



**HAL**  
open science

## Environmental News Emotion and Air Pollution in China

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

► **To cite this version:**

Sébastien Marchand, Damien Cubizol, Elda Nasho Ah-Pine, Huanxiu Guo. Environmental News Emotion and Air Pollution in China. 2021. hal-03448375v1

**HAL Id: hal-03448375**

**<https://uca.hal.science/hal-03448375v1>**

Preprint submitted on 25 Nov 2021 (v1), last revised 13 Dec 2023 (v2)

**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.



CENTRE D'ÉTUDES  
ET DE RECHERCHES  
SUR LE DÉVELOPPEMENT  
INTERNATIONAL

## SÉRIE ÉTUDES ET DOCUMENTS

### Environmental News Emotion and Air Pollution in China

Sébastien Marchand  
Damien Cubizol  
Elda Nasho Ah-Pine  
Huanxiu Guo

*Études et Documents* n°29  
November 2021

To cite this document:

Marchand S., Cubizol D., Nasho Ah-Pine E., Guo H. (2021) "Environmental News Emotion and Air Pollution in China", *Études et Documents*, n°29, CERDI.

CERDI  
POLE TERTIAIRE  
26 AVENUE LÉON BLUM  
F- 63000 CLERMONT FERRAND  
TEL. + 33 4 73 17 74 00  
FAX + 33 4 73 17 74 28  
<http://cerdi.uca.fr/>

## The authors

Sébastien Marchand

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

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

Damien Cubizol

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

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

Elda Nasho Ah-Pine

Associate Professor, CleRMA, ESC, Clermont Business School

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

Huanxiu Guo

Associate Professor, Institute of Economics and Finance, Nanjing Audit University

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

**Corresponding author:** Sébastien Marchand



This work was supported by the LABEX IDGM+ (ANR-10-LABX-14-01) within the program “Investissements d’Avenir” operated by the French National Research Agency (ANR).

*Études et Documents* are available online at: <https://cerdi.uca.fr/etudes-et-documents/>

Director of Publication: Grégoire Rota-Graziosi

Editor: Catherine Araujo-Bonjean

Publisher: Aurélie Goumy

ISSN: 2114 - 7957

## Disclaimer:

*Études et Documents* is a working papers series. Working Papers are not refereed, 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 2013, the Chinese central government launched a war on air pollution. As a new and major source of information, the Internet plays an important role in diffusing environmental news emotion and shaping people's perceptions and emotions regarding the pollution. How could the government make use of the environmental news emotion as an informal regulation of pollution? The paper investigates the causal relationship between web news emotion (defined by the emotional tone of web news) and air pollution (SO<sub>2</sub>, NO<sub>2</sub>, PM<sub>2.5</sub> and PM<sub>10</sub>) by exploiting the central government's war on air pollution. We combine daily monitoring data of air pollution at different levels (cities and counties, respectively the second and third administrative levels in China) with the GDELT database that allows us to have information on Chinese web news media (e.g. emotional tone of web news on air pollution). We find that a decrease of the emotional tone in web news (i.e. more negative emotions in the articles) can help to reduce air pollution at both city and county level. We attribute this effect to the context of China's war on air pollution in which the government makes use of the environmental news emotion as an informal regulation of pollution.

## **Keywords**

Air pollution, News emotion, Mass media, The Internet, Government, China

## **JEL Codes**

D83, D9, L82, O53, Q53, Q58

## **Acknowledgments**

We thank Camille Meneur and Alexis Repon for excellent research assistance. 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. We also thank seminar participants at CERDI (UCA - CNRS), Clerma (ESC Clermont), GRETHA (Bordeaux University - CNRS), Nanjing Audit University and 8th FAERE Annual Conference in Grenoble. All remaining errors are ours.

# 1 Introduction

Information is critical for decision-making in any socio-economic setting. However, the imperfect or asymmetric information in the real-world often hinders economic agents from making optimal choice and achieving market efficiency (Stigler, 1961; Grossman and Stiglitz, 1976). In the environmental setting, individuals and firms with imperfect information of pollution may under-estimate the pollution-related health impacts and under-invest in activities of self-protection and pollution abatement (Jalan and Somanathan, 2008a; Deschênes et al., 2017). Improved information access could thus help raise public awareness of pollution issues, change individuals' consumption patterns to avoid pollution exposure and increase their spending and willingness to pay (WTP) for environmental protection (Barwick et al., 2019; Greenstone et al., 2020; Ito and Zhang, 2020; Tu et al., 2020).

As a new and major source of information, the Internet plays an important role in diffusing environmental news and shaping people's perceptions and emotions regarding the pollution. In 2013, the Chinese government launched a nation-wide network of monitoring stations and data streaming system to improve the accuracy of pollution data and facilitate public access to the real-time pollution information on the Internet (Greenstone et al., 2020). Since then, mentions and reports of pollution issues have become frequent and common by traditional newspapers and new social media on the Internet. This latter has created unprecedented social network spreading environmental news and emotions regarding the pollution issues in China (Tu et al., 2020). Does the surging environmental news, in particular, the emotions disseminated in the news, alter polluter's behaviors and improve the environmental quality in China? Moreover, how could the government make use of the environmental news emotion as an informal regulation of pollution? Despite the widespread policy interests on these questions, the answer from existing empirical study is very limited. Yet, deeper understanding of these questions could shed lights on the micro foundation of the mass media and favor broader public participation in the informal regulation of pollution.

To fill in the blank of literature, this paper investigates the causal relationship between the emotions in environmental news published by online newspapers and air pollution in the context of China's war on pollution. Air pollution has always been a huge threat to the public health in China (Deschenes et al., 2020), which caused 15-25% of annual deaths. In 2013, the premier Li Keqiang declared a "war on pollution" and since then, China has undertaken comprehensive policies to combat the air pollution and the results are significant (Li et al., 2020; Liu et al., 2021). Five years after the peak in 2013, the national pollution levels of PM 2.5, SO<sub>2</sub> and NO<sub>2</sub> have fallen by 65 percent, 41 percent and 18 percent, respectively (Greenstone et al., 2021). However, the air pollution level remains high in comparison with international recommendations. For example, the national average of PM 2.5 concentration in 2018 was 40.1 ug/m<sup>3</sup>, which is more than 8 times higher than the World Health Organization guidelines (WHO, 2021). The still hard situation calls for more efforts from all stakeholders,

e.g., government, mass media, firms and households, to further improve the air quality of the country.

In this study, we collect data on the environmental news emotion from the Global Database of Events, Language and Tone (GDELT). The GDELT is an open-source database and provides daily information (theme, emotion, number of words, etc.) of news reports around the globe using technique of Natural Language Processing (NLP)<sup>1</sup>. For the pollution data, we collect daily monitoring data of air pollution for four pollutants (Nitrogen Dioxide (NO<sub>2</sub>), Sulfur Dioxide (SO<sub>2</sub>), Particulate Matter 2.5 and 10) at two different levels: from the China National Environmental Monitoring Center at city level (second level of the administrative hierarchy) and the Chinese Air Quality Reanalysis dataset at county level (third level of the administrative hierarchy)<sup>2</sup>. 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 but also 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 and its mechanisms at both city-level (286 cities in our sample) and county-level (2,408 counties in our sample).

