag)	0.31 (0.26)	0.38 (0.35)	1.81*** (0.34)	-2.11*** (0.37)	-1.92*** (0.44)	-0.047 (0.46)	0.64* (0.39)	1.63*** (0.46)	4.46*** (0.54)	-1.19*** (0.37)	0.71 (0.49)	3.62*** (0.54)
Shibor (lag)	0.32*** (0.035)	0.37*** (0.049)	0.76*** (0.081)	0.28*** (0.041)	0.28*** (0.065)	0.80*** (0.11)	0.71*** (0.054)	1.53*** (0.081)	1.64*** (0.14)	0.22*** (0.048)	0.57*** (0.078)	0.34*** (0.12)
Observations	635,412	642,644	642,645	635,412	642,644	642,645	635,412	642,644	642,645	635,412	642,644	642,645
R-squared	0.917	0.915	0.914	0.833	0.833	0.833	0.672	0.671	0.669	0.781	0.780	0.778
F-statistic	140	150	132	53.9	75.3	89	228	304	219	212	256	184
Panel A.2 - Effect of positive score at county level												
Positive score (lag)	0.025*** (0.0060)	0.054*** (0.0097)	0.0011 (0.013)	0.022** (0.010)	0.084*** (0.014)	-0.026 (0.024)	0.045*** (0.0096)	0.37*** (0.017)	0.25*** (0.027)	-0.025** (0.012)	0.17*** (0.018)	0.061** (0.024)
SHCOMP index (lag)	1.20*** (0.10)	1.58*** (0.14)	1.88*** (0.17)	1.07*** (0.15)	1.53*** (0.17)	2.14*** (0.21)	0.34** (0.17)	1.62*** (0.22)	3.05*** (0.33)	1.95*** (0.16)	3.00*** (0.21)	4.84*** (0.26)
Energy price (lag)	0.16** (0.079)	0.42*** (0.091)	0.84*** (0.10)	0.25** (0.096)	0.18* (0.10)	0.27** (0.11)	0.28*** (0.094)	0.25** (0.10)	0.53*** (0.10)	-1.04*** (0.097)	-0.84*** (0.11)	-0.51*** (0.12)
Exchange rate USD/RMB (lag)	-0.31 (0.27)	-0.024 (0.35)	1.74*** (0.34)	-2.30*** (0.37)	-2.11*** (0.44)	-0.057 (0.46)	0.020 (0.39)	0.91** (0.45)	4.38*** (0.54)	-1.72*** (0.37)	0.16 (0.47)	3.58*** (0.54)
Shibor (lag)	0.29*** (0.035)	0.33*** (0.049)	0.74*** (0.081)	0.28*** (0.041)	0.27*** (0.065)	0.80*** (0.11)	0.68*** (0.054)	1.51*** (0.082)	1.62*** (0.14)	0.18*** (0.048)	0.54*** (0.078)	0.33*** (0.12)
Observations	635,412	642,644	642,645	635,412	642,644	642,645	635,412	642,644	642,645	635,412	642,644	642,645
R-squared	0.916	0.915	0.914	0.833	0.833	0.833	0.672	0.671	0.669	0.781	0.779	0.778
F-statistic	137	149	127	52.2	77.1	88.7	232	307	218	224	254	187
Panel A.3 - Effect of negative score at county level												
Negative score (lag)	-0.100*** (0.011)	-0.21*** (0.020)	-0.14*** (0.018)	-0.029 (0.032)	-0.19*** (0.032)	0.0045 (0.050)	-0.15*** (0.017)	-0.81*** (0.035)	-0.58*** (0.051)	-0.060*** (0.020)	-0.43*** (0.037)	-0.18*** (0.047)
SSE index (lag)	1.11*** (0.100)	1.40*** (0.13)	1.69*** (0.17)	1.08*** (0.15)	1.48*** (0.16)	2.08*** (0.20)	0.24 (0.17)	1.45*** (0.22)	2.92*** (0.32)	1.79*** (0.16)	2.82*** (0.21)	4.75*** (0.26)
Energy price (lag)	0.15* (0.078)	0.43*** (0.091)	0.87*** (0.10)	0.26*** (0.096)	0.19* (0.10)	0.29*** (0.11)	0.28*** (0.093)	0.26*** (0.10)	0.50*** (0.10)	-1.08*** (0.096)	-0.83*** (0.11)	-0.50*** (0.11)
Exchange rate USD/RMB (lag)	0.081 (0.27)	0.21 (0.35)	1.79*** (0.34)	-2.26*** (0.37)	-1.99*** (0.44)	-0.042 (0.46)	0.57 (0.39)	1.41*** (0.45)	4.45*** (0.54)	-1.28*** (0.38)	0.50 (0.48)	3.61*** (0.54)
Shibor (lag)	0.31*** (0.035)	0.36*** (0.049)	0.76*** (0.081)	0.27*** (0.041)	0.28*** (0.065)	0.80*** (0.11)	0.71*** (0.054)	1.55*** (0.081)	1.64*** (0.14)	0.21*** (0.048)	0.57*** (0.078)	0.34*** (0.12)
Observations	635,412	642,644	642,645	635,412	642,644	642,645	635,412	642,644	642,645	635,412	642,644	642,645
R-squared	0.917	0.915	0.914	0.833	0.833	0.833	0.672	0.671	0.669	0.781	0.779	0.778
F-statistic	141	149	130	51.1	76	89.2	228	307	219	221	255	184
Number of counties (all panels)	1,805	1,805	1,805	1,805	1,805	1,805	1,805	1,805	1,805	1,805	1,805	1,805
Number of cities (all panels)	325	325	325	325	325	325	325	325	325	325	325	325

