MUNI ECON

WORKING PAPER

n. 2025-05
ISSN 2571-130X
DOI: 10.5817/WP_MUNI_ECON_2025-05

The Effects of Air Pollution on Mood: Evidence from Twitter

Michaela Kecskésová / Department of Economics, Masaryk University, Czech Republic
Štěpán Mikula / Department of Economics, Masaryk University, Czech Republic

The Effects of Air Pollution on Mood: Evidence from Twitter

Abstract

This paper investigates the effects of air pollution on public mood using sentiment analysis of geolocated social media data. Analyzing approximately 7 million twitter posts from the United States in July 2015, we examine how fluctuations in air quality caused by Canadian wildfires influence sentiment. We find robust evidence that higher exposure to particulate matter leads to decreased positive sentiment and increased negative sentiment. Given the importance of mood as a factor in labor productivity, our results suggest that the short-term psychological effects of air pollution, alongside its well-documented physical health impacts, should be considered in policy discussions, as negative shifts in public mood due to poor air quality could have far-reaching economic consequences.

Masaryk University
Faculty of Economics and Administration

Authors:
Michaela Kecskésová / Department of Economics, Masaryk University
Štěpán Mikula / Department of Economics, Masaryk University

Contact: michaela.kecskesova@econ.muni.cz

Creation date: 2025-05
Revision date:

Keywords: air pollution, particulate matter, mood, sentiment analysis, Twitter, wildfires
JEL classification: Q5, D9

Citation:
Kecskésová, M., Mikula, Š. (2025). *The Effects of Air Pollution on Mood: Evidence from Twitter*. MUNI ECON Working Paper n. 2025-05. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2025-05

(https://creativecommons.org/licenses/by/4.0/)
Licensing of the final text published in the journal is in no way conditional on this working paper licence.

MUNI ECON Working Paper n. 2025-05
ISSN 2571-130X

The Effects of Air Pollution on Mood: Evidence from Twitter*

Michaela Kecskésová[†a] and Stepan Mikula[‡a,b]

[a]Masaryk University, Brno, Czech Republic

[b]IZA, Bonn, Germany

May 13, 2025

Abstract

This paper investigates the effects of air pollution on public mood using sentiment analysis of geolocated social media data. Analyzing approximately 7 million twitter posts from the United States in July 2015, we examine how fluctuations in air quality caused by Canadian wildfires influence sentiment. We find robust evidence that higher exposure to particulate matter leads to decreased positive sentiment and increased negative sentiment. Given the importance of mood as a factor in labor productivity, our results suggest that the short-term psychological effects of air pollution, alongside its well-documented physical health impacts, should be considered in policy discussions, as negative shifts in public mood due to poor air quality could have far-reaching economic consequences.

Keywords: air pollution; particulate matter; mood; sentiment analysis; Twitter; wildfires

JEL Codes: Q5, D9

Declaration: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

*We thank Tom Brökel, Aline Bütikofer, Nicole Wägner, and participants in seminars at the Norwegian School of Economics in Bergen for valuable feedback. The work was supported from ERDF/ESF project Ageing of the population and related challenges for health and social systems (AGEING-CZ) (No. CZ.02.01.01/00/23_025/0008743). Michaela Kecskésová acknowledges financial support provided by the Masaryk University internal grant (MUNI/A/1740/2024). The data preparation and analysis for this paper was performed as part of a collaboration with the authors of Pfeffer and Morstatter (2016). Usual caveats apply.

†Corresponding author. Department of Economics, Masaryk University, Lipová 41a, 60200 Brno, Czech Republic. Email: michaela.kecskesova@econ.muni.cz. ORCID: 0009-0006-4687-9753

‡Department of Economics, Masaryk University, Lipová 41a, 60200 Brno, Czech Republic. Email: Stepan.Mikula@econ.muni.cz. ORCID: 0000-0003-1725-8561

1 Introduction

Mood and emotions are frequently overlooked, yet crucial elements of economic behavior and decision-making. Traditional economic models often assume that individuals act rationally, basing decisions solely on logical considerations of costs and benefits. However, behavioral economists and psychologists have repeatedly challenged this assumption, demonstrating that mood and emotions significantly influence behavior, especially in situations involving risk and uncertainty. Emotional states interact with cognitive processes, shaping how decisions are made. For example, individuals in a positive mood are more likely to take risks (Stanton et al. 2014), and people in a good mood tend to be more generous and less reciprocal (Kirchsteiger, Rigotti, and Rustichini 2006). Moreover, mood is often transmitted through social interactions and can influence decision-makers across various domains (Nofsinger 2005). Lerner, Small, and Loewenstein (2004) further show that emotions such as disgust and sadness can directly influence consumer behavior, affecting both purchasing decisions and pricing strategies. There is also extensive literature in behavioral finance linking mood to trading decisions and stock market returns (Kamstra, Kramer, and Levi 2003; Nofsinger 2005). Bollen, Mao, and Zeng (2011) show that collective mood states, as measured through Twitter data, can improve predictions of changes in the Dow Jones Industrial Average (DJIA).