In particular, we focus on news about air pollution reported by online newspapers in China for the year of 2018. We focus on this year to tackle potential endogeneity problem. More precisely, to address the endogeneity problem of environmental news emotion, we rely on the Chinese context of the war on air pollution launched by the central government. The hypothesis of the identification strategy is that while the mass medias response immediately to the announcement of the policies of the central government, it usually takes longer time for local government to formulate and implement the general policies in the local context. As a result, the central government policies should not directly impact the air pollution but through the catalyst of mass media in a short time span. We thus collect information of Chinese central government’s policies about air pollution from the policy inventory of the Ministry of Ecology and Environment (MEE) of China and use the announcement dates to isolate a causal impact of environmental news emotion. More precisely, we use a regression discontinuity design (RDD) in which the announcement date is used as a cutoff to estimate the local impact (around the date) of the environmental news emotion. In our study, we use as a cutoff June 27th 2018

---

<sup>1</sup>The GDELT project was created by Kaveh Leetaru, an American internet entrepreneur, academic, and senior fellow at the George Washington University, along with Philip Schrodt, a political scientist that 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) are considered as cities in our sample whereas they have the same rank as provinces. As such, these municipalities are simultaneously 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.

which is the date of the announcement (and the release on the website of the State Council) of the three-years action plan of the central government to struggle air pollution. This cutoff allows to separately test whether the “action plan” caused discontinuous changes in media news emotion and in the air pollution immediately before and after the announcement.

Our study contributes to three strands of literature. First, a growing body of literature studied the impact of the pollution information disclosure in China. For example, [Barwick et al. \(2019\)](#) documented the profound impacts of the pollution information disclosure on household awareness about pollution issues and triggered the behavioral changes in household activities. [Greenstone et al. \(2020\)](#) argued that the information program significantly improved the quality of China’s air pollution data and led to more avoidance behaviors of Chinese households. [Ito and Zhang \(2020\)](#) found that the WTP estimate based on air purifier sales was larger after the information disclosure. Second, our study also contributes to the literature of the role of policy in media ([Brady, 2009](#)). [Stockmann and Gallagher \(2011\)](#) find that the Chinese media contribute to regime legitimacy by shaping citizens’ attitudes and behavior toward the legal system in the context of labor problem and labor rights. From a content analysis of two Shanghai newspapers and one television show, they show that media representation of labor conflict often promotes the successful use of the legal system by a normal citizen which enhances government’s legitimacy through rule of law. [Wang et al. \(2014\)](#) 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 for both newspapers, almost all of the articles presented informational or pro-nuclear statements in accordance with the Chinese government’s nuclear power policies, whereas little antinuclear opinion was expressed. Third, our study is also related to the role of mass media and especially the Internet as a powerful tool for shaping perceptions and awareness of environmental issues. People’s economic decision-making is hardly influenced by information ([Eil and Rao, 2011](#); [Jacobsen, 2011](#); [Jalan and Somanathan, 2008b](#); [Madajewicz et al., 2007](#)). Mass media have a strong role as an information-based instrument. [Tu et al. \(2020\)](#) finds that the release on the Internet of the environmental documentary “Under the Dome” in 2015 significantly increases public willingness to pay for better environmental quality and raises public awareness of environmental health risk in China. [Junsheng et al. \(2019\)](#) find that mass media (TV, newspaper and the Internet) has a role on awareness, attitudes and knowledge of climate change as well as environmental friendly behavior (e.g. use of public transport, reduction of the amount of waste, ...) in Malaysia. [Happer and Philo \(2016\)](#) argue that the mass media has a powerful role to shape public understanding of climate change. They find that people can change their awareness and attitude but often with unmet behavioural 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 found on the Internet has a significant role to stimulate pro-environmental behaviors (waste reduction, recycling and energy saving). Traditional media (tv, radio and newspapers) are found to influence mainly

energy saving. The authors suggest that the explanation relies on the maturity of the topic of energy saving in old media compared to the others such as recycling.

We find consistent results regarding the impact of the environmental news emotion on air pollution reduction (at both the city and county level) within the Chinese context of the war on air pollution launched in 2013 by the central government. We find a positive effect of environmental news emotion published until three weeks before the day of observed 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 city and county level and are robust to the RDD strategy. We attribute this effect to the context of China's war on air pollution in which the government makes use of the environmental news emotion as an informal regulation of pollution. On the one hand, a decrease of news emotion contributes to reduce air pollution. This fall of news emotion suggests that emotions become less positive and more negative. This change can be perceived by individuals as a more stringent control by officials with more inspections in firms for instance. Individuals anticipate this control and reduce their pollution. Environmental news emotion have here a restriction effect. On the other hand, an increase of news emotion contributes to extend air pollution. This rise implies that news emotion become more positive and less negative. Here individuals can perceived this change as an upcoming less stringent control. As a consequence they release their effort and increase their pollution. Environmental news emotion have here a slackening effect. Overall, our results suggest that web news through the emotions they disseminate act as an informal regulation on air pollution.

The remainder of the paper is structured as follows. Section 2 describes the data, Section 3 discusses the econometric framework, and Section 4 presents the main results.

## 2 Data

We collect daily mass media information on air pollution news and link them to daily air pollution emissions at different levels. We use other various databases for control variables and the identification strategy.

### 2.1 Media Data and the GDELT project

To collect mass media information on air pollution news we use the Global Database of Events, Language, and Tone (GDELT) project (<https://www.gdelproject.org/>). The GDELT project extracts events, themes, emotions from media articles, the persons/organizations involved in, and creates the networks that tie them together. It covers web, print and broadcast news media in many languages from every country in the world, and the database is updated every 15 minutes. With such a large covering and high frequency of updating it is considered as big data, necessitating more convenient methods of processing and extraction (such as BigQuery platform).



In this study we focus on web articles in China (in English and Chinese) related to air pollution in 2018. It includes official and non-official journals (but not social media), providing information to a large audience (e.g. China Daily, South China Morning Post; 80% of the sample) or to a smaller audience (e.g. China Briefing, China Dialogue, Ecns (English website of China National News)).

We extract the following variables for all articles of our selection:

- Emotion of the article: This is our main variable of interest that measures the tone of an article. The tone is as a combination of emotional connotation (positive or negative) transmitted by an article through the words used. GDELT defines the tone of an article as “a score ranges from -100 (extremely negative) to +100 (extremely positive). Common values range between -10 and +10, with 0 indicating neutral. This is calculated as Positive Score minus Negative Score” (Source: GDELT). For example, an article explaining that air pollution has been reduced these last months and that environmental policies are a success, will have a positive tone. However, an article describing the current critical level of air pollution in a big city, will have a negative tone.
- Positive and negative scores: “This is the percentage of all words in the article that were found to have a positive (resp. negative) emotional connotation. Ranges from 0 to +100” (Source: GDELT). Positive (resp. negative) score measures the degree to which the article contains words with positive (resp. negative) connotation.
- Polarity: “This 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 polarity is high, but tone is neutral, this suggests the text was highly emotionally charged, but had roughly equivalent numbers of positively and negatively charged emotional words” (Source: GDELT). Note that an article can have a tone close to zero (the net emotional connotation of words, positive-negative) 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.
- Number of words of the article.