Note: In each specification, other control variables not presented are: (1) polarity, count, number of articles and percentage of official articles (all media variables are lagged); (2) Htmp (highest temperature), Ltmp (lowest temperature), wind and weather (all weather variables are both not lagged and lagged); (3) a time trend and its squared term; (4) county fixed effects, weekday fixed effects, county-weekday fixed effects, county-month fixed effects. Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1



Table A5: Dynamic effects of tone on air pollution with new lag calculation

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
	lnno2	lnso2	lnpm25	lnpm10	lnno2	lnso2	lnpm25	lnpm10	
		Panel A - city level					Panel B - county level		
Tone (1-7)	0.041*** (0.0094)	0.045*** (0.0086)	0.058*** (0.021)	0.054*** (0.019)	0.043*** (0.0049)	0.037*** (0.0059)	0.11*** (0.0080)	0.077*** (0.0080)	
Tone (8-14)	0.032*** (0.010)	0.043*** (0.015)	0.17*** (0.021)	0.13*** (0.016)	0.023*** (0.0065)	0.039*** (0.0093)	0.26*** (0.011)	0.13*** (0.010)	
Tone (15-21)	0.0091 (0.0091)	0.012 (0.0074)	0.058*** (0.019)	0.035** (0.016)	0.0061 (0.0043)	0.032*** (0.0083)	0.10*** (0.010)	0.046*** (0.0092)	
Tone (22-28)	-0.054*** (0.0077)	-0.021** (0.0080)	0.018 (0.012)	-0.015 (0.0098)	-0.026*** (0.0051)	0.0026 (0.0067)	0.088*** (0.0082)	-0.024*** (0.0078)	
Tone (29-35)	0.016* (0.0079)	0.015** (0.0063)	-0.0092 (0.017)	-0.024 (0.015)	0.034*** (0.0043)	0.025*** (0.0055)	-0.014** (0.0058)	-0.037*** (0.0064)	
Observations	90,086	90,086	90,086	90,086	588,028	588,028	588,028	588,028	
R-squared	0.762	0.763	0.630	0.659	0.922	0.832	0.681	0.784	
Number of ID	286	286	286	286	1,805	1,805	1,805	1,805	
Number of upper ID	31	31	31	31	325	325	325	325	

Note: In each specification, control variables not presented are: (1) polarity, count, number of articles and percentage of official articles (all media variables are lagged in the same way than tone variable); (2) Htmpt (highest temperature), Ltmpt (lowest temperature), wind and weather (all weather variables are both not lagged and lagged (in the same way than tone variable)); (3) a time trend and its squared term; (4) city fixed effects (panel A), county fixed effects (panel B), weekday fixed effects (all panels), city-weekday fixed effects (panel A), county-weekday fixed effects (panel B), city-month fixed effects (panel A) and county-month fixed effects (panel B). Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A6: Dynamic effects of positive score on air pollution with new lag calculation