In addition to influencing decision-making processes, research in organizational psychology shows that better mood can enhance labor productivity in various contexts. For instance, Rothbard and Wilk (2011) show that mood influences the overall perception of work events, which can subsequently influence performance. Bellet, De Neve, and Ward (2024) show that a good mood significantly boosts sales performance for workers in a telecommunications company, largely through workers converting more calls into sales. Similarly, Oswald, Proto, and Sgroi (2015) show that positive emotions increase productivity by approximately 10–12 %, as measured by the number of accurate additions of two numbers in a lab experiment. Positive effect is also seen in the time taken by assembly line workers to complete tasks, with happier individuals processing tasks more quickly (Pakdamanian, Shiyamsunthar, and Claudio 2016).

The mechanisms through which mood influences productivity are twofold. The first mechanism is from a social point of view, specifically through increased cooperation, as happier workers are more likely to engage in voluntary cooperation and effective sharing of human capital, leading to heightened productivity (Isham, Mair, and Jackson 2020). Positive affect also contributes to individuals becoming more likable and helpful, and these positive affects tend to spread among colleagues through social influence, creating a more supportive and productive work environment (Staw, Sutton, and Pelled 1994). Second, positive emotions enhance cognitive processes, mainly by improving creative ingenuity, broadening attention, and boosting cognitive flexibility (Krekel, Ward, and De Neve 2019; Isen, Daubman, and Nowicki 1987; Amabile et al. 2005). Higher cognitive flexibility translates to higher probability of making connections between new ideas, which is, along with creativity, a crucial skill for better productivity. Coviello et al. (2024), however, stress the importance of the specific setting and the incentive scheme and argue that the positive effects do not hold for workers with a fixed wage who are not paid by performance.

However, for being such an important factor for productivity, mood is a highly subjective and very easily influenced psychological state. Sometimes it can take as little as a passing comment or a change in the weather for our mood to shift rapidly. We investigate a subtle but dangerous environmental stressor – air pollution, and its effect on public mood, measured by sentiment analysis of social media posts. Our study contributes to the relatively sparse strand of literature on the short-term psychological effects of air pollution by assessing its impact on public mood using social media data, using a similar approach to Zheng et al. 2019. We focus on particulate matter, a mixture of solid and liquid particles of various sizes, such as dust, pollen, ash, or smoke suspended in the air. These particles, especially those smaller than 2.5 micrometers ($PM_{2.5}$), pose a particular danger due to their ability to absorb various chemicals, including carcinogens, and increase their penetration and longevity in the lungs. Unlike other common air pollutants, $PM_{2.5}$ can easily penetrate indoors, meaning that going inside does not significantly reduce exposure.

As urbanization and industrialization accelerate globally, the concentration of pollutants in the atmosphere has increased significantly, posing severe risks to human health and well-being. It is estimated that air pollution alone is responsible for approximately 6 to 9 million deaths annually, surpassing the combined impact of all other known environmental risk factors on global morbidity and mortality (Landrigan 2017; Fuller et al. 2022). Among the most well-known negative effects of air pollution are shorter life expectancy (Brunekreef and Holgate 2002), cardiovascular and respiratory diseases (Dockery and Pope 1994; Hoek et al. 2013; Landrigan 2017; Newell et al. 2018; Al-Kindi et al. 2020), and impeded child development (Currie, Neidell, and Schmieder 2009; Currie et al. 2014; Greenstone and Hanna 2014). Additionally, air pollution has been shown to reduce labor productivity (Graff Zivin and Neidell 2012; Neidell 2017; Chang et al. 2016). We hypothesize that mood can work as one of the channels through which air pollution affects productivity.