For a deeper understanding of these media variables, we describe below three examples of articles.

An article from China Daily (a major and official media in China) published on April 7, 2018, strengthens the improvements of air quality due to the use of more natural gas and less coal<sup>3</sup>. The tone is positive with a value of 1.1 (positive score minus negative score). The positive score is 3.3 and the negative score 2.2: the positive emotional connotations of words

<sup>3</sup>Available at: <http://www.chinadaily.com.cn/a/201804/07/WS5ac805b2a3105cdcf65168a7.html>.

(“improvement”, “enjoy”, “higher quality”, etc.) exceed the negative emotional connotations of words (“worse”, “hazardous”, etc.). The polarity (positive score plus negative score) is above 5, 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 major media in China but not official) published on April 11, 2018, alerts on ozone pollution that is worsening in Northern China<sup>4</sup>. The tone is strongly negative (-5.8) but the positive score is not null; the latter is equal to 1.8 while the negative score (7.6) makes the tone becomes negative. There are a lot of words with negative connotations (“worsened”, “hazardous”, etc.) and this text is highly emotionally charged with a polarity at 9.4 (positive and negative).

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>5</sup>. Many words with negative emotional connotations increase the negative score (“killing”, “costing”, “damage”, etc.). The few words with positive emotional connotations do not pull upward the tone that remains negative (-3.9).

For illustration we also propose word clouds bringing out articles’ words that emerge the most from important press media in China (Figure 1). We include the period 2018 with both press media in 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).

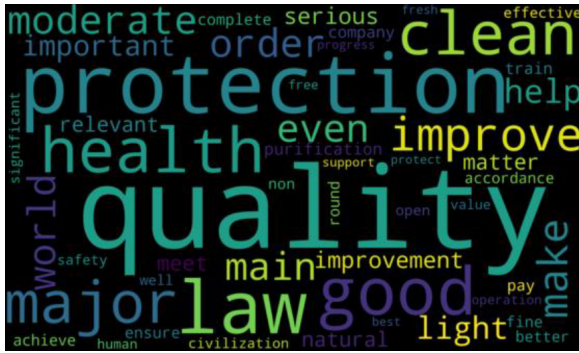
In the left-top cloud, the words with positive connotations that stand out the most are “quality”, “protection”, “health”, “improve”, “law”, “clean”, “health”, etc.; in the right-top cloud, the words with negative connotations that stand out the most are “pollution”, “severe”, “charge”, “exhaust”, “problem”, “fight”, “discharge”, “fine”, “damage”, “battle”, “waste”, etc.

Moreover, the left-bottom cloud reports words that emerge the most without considering their positive or negative connotation. The words “pollution”, “air”, “emission”, “China”, “coal”, “environmental”, “Beijing”, “protection” stand out the most. Those words suggest both the relevancy of the articles used in our analysis and the presence of many words without emotional tone like “China”, “emission”, “air”. The right-bottom cloud includes verbs and shows that “will” and “said” strongly stand out from word clouds. This result highlights the various announcements of central government plans to fight pollution.

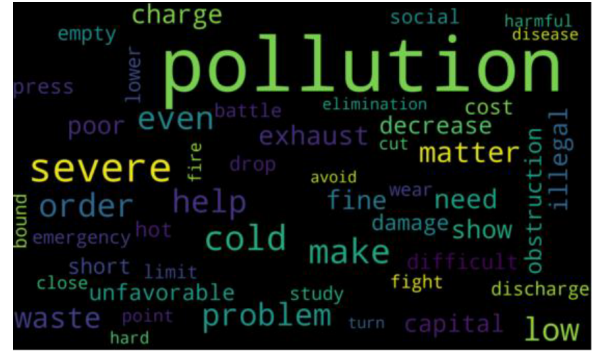
---

<sup>4</sup>Available at: <http://www.scmp.com/news/china/policies-politics/article/2141259/harmful-ozone-pollution-worsening-northern-china-says>.

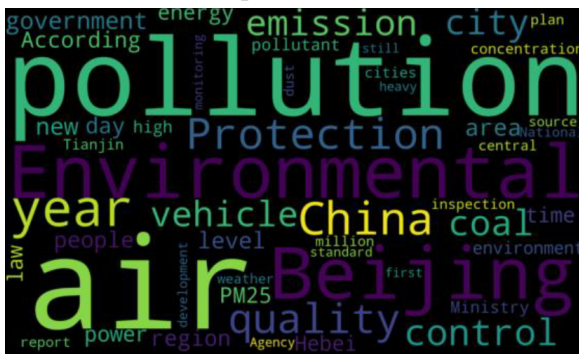
<sup>5</sup>Available at: <https://www.scmp.com/news/china/science/article/2166542/air-pollution-killing-1-million-people-and-costing-chinese>.



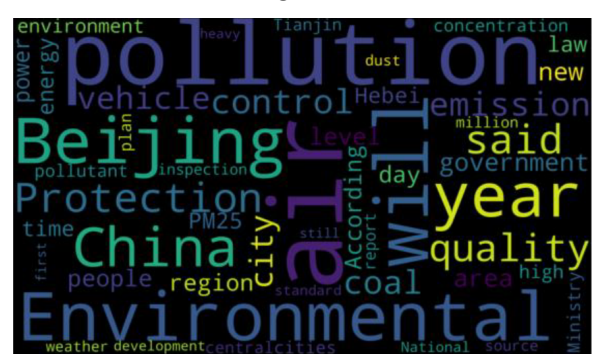
Words with positive connotation



Words with negative connotation



All words (without verbs)



All words (with verbs)

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

The following network graph (Figure 2) 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 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.