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	lnno2	lnso2	lnpm25	lnpm10	lnno2	lnso2	lnpm25	lnpm10
	Panel A - city level				Panel B - county level			
Positive score (1-7)	0.061*** (0.015)	0.056*** (0.015)	0.072** (0.032)	0.048 (0.029)	0.081*** (0.0082)	0.084*** (0.010)	0.16*** (0.012)	0.089*** (0.014)
Positive score (8-14)	0.059*** (0.018)	0.039* (0.021)	0.21*** (0.030)	0.15*** (0.024)	0.026** (0.011)	0.070*** (0.015)	0.43*** (0.017)	0.19*** (0.016)
Positive score (15-21)	0.018 (0.023)	0.019 (0.016)	0.081** (0.036)	0.069** (0.031)	-0.028*** (0.0075)	0.019 (0.014)	0.13*** (0.014)	0.049*** (0.013)
Positive score (22-28)	-0.092*** (0.014)	-0.060*** (0.016)	-0.017 (0.025)	-0.054** (0.021)	-0.079*** (0.0090)	-0.037*** (0.014)	0.088*** (0.012)	-0.095*** (0.012)
Positive score (29-35)	0.056*** (0.014)	0.027** (0.011)	-0.029 (0.032)	-0.049* (0.025)	0.059*** (0.0074)	0.063*** (0.0088)	-0.023** (0.0097)	-0.059*** (0.010)
Observations	90,086	90,086	90,086	90,086	588,028	588,028	588,028	588,028
R-squared	0.762	0.763	0.628	0.657	0.922	0.832	0.680	0.783
Number of ID	286	286	286	286	1,805	1,805	1,805	1,805
Number of upper ID	31	31	31	31	325	325	325	325

Note: In each specification, control variables not presented are: (1) polarity, count, number of articles and percentage of official articles (all media variables are lagged in the same way than positive score variable); (2) Htmpt (highest temperature), Ltmpt (lowest temperature), wind and weather (all weather variables are both not lagged and lagged (in the same way than positive score variable)); (3) a time trend and its squared term; (4) city fixed effects (panel A), county fixed effects (panel B), weekday fixed effects (all panels), city-weekday fixed effects (panel A), county-weekday fixed effects (panel B), city-month fixed effects (panel A) and county-month fixed effects (panel B). Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A7: Dynamic effects of negative score on air pollution with new lag calculation

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	lnno2	lnso2	lnpm25	lnpm10	lnno2	lnso2	lnpm25	lnpm10
	Panel A - city level				Panel B - county level			
Negative score (1-7)	-0.11*** (0.028)	-0.12*** (0.023)	-0.18*** (0.058)	-0.17*** (0.055)	-0.11*** (0.014)	-0.059*** (0.016)	-0.26*** (0.021)	-0.25*** (0.022)
Negative score (8-14)	-0.10*** (0.031)	-0.10** (0.042)	-0.50*** (0.061)	-0.37*** (0.047)	-0.12*** (0.020)	-0.12*** (0.028)	-0.78*** (0.034)	-0.37*** (0.031)
Negative score (15-21)	-0.033 (0.029)	-0.029 (0.020)	-0.18*** (0.060)	-0.11** (0.050)	-0.027** (0.013)	-0.060** (0.025)	-0.29*** (0.029)	-0.16*** (0.027)
Negative score (22-28)	0.16*** (0.027)	0.073*** (0.022)	-0.074* (0.043)	0.038 (0.033)	0.037** (0.016)	-0.063*** (0.023)	-0.36*** (0.027)	-0.0021 (0.025)
Negative score (29-35)	-0.058** (0.025)	-0.045** (0.020)	0.033 (0.054)	0.071 (0.045)	-0.089*** (0.014)	-0.061*** (0.020)	0.0031 (0.020)	0.14*** (0.022)
Observations	90,086	90,086	90,086	90,086	588,028	588,028	588,028	588,028
R-squared	0.762	0.763	0.630	0.658	0.922	0.831	0.681	0.784
Number of ID	286	286	286	286	1,805	1,805	1,805	1,805
Number of upper ID	31	31	31	31	325	325	325	325