From the psychological perspective, long-term exposure to high concentrations of pollutants has been associated with higher levels of psychiatric distress (Rotton and Frey 1984), headaches (Nattero and Enrico 1996), negative psychological states such as nervousness and powerlessness (Gu et al. 2020), and a heightened risk of depression (Levinson 2012; Szyszkowicz et al. 2016; Zhang, Zhang, and Chen 2017). While such long-term effects can incur substantial medical costs and potentially result in lower human capital accumulation, there are also short-term economic implications of air pollution's effects on human brain. Notably, air pollution has been shown to alter investor behavior, leading to negative trade performance and significantly lower stock market returns (Heyes, Neidell, and Saberian 2016; Huang, Xu, and Yu 2020). There is also evidence suggesting that higher levels of air pollution cause increases in violent crime (Herrnstadt et al. 2021; Bondy, Roth, and Sager 2020; Lu et al. 2020; Burkhardt et al. 2019).

Particulate matter (PM) is considered one of the most harmful pollutants, alongside nitrogen oxides, sulfur dioxide, and ground-level ozone. While there are natural sources of PM, such as wildfires, the majority of the world's PM concentrations originate from human activities, including industrial processes and fossil fuel combustion. Particulate matter is especially relevant for Central Nervous System (CNS) diseases, as it is highly toxic to lung tissue and, due to its small diameter, can cross the blood-air barrier of the lungs. Particles can then either enter the brain this way through systemic circulation, or be transported directly via the olfactory axon, leading to neuroinflammation and oxidative stress, which can contribute to more serious CNS diseases (Block and Calderón-Garcidueñas 2009; Babadjouni et al. 2017; Thomson 2019).

The exact mechanisms by which air pollution could affect behavior in a short time window are not yet fully understood, though medical research suggests a possible pathway through activation of immune cells resident in the brain, called microglia, and subsequent inflammation processes. Neuroinflammation in the prefrontal cortex, which is responsible for regulation of behavior and cognition, can result in low mood, anxiety, and delirium (Jayaraj et al. 2017; Mittli 2023). These psychological effects can in turn affect behavior, making air pollution a potential risk factor for substance abuse (Szyszkowicz et al. 2018) and suicide (Kim et al. 2010; Szyszkowicz et al. 2010).

We use satellite measurements of $PM_{2.5}$ concentrations and other environmental variables, such as temperature and precipitation, combined with approximately 7 million geolocated Twitter posts from the United States in July 2015 to assess how fluctuations in air quality influence mood. We focus specifically on the Canadian wildfire season—a natural event occurring in Canada every summer,[1] which creates an exogenous variation in air pollution over the United States. Using multiple specifications of a two-way fixed effects model, we find robust evidence that higher exposure to particulate matter leads to decreased positive sentiment and increased negative sentiment.

The rest of the paper is structured as follows. In Section 2 we describe the data sources used in this study. In Section 3 we set out our empirical strategy. Section 4 presents the main results and Section 4.1 presents additional robustness and sensitivity checks. Section 5 concludes.

2 Data

Evaluating mood on a large scale is challenging for several reasons. Mainly, methods for measuring mood, such as surveys and questionnaires, can be intrusive and time-consuming. Moreover, such surveys usually cannot be done accurately on a large scale, while they also rely on self-reporting, which can be biased or inaccurate. With this approach, it can also prove difficult to distinguish between long-term happiness and mood in real-time. Given these challenges, we turn to social media posts and natural language processing tools to infer public mood. In recent years, social media platforms have emerged as a valuable source of real-time, user-generated data that captures public sentiment and opinions. Among these platforms, Twitter stands out due to its vast user base, instantaneous nature, and location-tagged posts. Consequently, posts on Twitter can be used to track individuals' moods in real time by using natural language processing tools, such as language models trained for sentiment analysis.

2.1 Twitter data

Since its founding in 2006, Twitter has become one of the most popular social media platforms, mainly in the United States. With more than 300 million active users worldwide, it became an invaluable data source for researchers in social sciences.[2] We use a historical database of representative, geolocated tweets from the United States, collected by Pfeffer

1. Although even these naturally occurring wildfires have been getting progressively worse due to climate change.

2. However, in 2023, Twitter implemented significant changes to its data access policies and API pricing, driven primarily by concerns over data scraping and misuse of the platform's data. In addition to limits for viewing posts, Twitter significantly increased the prices for accessing its API, which made scraping a sufficient amount of tweets for research purposes inaffordable. Given these recent limitations, it has become difficult to acquire big, up-to-date samples of tweets.

and Morstatter (2016) and obtained from the GESIS Leibniz Institute for the Social Sciences. This dataset was collected specifically to serve as a reference dataset for Twitter posts that were tagged with a GPS location within the U.S. geographic bounding box. Only tweets for which users have consented to share their location information are geolocated, accounting for approximately 1 % of the total number of posts. The location information is assigned by post and can contain either exact GPS coordinates or a more general location, such as a neighborhood or a city. In our dataset, only posts with exact GPS coordinates are used. Our sample spans every day of July 2015, resulting in a total of 10 million tweets.