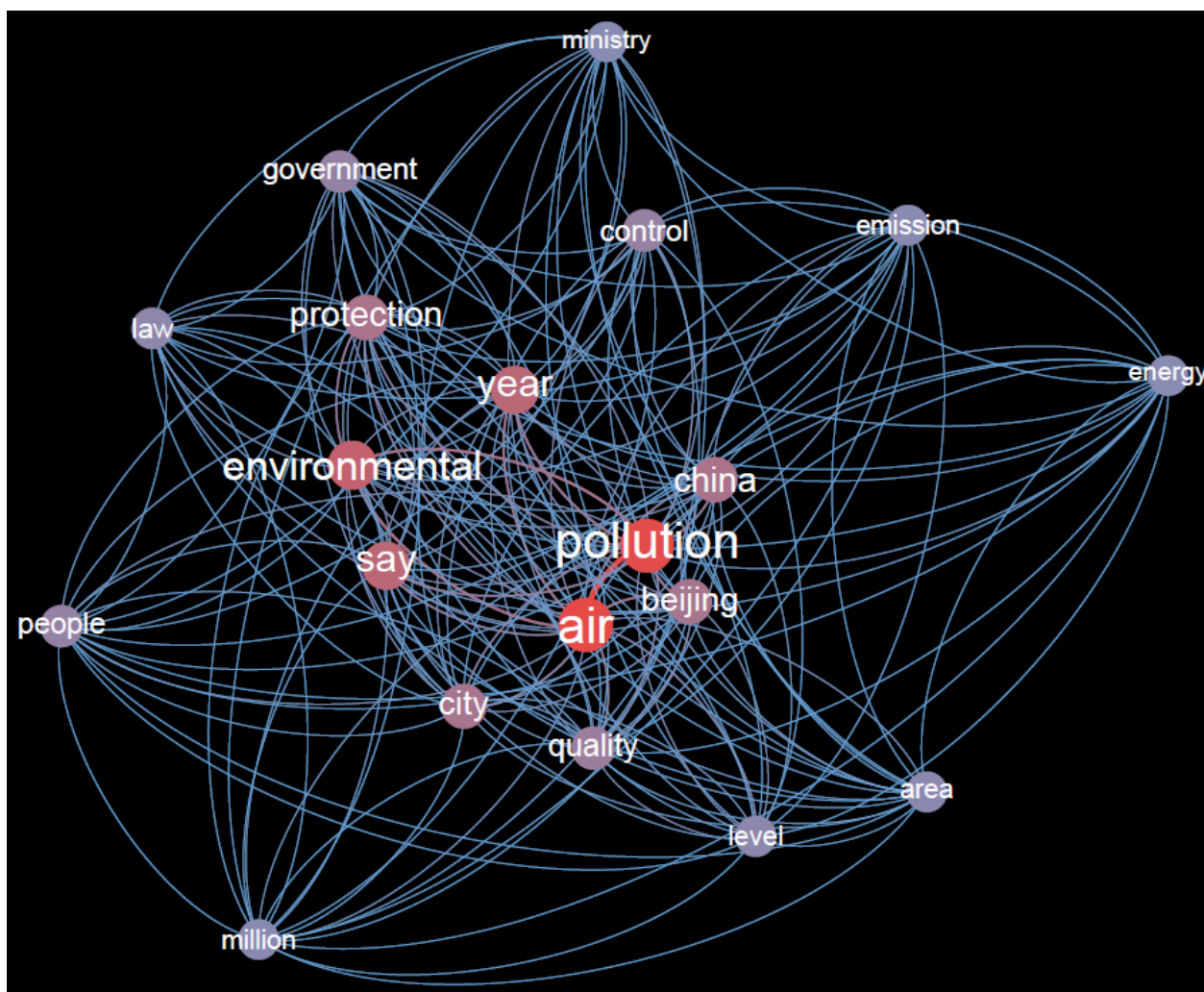


Figure 2: Network graph (most quoted words in media articles dataset)

## 2.2 Pollution Data

We use air pollution databases at county-level (third level of the administrative hierarchy) and city-level (second level of the administrative hierarchy)<sup>6</sup>:

- **City level:** for about 300 cities, we collect daily monitoring data on main air pollutants (Aqi index, PM2.5, PM10, SO2, NO2) from China National Environmental Monitoring Center (2018)<sup>7</sup>.
- **County level:** for about 2400 counties, we collect high resolution daily data on main air pollutants (PM2.5, PM10, SO2, NO2, CO and O3) from Chinese Air Quality Reanalysis dataset (Tang et al., 2020).

We focus on pollutants that are common to all databases and major contributors of air pollution: fine particles pollution (PM2.5 which are fine inhalable particles, with diameters

<sup>6</sup>A further description of the databases is available in the Appendix.

<sup>7</sup>The city level data was collected and aggregated by the Geographic Data Sharing Infrastructure, College of Urban and Environmental Science, Peking University.

that are generally 2.5 micrometers and smaller, and PM10 which are inhalable particles, with diameters that are generally 10 micrometers and smaller), Nitrogen Dioxide (NO<sub>2</sub>) and Sulfur Dioxide (SO<sub>2</sub>). These pollutants are mainly emitted by the burning of fossil fuels (coal and petroleum) and also by the smelting of iron ores for SO<sub>2</sub> from power plants, industries and automobiles.

## 2.3 Other Data

We use other data related to:

- Weather database (daily data at county and city level): day highest temperature, day lowest temperature, wind speed and an index of weather (source: [China National Environmental Monitoring Center](#)).
- The main strategical announcements of environmental policies by the Ministry of Ecology and Environment of China (MEE) during the year 2018.

## 3 The econometric model

### 3.1 The baseline model

We implement the following econometric model on our two panel datasets:

$$E_{i,p,t} = \alpha + \beta_1 \mathbf{ Tone}_{t-n} + \beta_2 \mathbf{ Polarity}_{t-n} + \beta_3 \mathbf{ Count}_{t-n} + \gamma \mathbf{ weather}_{c,t} + mFE + \epsilon_{i,t}, \quad (1)$$

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  $p$  (SO<sub>2</sub>, NO<sub>2</sub>, PM<sub>2.5</sub> and PM<sub>10</sub>). Our variable of interest is the daily media 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  $Var_{t-n}$  is not the value of the variable at  $t - n$  but the mean of the variable values until  $t - n$  ( $\text{Mean}(Var_{i,p,t-1}; \dots; Var_{i,p,t-n})$ ,  $n = 1, 7, 14, 21, 28, 35$ ).

We control for the daily weather at county and city levels: day highest/lowest temperature, weather (an index variable measuring the type of weather (rainy, sunny, etc.)) and wind speed (see Section 2 for data sources).  $mFE$  are multi-level fixed effects (ID, weekday and upper level ID\*month)<sup>8</sup>. Those multiple fixed effects enable us to control for many characteristics at

---

<sup>8</sup>We use the *reghdfe-ivreghdfe* 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).

different levels (for example in the county database, upper level ID\*month can control for an intensive use of heating in a city during a winter month). Variables are in level or logarithm (no 1st-difference because variables are stationary). For the calculation of standard errors, we cluster both at ID and upper level ID (province for city, and city for county) in order to take into account serial and spatial correlations.

We deeply analyze the potential impact of environmental news emotion on air pollution in two steps:

1. **Step 1:** our baseline model is to estimate the model with an one week average lag ( $n = 7$ ) for news emotion and other media variables only at city and country levels in order to find or not a significant effect of news emotion on air pollution while considering all potential sources of air pollution. We assume that a one week average lag for media variables is a convenient and relevant time-span.
2. **Step 2:** still at county and city level, we introduce different lags on media variables in addition to one week average lag ( $n = 1, 14, 21, 28, 35$ ); in our context, for  $n = 14, 21, 28, 35$ , it is equivalent to step by step add older news in the average of media variables while keeping the previous values (until five weeks ( $n = 35$ ) before the observed day of pollution ( $t$ )); by doing this we try to capture a memory effect of environmental news emotion on air pollution.

### 3.2 Positive and negative emotions

Positive and negative emotions do not shape the actions of individuals in the same way. In the Chinese context, individuals should be more sensitive to negative emotions transmitted by online newspapers. Regarding the stringency of the environmental policy in terms of restrictions and controls to tackle air pollution but also the huge effort to develop the awareness of detrimental impacts of air pollution on human health, individuals should more react to negative emotions than positive emotions.

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

### 3.3 The causal effect

In the baseline model, we regress air pollution on media tone by considering the latter as exogenous. However, we can assume that the tone is endogenous for several reasons. There can 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 pollution critical levels (only few observations

can be concerned by a potential impact of air pollution on tone articles). Moreover, there may be an omitted variables bias that can plague the estimation of the effect of the tone.