Note: In each specification, control variables not presented are: (1) polarity, count, number of articles and percentage of official articles (all media variables are lagged in the same way than negative score variable); (2) Htmpt (highest temperature), Ltmpt (lowest temperature), wind and weather (all weather variables are both not lagged and lagged (in the same way than negative score variable)); (3) a time trend and its squared term; (4) city fixed effects (panel A), county fixed effects (panel B), weekday fixed effects (all panels), city-weekday fixed effects (panel A), county-weekday fixed effects (panel B), city-month fixed effects (panel A) and county-month fixed effects (panel B). Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A8: Non-parametric RD estimates of the (LATE) impact of the tone of news media on drugs at the city and county levels – with fixed effects

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidth selection	lnno2	lnso2	lnpm25	lnpm10	lnno2	lnso2	lnpm25	lnpm10
	Two CER				Common CER			
	Panel A - City level							
Ln( $\text{tone}_{t-1}$ )	0.0573 (0.105)	0.0871 (0.0619)	0.120** (0.0514)	0.0455 (0.0570)	0.00943 (0.0710)	0.0190 (0.0475)	0.0697 (0.0437)	0.0501 (0.0352)
Observations	64,669	64,669	64,669	64,669	64,669	64,669	64,669	64,669
bandwidth below c	5.362	11.10	5.969	5.388	5.713	9.817	10.32	8.675
bandwidth above c	23.10	22.31	12.74	11.71	5.713	9.817	10.32	8.675
	Panel B - County level							
Ln( $\text{tone}_{t-1}$ )	0.00774 (0.0164)	-0.496*** (0.0380)	.	.	0.309* (0.163)	0.00133 (0.0268)	0.00488 (0.0207)	-0.0423*** (0.00891)
Observations	405,690	405,690	.	.	405,690	405,690	405,690	405,690
bandwidth below c	12.48	24.71	.	.	4.385	6.368	6.599	2.551
bandwidth above c	9.141	4.351	.	.	4.385	6.368	6.599	2.551

Note: The estimation sample excludes cities and counties targeted by the government’s June 27 announcement, whose date is used as the cutoff. RD estimates with non-parametric local linear regressions. Conventional local linear regression robust standard errors are clustered at the city (panel A)/county (panel B) levels are reported in parentheses. CER = coverage error-rate. The optimal bandwidth is selected using the approach of [Calonico et al. \(2014\)](#). Fixed effects (FE) controlled for are: ID FE (city in panel A and county in panel B), weekday FE (both panel), province\*month FE (panel A) and city\*month (panel B). No estimation results in cols. 3 and 4 of Panel B due to computation problems. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A9: Dynamic effect of tone on firms' SO2 emission