2.1.1 Bots

A significant part of geolocated tweets consists of automatic posts made by bots. A bot is a type of software that can autonomously perform actions such as posting, liking, following, or messaging other accounts, and does not change its location. These include traffic bots, weather bots, job openings, etc. It is obvious that posts made by such bot accounts should not be affected by air pollution via biological mechanisms. We therefore attempt to identify these posts and remove them from the main sample.[3]

Although it is not trivial to exactly determine which posts were made by bots or their exact proportion, simple rule-based detection methods can be employed to identify these accounts. Bot accounts usually post more frequently than regular users, often within short time frames. Baylis (2020) includes only users who tweet fewer than 25 times a day to reduce the amount of bots in the sample used in his analysis. We employ two approaches to detect bot accounts – identifying users who posted more than 5 posts within 5 seconds at least once, and users who posted more than 50 times in a single day. The first rule proved to identify many spam bots, while the second rule identified automated accounts such as those posting police, weather, or job updates. Sample of a few tweets that were identified as bot-generated can be found in Table 12 in Appendix A.4. A total of 3.4 millions of tweets from the sample were identified as bot-generated, leaving a sample of approximately 7 million user-generated posts.

The top panel of Figure 1 shows the distribution of tweets identified as user-generated, while the bottom panel shows the distribution of tweets that we identified as bot-generated. In the top panel, apart from cities, major U.S. highways and frequent roads are clearly visible, reflecting movement patterns of people. The bot-generated tweets exhibit a higher concentration in larger cities, corresponding to the greater presence of police and weather

3. In Section 4.1, we use these bot-generated posts for a placebo test.

stations, as well as a higher volume of job postings in these urban areas. Unlike the top panel, highways and roads are not clearly visible in this panel.

2.2 Measuring Twitter sentiment

Sentiment analysis is a set of Natural Language Processing (NLP) techniques that evaluate the emotional tone of digital text. These tools aim to determine whether the expressed sentiment in the text is positive, negative, or neutral, and often measure the intensity of these sentiments. There is a wide variety of approaches and pre-trained models for sentiment analysis. Each model is developed using different techniques, to address different linguistic features and complexities found in textual data. As such, the choice of a sentiment analysis model is largely determined by the specific requirements of the task at hand.

One specific issue in sentiment analysis of social media posts is the informal type of medium and the short length of posts. Twitter, for example, limited the length of its posts to 140 characters at the time our data sample was collected. These posts also contain textual peculiarities including emphatic uppercasing, lengthening, abbreviations, emojis, and the use of slang. They also contain a lot of noise due to the (often deliberate) use of incorrect English, sarcasm, and misspellings. This phenomenon has an impact on the overall performance of sentiment analysis models. Our model of choice for sentiment analysis is BERTweet (Nguyen, Vu, and Tuan Nguyen 2020; Pérez et al. 2021), a model based on BERT (Bidirectional Encoder Representations from Transformers), developed by researchers at Google (Alaparthi and Mishra 2020) and fine-tuned for analyzing Twitter posts. BERTweet uses a transformer architecture, which leverages attention mechanisms that weigh the influence of different words within a sentence, allowing it to understand context more deeply than models based on earlier techniques. The output of BERTweet are probabilities of the corresponding tweet being positive, negative, or neutral[4].

Figure 2 shows mean sentiment probabilities by county during July 2015, as evaluated by BERTweet. It is clearly visible that overall sentiment is much more positive than negative.

The first panel of Table 1 describes the summary statistics of unstandardized BERTweet measures on the level of individual posts. Examples of sentiment classification for several tweets can be found in Table 11 in the Appendix. We standardize each sentiment measure prior to analysis for comparability.