Therefore, to isolate a causal impact of the tone on air pollution, we implement a Regression Discontinuity Design (RDD) which has been developed 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 “three-years action plan” on June 27, 2018 as a cutoff to separately test whether the “action plan” caused discontinuous changes in media news emotional tone and in the air pollution immediately before and after the announcement.

The notice of this plan was released on the State council website by June 27th 2018. This plan proposes new targets for 2021 in terms of indicators for SO<sub>2</sub>, NO<sub>2</sub> and PM as well as the scope of key areas concerned by the plan (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 promotes the green development of the industry (e.g. the control of production capacities of some manufactured goods (steel, etc.), of the energy system and of the transportation system<sup>9</sup>.

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

$$\begin{aligned} Tone_t &= \alpha_0 + \alpha_1 Announce_t + \alpha_2 Day_t + \alpha_3 Announce_t \times Day_t + \iota + \omega + \iota \times \tau + u_{it} \\ \text{s.t. } & -h \leq Day_t \leq h \end{aligned} \quad (2)$$

$$\begin{aligned} E_{ipt} &= \alpha_0 + \alpha_1 Announce_t + \alpha_2 Day_t + \alpha_3 Announce_t \times Day_t + \iota + \omega + \iota \times \tau + u_{it} \\ \text{s.t. } & -h \leq Day_t \leq h \end{aligned} \quad (3)$$

Where  $Announce_t$  is an indicator variable of 1 for period after the announcement of “action plan”.  $Day_t$  is the distance of days within  $h$  days around the announcement. We rely on the robust CCT selector as proposed by Calonico et al. (2014) to select the optimal bandwidth  $h$ . 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 FE, weekday FE, city/county by month FE.

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

---

<sup>9</sup>For more information, see the official announcement through the release of the notice on the official website of the State Council: [http://www.gov.cn/zhengce/content/2018-07/03/content\\_5303158.htm](http://www.gov.cn/zhengce/content/2018-07/03/content_5303158.htm).

$$E_{ipt} = \beta_0 + \beta_1 \widehat{Tone}_t + Day_t + Announce_t \times Day_t + \iota + \omega + \iota \times \tau + \varepsilon_{it} \quad (4)$$

s.t.  $-h \leq Day_t \leq h$

Where  $\widehat{Tone}_t$  represents the fitted values from equation (1). In the non-parametric set up, 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. This approach is analogous to a “fuzzy RD”, and the optimal bandwidth is selected by the CCT selector as previously.

## 4 Results

### 4.1 Effect of news emotion on air pollution

We first propose to investigate the impact of news emotion at county and city levels. To do so, we focus to begin on one week average lag ( $n = 7$ ).

Tables 1 and 2 show respectively estimation results at county and city levels. In each table, we report results for both 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 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 as well as upper-level ID×months fixed effects.

We find a significant and positive effect of the tone of web news related to air pollution on the emissions of the four pollutants either at city level or county level (except for PM 2.5 at city level). More precisely, at city level (county level), an one percent increase of the tone reduces air pollutants between 0.0282% (0.0279%) and 0.0454% (0.0495%).

### 4.2 Older news emotion and memory effect

Secondly, at county and city levels, we introduce different lags on media variables in addition to one week average lag in step 1 ( $n = 1, 14, 21, 28, 35$ ). The goal is to investigate if there is a memory effect of the tone of 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 4 and 3 show the estimation results after plotting the coefficient of the tone variable according to its lag ( $n = 1, 7, 14, 21, 28, 35$ ). Also, we plot the estimation results according to the pollutant (upper graphs for NO<sub>2</sub> and S0<sub>2</sub>, lower graphs for PM<sub>2.5</sub> and PM<sub>10</sub>)<sup>10</sup>.

We find a positive effect of tone until  $n = 21$  (that is to say until three weeks before the day of observed air pollution) which becomes insignificant when we take into account older environmental web news until  $n = 35$  (that is to say until five weeks before the day of observed air pollution). This result is robust to the four pollutants at both county and city level.

<sup>10</sup>For  $n = 7$ , the plotted coefficient is the coefficient of the tone reported in Tables 1 and 2.



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

Dep. variable	(1) lnno2	(2) lnno2	(3) lnso2	(4) lnso2	(5) lnpm25	(6) lnpm25	(7) lnpm10	(8) lnpm10
lnaptone7	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)
lnappolarity7	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)
lnapcount7	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	Y	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y	Y
Weekday FE	Y	Y	Y	Y	Y	Y	Y	Y
Province*Month FE	Y	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y	Y
IV	N	N	N	N	N	N	N	N
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: Robust standard errors in parentheses ; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

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

Dep. variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	lnno2	lnno2	lnso2	lnso2	lnpm25	lnpm25	lnpm10	lnpm10
lnaptone7	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)
lnappolarity7	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)
lnapcount7	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	Y	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y	Y
Weekday FE	Y	Y	Y	Y	Y	Y	Y	Y
Province x Month FE	Y	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y	Y
IV	N	N	N	N	N	N	N	N
F-statistic	123.5	93.38	12.85	34.07	15.34	160.8	81.10	116.4
Number of county	2408	2408	2408	2408	2408	2408	2408	2408

Note: Robust standard errors in parentheses ; \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Regarding the positive effect of the tone of news during the three past weeks, we attribute this effect to the context of China's war on air pollution. More precisely, the central government makes use of the emotion of air pollution news as an informal regulation of pollution. This relationship between the government and released news on the Internet is allowed by the nation-wide network of monitoring stations and data streaming system that facilitate public access to the real-time pollution information. Thanks to that, mentions and reports of air pollution issues are more frequent and common by news paper on the Internet. In a way, the government incites the media to communicate on air pollution. Thus, web news through the emotions they disseminate (negative/neutral/positive emotional connotation) can act as an informal regulation on air pollution.

The positive effect of the tone on air pollution can be understood at the light of this context. On the one hand, a decrease of news emotion contributes to reduce air pollution. This fall of news emotion suggests that emotions become less positive and more negative. This change can be perceived by individuals as a more stringent control by officials with more inspections in firms for instance. Individuals anticipate that and reduce their pollution. Environmental news emotion act as a restriction effect. On the other hand, an increase of news emotion contributes to extend air pollution. This rise implies that news emotion become more positive and less negative. Here individuals can perceived this change as an upcoming less stringent control. As a consequence they release their effort and increase their pollution. Environmental news emotion acts here as a slackening effect.