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lag	lnso2 t=0	lnso2 t=-1	lnso2 t=-1 to -2	lnso2 t=-1 to -3	lnso2 t=-1 to -4	lnso2 t=-1 to -5	lnso2 t=-1 to -6	lnso2 t=-1 to -7
Lntone (lagged)	0.0046** (0.0020)	0.0092*** (0.0025)	0.0091*** (0.0028)	0.0066*** (0.0023)	0.0092*** (0.0025)	0.0070** (0.0033)	0.010** (0.0042)	0.011** (0.0047)
Lnpolarity (lagged)	0.0094 (0.0072)	0.036*** (0.0094)	0.025*** (0.0097)	0.014 (0.0091)	0.024** (0.010)	0.019* (0.011)	0.022** (0.011)	0.022* (0.012)
D.Lnenergy (not lagged)		-0.061 (0.11)	-0.015 (0.091)	-0.19*** (0.068)	-0.16** (0.065)	-0.15** (0.070)	-0.17*** (0.064)	-0.17*** (0.064)
D.Lnenergy (lagged)	-0.28*** (0.089)	-0.029 (0.11)	0.40*** (0.14)	-0.035 (0.065)	0.10 (0.093)	0.11 (0.12)	0.29* (0.16)	0.16 (0.16)
D.lnSEE index (not lagged)		0.17 (0.14)	0.32*** (0.11)	0.17** (0.084)	0.15* (0.082)	0.13* (0.078)	0.15* (0.082)	0.15* (0.082)
D.lnSEE index (lagged)	0.20 (0.12)	0.30** (0.15)	0.034 (0.22)	0.0093 (0.14)	-0.096 (0.15)	0.12 (0.20)	0.16 (0.24)	0.18 (0.27)
D.Lnexchange rate USD/RMB (not lagged)		-0.27 (0.42)	-1.26*** (0.43)	-0.27 (0.33)	-0.31 (0.31)	-0.29 (0.31)	-0.30 (0.32)	-0.39 (0.32)
D.Lnexchange rate USD/RMB (lagged)	-0.50 (0.41)	0.70 (0.58)	-0.65 (0.90)	1.49** (0.61)	0.86 (0.64)	-0.056 (0.67)	-1.40 (0.85)	-1.28 (1.04)
Htmpt (not lagged)	0.00074 (0.00063)	0.0015** (0.00069)	0.00077 (0.00059)	0.00058 (0.00053)	0.00062 (0.00052)	0.00054 (0.00051)	0.00042 (0.00052)	0.00045 (0.00052)
Ltmpt (not lagged)	-0.0043*** (0.00082)	-0.0024** (0.00096)	-0.0032*** (0.00079)	-0.0026*** (0.00069)	-0.0029*** (0.00068)	-0.0028*** (0.00067)	-0.0026*** (0.00067)	-0.0029*** (0.00066)
Wind (not lagged)	-0.0015 (0.0014)	-0.00098 (0.0016)	0.00054 (0.0013)	0.00036 (0.0011)	0.00026 (0.0011)	0.00026 (0.0011)	0.00068 (0.0011)	0.00052 (0.0011)
Weather (not lagged)	-0.000012 (0.00015)	0.00018 (0.00018)	0.000083 (0.00015)	-0.000059 (0.00013)	-0.000043 (0.00013)	-0.000036 (0.00013)	-0.000031 (0.00012)	-0.000037 (0.00013)
Htmpt (lagged)		0.00044 (0.00076)	-0.00065 (0.00085)	-0.0011 (0.00086)	-0.0012 (0.00093)	-0.0017* (0.00099)	-0.0019* (0.0010)	-0.0019* (0.0011)
Ltmpt (lagged)		-0.0032*** (0.00100)	-0.00084 (0.0010)	-0.0014 (0.0011)	-0.0012 (0.0011)	-0.00096 (0.0012)	-0.0012 (0.0012)	-0.0011 (0.0013)
Wind (lagged)		-0.00048 (0.0017)	-0.0027 (0.0021)	-0.0017 (0.0024)	-0.000084 (0.0028)	0.0015 (0.0032)	0.0016 (0.0035)	0.00095 (0.0039)
Weather (lagged)		0.000073 (0.00018)	-0.00030 (0.00024)	-0.00029 (0.00026)	-0.00020 (0.00030)	-0.00029 (0.00034)	-0.00048 (0.00039)	-0.00060 (0.00044)
Time	-0.0021*** (0.00054)	-0.0018*** (0.00063)	-0.0019*** (0.00060)	-0.0017*** (0.00057)	-0.0017*** (0.00057)	-0.0015*** (0.00058)	-0.0015*** (0.00059)	-0.0014** (0.00059)
Time2	2.2e-06* (1.2e-06)	2.1e-06 (1.4e-06)	1.9e-06 (1.3e-06)	1.3e-06 (1.2e-06)	1.3e-06 (1.2e-06)	9.3e-07 (1.3e-06)	9.4e-07 (1.3e-06)	6.5e-07 (1.3e-06)
Constant	3.16*** (0.059)	3.04*** (0.069)	3.11*** (0.070)	3.14*** (0.069)	3.12*** (0.070)	3.12*** (0.069)	3.13*** (0.068)	3.12*** (0.069)
Observations	276,826	201,430	279,123	376,538	400,208	415,690	424,836	426,410
R-squared	0.765	0.768	0.765	0.764	0.763	0.762	0.762	0.763
F-statistic	11	5.78	6.05	6.29	6.37	5.97	6.07	6.03
Number of firms	4537	4456	4513	4531	4536	4548	4556	4556

Note: The dependent variable is SO2 emissions of industrial firms. Robust clustered standard errors in parentheses. Firm fixed effects, weekday fixed effects, city-weekday fixed effects and city-month fixed effects are controlled for in all columns. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A10: Dynamic effect of tone on firms' NOx emission