4. In Section 4.1, we evaluate sentiment using two other models and reestimate our equations as a sensitivity check.

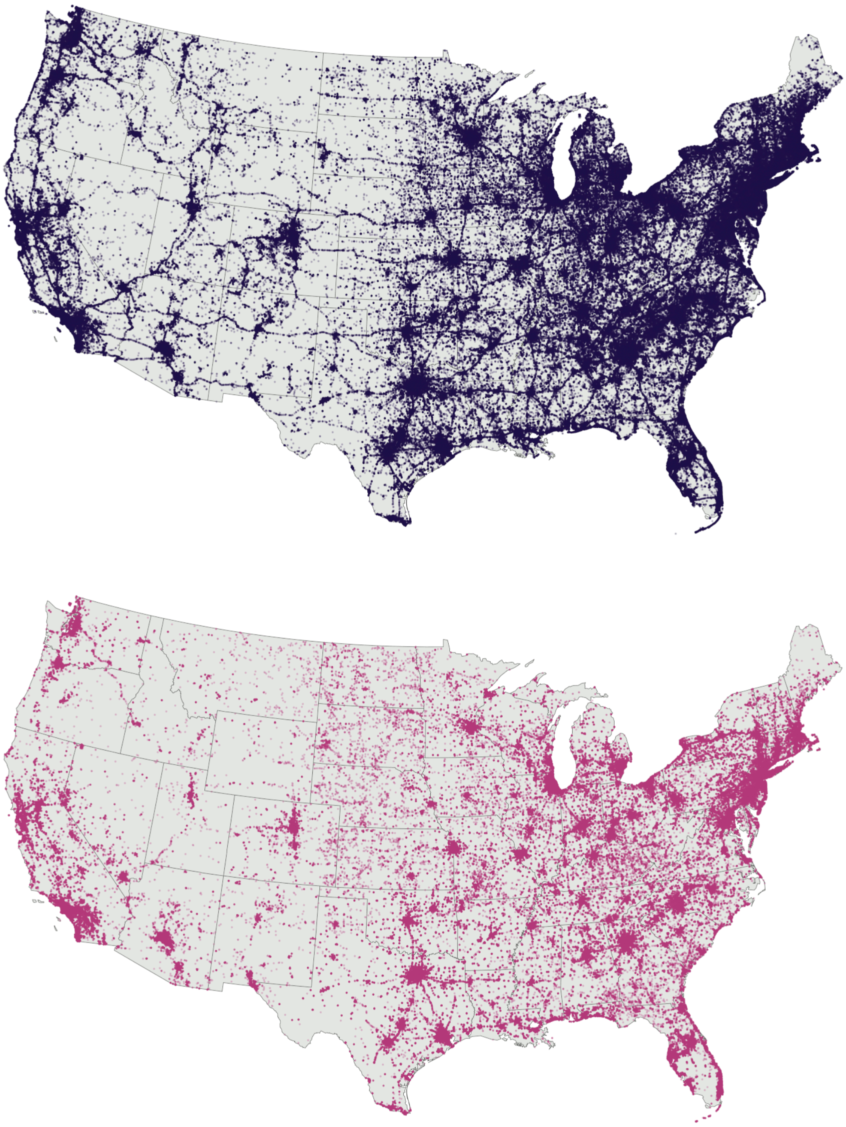

Figure 1: Spatial distribution of tweets based on GPS coordinates. Top panel: Tweets identified as user-generated. Bottom panel: Tweets identified as bot-generated.

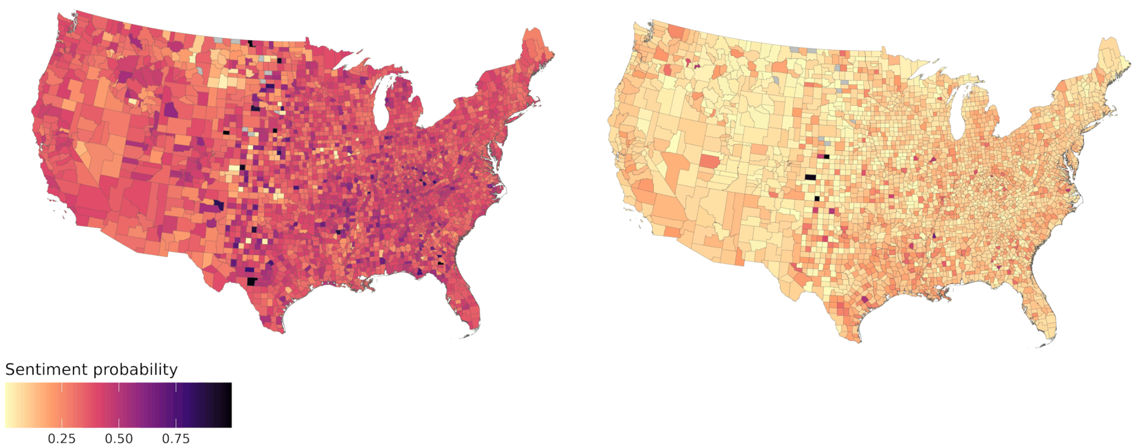

Figure 2: Mean BERTweet positive (left) and negative (right) sentiment probability by county during July 2015.