Figure 3: Dynamic effect of tone at city level

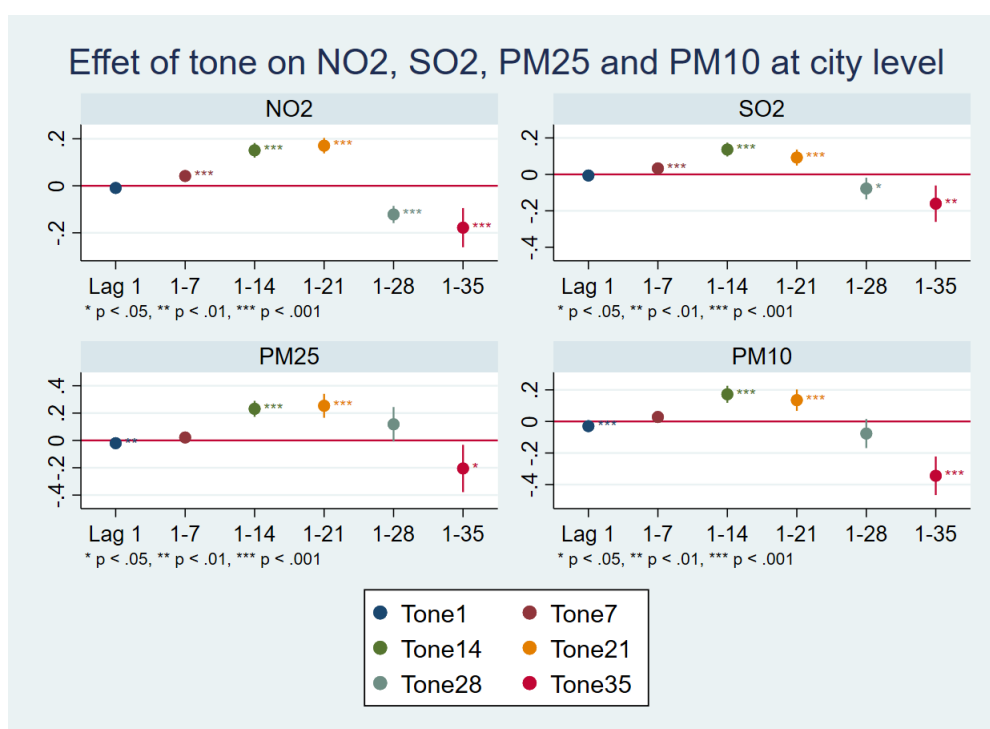
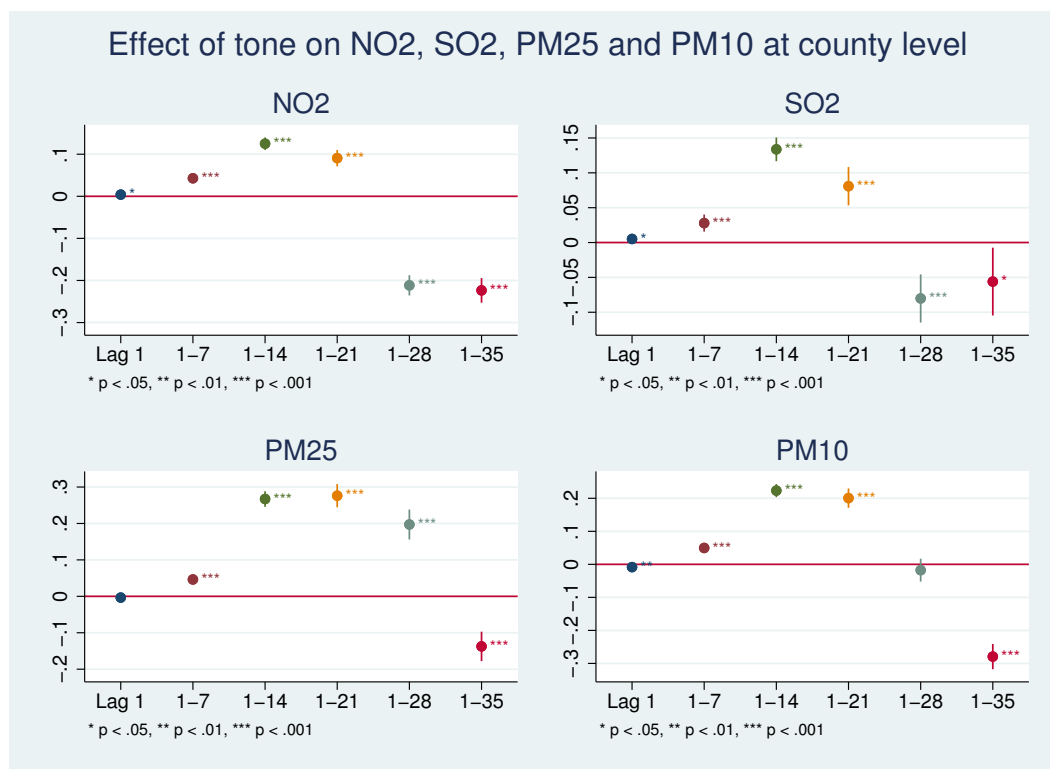


Figure 4: Dynamic effect of tone at county level



### 4.3 Positive emotions and negative emotions

To investigate this differential effect of positive emotions versus negative emotions, we replicate the specifications of Figures 4 and 3 by replacing alternatively the emotional tone by the positive score and the negative score. Those scores give 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 should 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.

Two main results are worth noting. Firstly, both positive scores and negative scores of web news published until  $n = 21$  impact air pollution in the expected ways. An increase of the positive score significantly pushes up air pollution both at city level (Figure A1) and county level (Figure A2) while an increase of the negative score significantly reduces air pollution both at city level (Figure A3) and county level (Figure A4). Those results go in the same direction than those ones 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 articles of online newspapers.

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 an higher variation of air pollution than positive emotions. For instance, at city level, while the highest effect of the positive score is ranged between 0.2 and 0.5, the strongest effect of the negative score fluctuates between -0.4 and 1. As explained, we attribute those results to the specific Chinese context made up both stringent environmental policy and strong awareness on air pollution.

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

Table 3 displays the LATE impact of tone on air quality (SO<sub>2</sub>, NO<sub>2</sub>, PM<sub>2.5</sub> and PM<sub>10</sub>) at city level (panel A) and county level (panel B). The LATE impact is estimated by using the cutoff created by the announcement on June 27th 2018 of the central government regarding the triennial strategy in terms of 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. Estimation sample excludes cities and counties targeted by the announcement of three-years action plan of the central government. 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. At both city and county level, we still find a significant and positive effect of tone while the coefficient is not significant in some specifications. More precisely, the effect tone is always positive and significant for NO<sub>2</sub>. Regarding SO<sub>2</sub>, the effect is significant only in one over four specifications (with the two CER criteria at city level). PM 2.5 and PM 10 are found to be positively impacted by tone expect for the specifications with common CER at city level.

Table 4 displays the same LATE impact of 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\*month FE (panel A) and city\*month (panel B). Except for SO<sub>2</sub> at county level, we still find a positive impact of the tone on air pollution with the two criteria used to select the optimal bandwidth.