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lag	lnnox t=0	lnnox t=-1	lnnox t=-1 to -2	lnnox t=-1 to -3	lnnox t=-1 to -4	lnnox t=-1 to -5	lnnox t=-1 to -6	lnnox t=-1 to -7
Lntone (lagged)	0.0048** (0.0021)	0.011*** (0.0025)	0.015*** (0.0030)	0.0073*** (0.0025)	0.0092*** (0.0027)	0.010*** (0.0034)	0.015*** (0.0043)	0.015*** (0.0048)
Lnpolarity (lagged)	0.010 (0.0074)	0.041*** (0.0095)	0.038*** (0.010)	0.012 (0.0099)	0.019* (0.011)	0.023* (0.012)	0.025** (0.012)	0.022* (0.013)
D.Lnenergy (not lagged)	-0.40*** (0.091)	0.043 (0.11)	-0.061 (0.087)	-0.25*** (0.066)	-0.28*** (0.063)	-0.33*** (0.068)	-0.32*** (0.063)	-0.31*** (0.063)
D.Lnenergy (lagged)		-0.50*** (0.11)	0.25* (0.14)	-0.13* (0.069)	-0.13 (0.10)	-0.11 (0.13)	0.15 (0.17)	0.0029 (0.18)
D.lnSEE index (not lagged)	0.12 (0.13)	0.36*** (0.14)	0.55*** (0.11)	0.27*** (0.086)	0.23*** (0.084)	0.21*** (0.080)	0.24*** (0.084)	0.21** (0.084)
D.lnSEE index (lagged)		0.40*** (0.15)	0.51** (0.23)	0.084 (0.15)	-0.12 (0.16)	-0.25 (0.22)	-0.24 (0.26)	-0.49* (0.29)
D.Lnexchange rate USD/RMB (not lagged)	-0.64 (0.43)	0.058 (0.43)	-1.50*** (0.44)	-0.22 (0.33)	-0.15 (0.32)	-0.31 (0.32)	-0.33 (0.32)	-0.47 (0.33)
D.Lnexchange rate USD/RMB (lagged)		-0.48 (0.58)	-0.73 (0.93)	1.24* (0.66)	0.48 (0.67)	-0.61 (0.70)	-1.92** (0.90)	-2.04* (1.10)
Htmpt (not lagged)	-0.0014** (0.00066)	0.00053 (0.00067)	-0.00080 (0.00058)	-0.00089* (0.00052)	-0.00090* (0.00051)	-0.0011** (0.00050)	-0.0013** (0.00052)	-0.0013** (0.00052)
Ltmpt (not lagged)	-0.0025*** (0.00085)	-0.00097 (0.00094)	-0.0017** (0.00081)	-0.0015** (0.00069)	-0.0017** (0.00067)	-0.0017*** (0.00066)	-0.0015** (0.00067)	-0.0019*** (0.00066)
Wind (not lagged)	-0.0022 (0.0015)	-0.00058 (0.0016)	-0.00065 (0.0013)	-0.00063 (0.0012)	-0.00055 (0.0011)	-0.00067 (0.0011)	-0.00040 (0.0011)	-0.00079 (0.0011)
Weather (not lagged)	-0.00023 (0.00016)	-0.00031* (0.00017)	-0.00015 (0.00015)	-0.00019 (0.00013)	-0.00016 (0.00012)	-0.00018 (0.00012)	-0.00016 (0.00012)	-0.00019 (0.00012)
Htmpt (lagged)		-0.00076 (0.00074)	-0.0011 (0.00085)	-0.0021** (0.00085)	-0.0027*** (0.00091)	-0.0029*** (0.00096)	-0.0032*** (0.0010)	-0.0034*** (0.0011)
Ltmpt (lagged)		-0.0030*** (0.00097)	-0.00072 (0.0010)	-0.00035 (0.0011)	0.00052 (0.0011)	0.00061 (0.0012)	0.00086 (0.0012)	0.00098 (0.0013)
Wind (lagged)		0.0011 (0.0018)	-0.00091 (0.0021)	-0.0035 (0.0023)	-0.0042 (0.0027)	-0.0046 (0.0031)	-0.0040 (0.0035)	-0.0062 (0.0038)
Weather (lagged)		-0.00015 (0.00018)	-0.00067*** (0.00025)	-0.00093*** (0.00028)	-0.0011*** (0.00032)	-0.0014*** (0.00037)	-0.0016*** (0.00042)	-0.0018*** (0.00048)
Time	-0.0016*** (0.00058)	-0.00092 (0.00066)	-0.0013** (0.00063)	-0.0011* (0.00061)	-0.0012* (0.00061)	-0.0011* (0.00061)	-0.0012* (0.00063)	-0.00098 (0.00063)
Time2	6.8e-07 (1.3e-06)	-1.6e-08 (1.5e-06)	9.1e-07 (1.4e-06)	1.2e-07 (1.3e-06)	2.1e-07 (1.3e-06)	-1.4e-07 (1.4e-06)	2.0e-07 (1.4e-06)	-3.6e-07 (1.4e-06)
Constant	4.18*** (0.066)	4.02*** (0.075)	4.08*** (0.077)	4.16*** (0.075)	4.18*** (0.076)	4.19*** (0.075)	4.21*** (0.075)	4.20*** (0.076)
Observations	273,690	199,081	275,946	372,241	395,637	410,945	419,996	421,568
R-squared	0.763	0.766	0.763	0.762	0.762	0.761	0.761	0.762
F-statistic	11.7	8.19	7.49	7.56	7.81	8.24	8.22	8.37
Number of firms	4505	4432	4481	4498	4503	4516	4523	4523