To sum up, we attribute the positive effect of tone to the context of China's war on air pollution. The government makes use of the environmental news emotion as an informal regulation of pollution. The announcement of the three-years action plan is a set of measures related to command-and-control measures to improve the green development of the industry, transportation and energy system. Also, the plan settles on stringent thresholds in terms of reduction of air pollution. On the whole, the emotional tone disseminated in this plan and transmitted by newspapers is negative. As a consequence, the positive effect suggest that a reduction of the tone helps to reduce air pollution. Put differently, the war on air pollution is relayed by mass media through emotional tone which influences air pollution. Lastly, while we

Table 3: Non-parametric RD estimates of the (LATE) impact of the tone on air quality at city and county level - without fixed effect

VARIABLES	(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							
Tone <sub>t-1</sub>	0.142*** (0.0515)	0.982** (0.394)	0.890*** (0.257)	0.760*** (0.185)	0.225*** (0.0789)	3.571 (2.469)	-0.539 (0.346)	-0.197 (0.154)
Obs	45,230	45,230	45,230	45,230	45,230	45,230	45,230	45,230
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
Bw below c	2.745	5.692	6.028	5.623	6.553	15.62	20.96	24.17
Bw above c	12.36	12.93	8.699	6.527	6.553	15.62	20.96	24.17
	Panel B - County level							
Tone <sub>t-1</sub>	0.436*** (0.122)	-0.0126 (0.0293)	0.368*** (0.0657)	0.460*** (0.0848)	0.107** (0.0467)	0.0275 (0.0390)	0.0936* (0.0565)	0.148** (0.0593)
Obs	361,380	361,380	361,380	361,380	361,380	361,380	361,380	361,380
Nb of counties								
Cutoff - c	June 27	June 27	June 27	June 27	June 27	June 27	June 27	June 27
Bw below c	3.457	2.903	1.536	1.241	7.903	7.542	1.979	1.801
Bw above c	16.13	7.006	16.24	11.93	7.903	7.542	1.979	1.801

Note: Estimation sample excludes cities and counties targeted by the announcement of the central government whose the date is used as the cutoff. RD estimates with non-parametric local linear regressions. Conventional local linear regression robust standard errors clustered at the city (panel A)/county (panel B) level are reported in parentheses. CER = coverage error-rate. The optimal bandwidth is selected using the approach of [Calonico et al. \(2014\)](#). \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

attribute the effect of tone on air pollution to the announcement of the action plan, remember that the effect is only a local average treatment effect (LATE). Put differently, this is the effect of the tone around the date of the announcement. While the inconvenient is to estimate only a LATE of tone, the advantage is to estimate a causal impact thanks to the context of the central government's war on pollution.

Table 4: Non-parametric RD estimates of the (LATE) impact of the tone on air quality at city and county level – with fixed effect

VARIABLES	(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							
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								
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							
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								
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: Estimation sample excludes cities and counties targeted by the announcement of the central government whose the date is used as the cutoff. RD estimates with non-parametric local linear regressions. Conventional local linear regression robust standard errors clustered at the city (panel A)/county (panel B) level are reported in parentheses. CER = coverage error-rate. The optimal bandwidth is selected using the approach of [Calonico et al. \(2014\)](#). Fixed effect (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.

## 5 Conclusion

This paper investigates the causal relationship between environmental news emotion and air pollution in the context of China’s war on air pollution launched in 2013 by the central government. In this battle, the Chinese government launched a nation-wide network of monitoring stations and data streaming system to improve the accuracy of pollution data and facilitate public access to the real-time pollution information. From that moment, mentions and reports of pollution issues have become frequent and common by traditional news paper and new social media on the Internet. This latter has thus created unprecedented social network spreading environmental news and emotions regarding the pollution issues in China, and shaping people’s perceptions and emotions regarding the pollution. In this context, the central government uses the Internet and more particularly emotions disseminated in news on air pollution as an informal regulation of pollution. The goal of this paper is thus to investigate the impact of news emotion on air pollution in this context of the Chinese war on air pollution.

From daily data over the entire 2018 year on environmental news emotion and pollution data at county and city levels, we find consistent results regarding the impact of the environmental news emotion on air pollution reduction. We find a positive effect of environmental news emotion on air pollution published until three weeks before the day of observed air pollution. This results is robust to several air pollutants used in this study (NO<sub>2</sub>, SO<sub>2</sub>, PM 2.5 and PM<sub>10</sub>) both at city and county level.

We attribute this effect to the context of China’s war on air pollution in which government makes use of the environmental news emotion as an informal regulation of pollution. To confirm this story, we use a RDD to isolate a causal and local impact of news emotion on air pollution. Specifically, we exploit the official announcement of the “three-years action plan” on June 27, 2018 as a cutoff to separately test whether the “action plan” caused discontinuous changes in media news emotion and in the air pollution immediately before and after the announcement. The RDD strategy confirms the positive effect of environmental news emotion on air pollution.

The explanation of this positive effect relies on the central government’s war on pollution. On the one hand, a decrease of news emotion contributes to reduce air pollution. This fall of news emotion suggests that emotions become less positive and more negative. This change can be perceived by individuals as an upcoming more stringent control by officials with more inspections in firms for instance. Individuals anticipate that and reduce their pollution. Environmental news emotion act as a restriction effect. On the other hand, an increase of news emotion contributes to extend air pollution. This rise implies that news emotion become more positive and less negative. Here individuals can perceived this change as an upcoming less stringent control. As a consequence they release their effort and increase their pollution. Environmental news emotion play a slackening effect. Overall, in the context of the Chinese war on air pollution, our results suggest that web news through the emotion they disseminate act as an informal regulation on air pollution.



Moreover, we estimate separately the impact of positive emotions and negative emotions. We assume that positive and negative emotions should not shape the actions of people in the same way regarding the specific Chinese context. The stringency of the environmental policy and the increase of awareness of detrimental impacts of air pollution on human health should push people to more react to negative emotions than positive emotions. We first find that positive scores and negative scores of web news published until  $n = 21$  impact air pollution in the expected ways. Positive emotions significantly increase air pollution while negative emotions significantly decrease air pollution. Those results suggest that individuals react both to the positive and negative emotions transmitted by articles of online newspapers. However, we find that those two types of emotions do not influence air pollution with the same magnitude. In all regressions, negative emotions lead to an higher variation of air pollution than positive emotions. As explained, we attribute those results to the specific Chinese context made up both stringent environmental policy and strong awareness on air pollution.

To conclude, the Internet plays an important role in diffusing environmental news and disseminating emotions related to air pollution. As a consequence, the Internet contributes to shape people's perceptions and emotions regarding the pollution, and in return, to act on people's air pollution emissions. The investigation of the micro foundation of the Internet as an informal regulation of pollution is a promising issue for future researches, especially in the context of China's war on pollution.

## References

- Barwick, P. J., Li, S., Lin, L., and Zou, E. (2019). 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.
- 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.
- Deschênes, O., Greenstone, M., and Shapiro, J. S. (2017). Defensive investments and the demand for air quality: Evidence from the nox budget program. *American Economic Review*, 107(10):2958–89.
- Deschenes, O., Wang, H., Wang, S., and Zhang, P. (2020). The effect of air pollution on body weight and obesity: Evidence from china. *Journal of Development Economics*, 145:102461.
- Eil, D. and Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–38.
- 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.
- Grossman, S. J. and Stiglitz, J. E. (1976). Information and competitive price systems. *The American Economic Review*, pages 246–253.
- 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.
- 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.
- Jacobsen, G. D. (2011). The al gore effect: An inconvenient truth and voluntary carbon offsets. *Journal of Environmental Economics and Management*, 61(1):67–78.

- Jalan, J. and Somanathan, E. (2008a). The importance of being informed: Experimental evidence on demand for environmental quality. *Journal of development Economics*, 87(1):14–28.
- Jalan, J. and Somanathan, E. (2008b). The importance of being informed: Experimental evidence on demand for environmental quality. *Journal of Development Economics*, 87(1):14–28.
- Junsheng, H., 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.
- Li, P., Lu, Y., and Wang, J. (2020). The effects of fuel standards on air pollution: Evidence from china. *Journal of Development Economics*, 146:102488.
- Liu, M., Tan, R., and Zhang, B. (2021). The costs of “blue sky”: environmental regulation, technology upgrading, and labor demand in china. *Journal of Development Economics*, 150:102610.
- Madajewicz, M., Pfaff, A., van Geen, A., Graziano, J., Hussein, I., Momotaj, H., Sylvi, R., and Ahsan, H. (2007). Can information alone change behavior? response to arsenic contamination of groundwater in bangladesh. *Journal of Development Economics*, 84(2):731–754.
- Stigler, G. J. (1961). The economics of information. *Journal of political economy*, 69(3):213–225.
- Stockmann, D. and Gallagher, M. E. (2011). Remote control: How the media sustain authoritarian rule in china. *Comparative Political Studies*, 44(4):436–467.
- Tang, X., Kong, L., Zhu, J., Wang, Z., Li, J., Wu, H., Wu, Q., Chen, H., Zhu, L., Wang, W., et al. (2020). 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, 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.

## A Appendix

### A.1 Additional information - Data

We collect information on air pollution at county and city level from two different database.

#### 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., 2020](#)).
- The dataset was produced by the chemical data assimilation system (ChemDAS) developed by 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. PM<sub>2.5</sub>, PM<sub>10</sub>, SO<sub>2</sub>, NO<sub>2</sub>, CO and O<sub>3</sub>) 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

Figure A1: Positive emotions and air pollution at city level

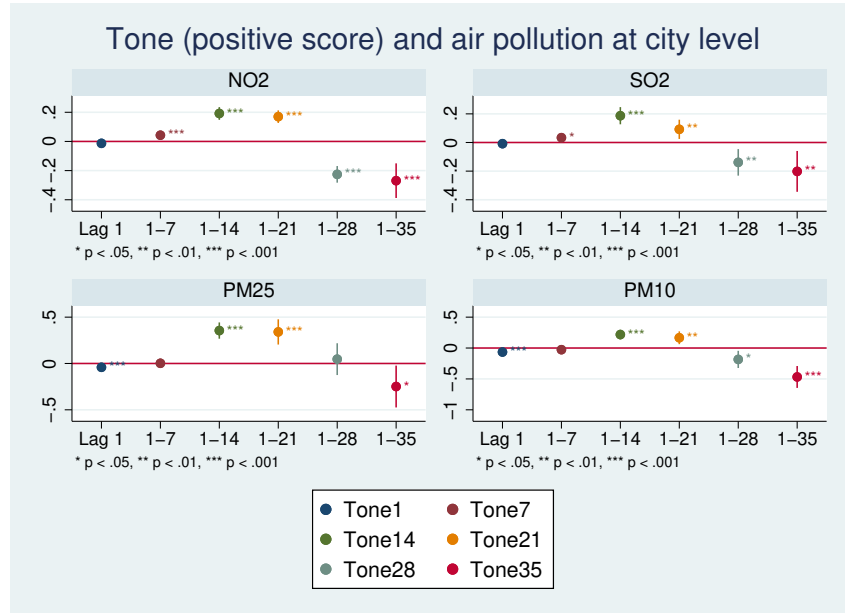


Figure A2: Negative emotions and air pollution at city level

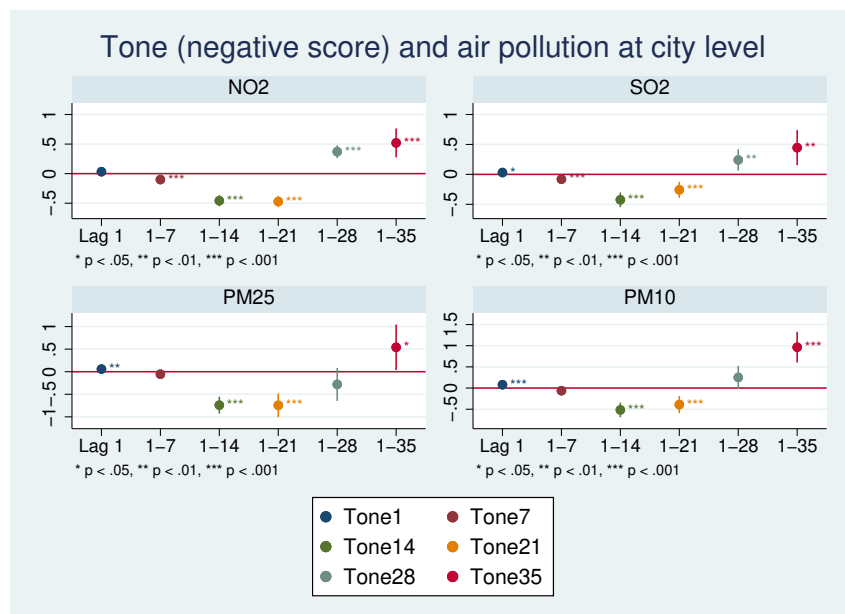


Figure A3: Positive emotions and air pollution at county level

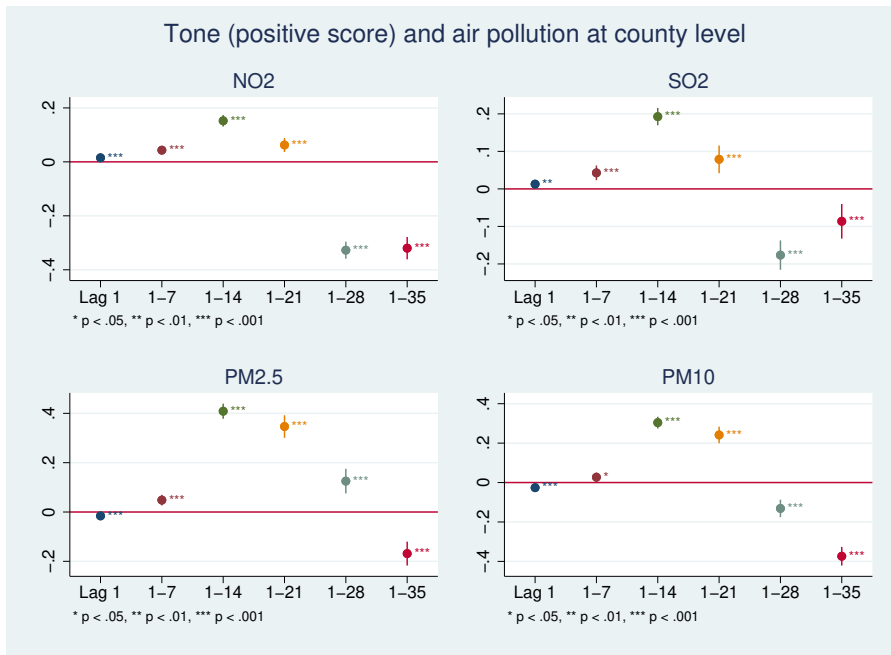


Figure A4: Negative emotions and air pollution at county level