Note: The dependent variable is NOx emissions of industrial firms. Robust clustered standard errors in parentheses. Firm fixed effects, weekday fixed effects, city-weekday fixed effects and city-month fixed effects are controlled for in all columns. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table A11: Impact of media tone on air pollution - Sample 1: cities with a higher number of sanctions (after violation of air pollution standard)

Dep. Var.	(1) pm10	(2) pm10	(3) pm10	(4) pm25	(5) pm25	(6) pm25	(7) no2	(8) no2	(9) no2	(10) so2	(11) so2	(12) so2
Lag	1-7	1-14	1-21	1-7	1-14	1-21	1-7	1-14	1-21	1-7	1-14	1-21
Tone	0.059*** (0.016)	0.11*** (0.028)	0.19*** (0.041)	0.050*** (0.012)	0.12*** (0.025)	0.21*** (0.040)	0.029*** (0.0090)	0.063*** (0.011)	0.043** (0.018)	0.023*** (0.0069)	0.051*** (0.015)	0.069** (0.031)
Observations	48,783	49,478	49,478	48,783	49,478	49,478	48,783	49,478	49,478	48,783	49,478	49,478
R-squared	0.642	0.642	0.644	0.627	0.628	0.630	0.763	0.761	0.762	0.771	0.770	0.771
F-statistic	453	181	199	765	256	152	620	492	897	160	129	517
Nb. of cities	139	139	139	139	139	139	139	139	139	139	139	139
Nb. of provinces	25	25	25	25	25	25	25	25	25	25	25	25

Note: The dependent variable is air pollution. In addition to the emotional tone, we add other media variables (polarity, count, the number of articles), weather variables and time trends; multi-level fixed effects: city, weekday, city\*weekday, city\*month; clustered at city and province levels. Robust standard errors are in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Further details of the results are available upon request.

Table A12: Impact of media tone on air pollution - Sample 2: cities with a lower number of sanctions (after violation of air pollution standard)

Dep. Var.	(1) pm10	(2) pm10	(3) pm10	(4) pm25	(5) pm25	(6) pm25	(7) no2	(8) no2	(9) no2	(10) so2	(11) so2	(12) so2
Lag	1-7	1-14	1-21	1-7	1-14	1-21	1-7	1-14	1-21	1-7	1-14	1-21
Tone	0.030 (0.022)	0.076* (0.040)	0.10 (0.060)	0.025 (0.022)	0.10** (0.039)	0.15** (0.059)	0.015** (0.0063)	0.049*** (0.011)	0.015 (0.019)	0.015*** (0.0046)	0.037*** (0.013)	-0.00020 (0.029)
Observations	49,031	49,732	49,732	49,031	49,732	49,732	49,031	49,732	49,732	49,031	49,732	49,732
R-squared	0.628	0.629	0.630	0.638	0.639	0.640	0.760	0.759	0.759	0.746	0.745	0.745
F-statistic	139	251	375	38.7	127	368	606	385	334	40	36.3	87.6
Nb. of cities	140	140	140	140	140	140	140	140	140	140	140	140
Nb. of provinces	25	25	25	25	25	25	25	25	25	25	25	25

Note: The dependent variable is air pollution. In addition to the emotional tone, we add other media variables (polarity, count, the number of articles), weather variables and time trends; multi-level fixed effects: city, weekday, city\*weekday, city\*month; clustered at city and province levels. Robust standard errors are in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Further details of the results are available upon request.