Table 1: Summary statistics.

Variable	Mean	Median	SD	Min	Max	Count
BERTweet (positive)	0.31	0.08	0.37	0.00	0.99	9,814,781
BERTweet (negative)	0.08	0.00	0.23	0.00	0.98	9,814,781
$PM_{2.5}$ ($\mu g/m^3$)	18.71	15.61	13.90	0.00	300.13	86,066
Temperature (°C)	27.08	27.26	4.43	1.66	43.33	86,066
Precipitation (mm)	1.35	0.01	4.30	0.00	69.70	86,066
Visibility (m)	30,176	29,599	7,452	78	51,544	86,066
Wind strength (m/s)	2.77	2.52	1.41	0.03	15.36	86,066
Tweets per user	12	2	152	1	34,161	816,483
Tweets per county	3,148	350	15,251	1	471,689	3,106

Notes: The first part of the table summarizes unstandardized measures of sentiment from BERTweet on the level of individual posts. Second part summarizes the variable of interest ($PM_{2.5}$) and weather controls, aggregated on a county-date level due to their large grid. Third part describes the number of posts per individual user and per county.

2.3 Pollution and weather data

To approximate local exposure to $PM_{2.5}$, we use the Copernicus Atmosphere Monitoring Service (CAMS) global reanalysis (EAC4) gridded dataset (Inness et al. 2019), provided by The European Centre for Medium-Range Weather Forecasts (ECMWF). This dataset is derived from satellite observations and has a resolution of approximately 80 km ($0.75° \times 0.75°$ lat/lon grid) with data available every 3 hours. To obtain hourly $PM_{2.5}$ exposure data, we apply a linear interpolation.

Our data spans the month of July, 2015, which coincides with the Canadian wildfire season. Naturally occurring wildfires are common in forested and grassland regions of Canada from May to September and even have beneficial effects on native vegetation, animals, and ecosystems (Tymstra et al. 2020). However, these events also produce substantial amounts of smoke that can be transported hundreds of miles by prevailing winds, occasionally reaching as far as the southern regions of the United States. Figure 3 shows elevated levels of $PM_{2.5}$ as a result of smoke originating from Canadian wildfires on 6th July, 2015 being carried through the United States. Our strategy is to exploit this additional exogenous variation to identify the effects of air pollution on public mood.[5] This approach, utilizing drifting wildfire smoke as a source of variation in air pollution levels, has been increasingly used in economics literature to assess the effects of air pollution on labor market outcomes (Borgschulte, Molitor, and Zou 2022), educational outcomes (Wen and Burke 2022), and both physical and mental health (Jayachandran 2009; Rangel and Vogl 2019; Molitor, Mullins, and White 2023; Miller, Molitor, and Zou 2024; Cabral and Dillender 2024).

While the chemical composition of wildfire smoke differs from that of anthropogenic pollution, both contain similar harmful substances, primarily $PM_{2.5}$. Notably, wildfire smoke tends to pose greater health risks due to its higher oxidative potential (Aguilera et al. 2021). The increasing frequency and severity of wildfires, driven by climate change, make this study particularly relevant, as more and more people in the United States are not only increasingly exposed to anthropogenic air pollution, but also to long-distance wildfire smoke plumes.

5. To identify the causal effects of pollution on various outcomes, an often-used strategy is to leverage wind patterns as a source of exogenous variation, as demonstrated by studies such as Zheng et al. (2019) and Graff Zivin et al. (2023). This approach exploits shifts in wind direction or strength to isolate pollution exposure from confounding factors. However, we do not adopt this strategy in our analysis for two reasons. First, wind patterns often cause changes in weather and atmospheric pressure, both of which can independently affect mood, while wind strength itself can also affect mood directly, potentially violating the exclusion restriction. Second, the variation in air pollution caused by Canadian wildfire smoke already provides a natural experiment with exogenous variation that is independent of local human activities.

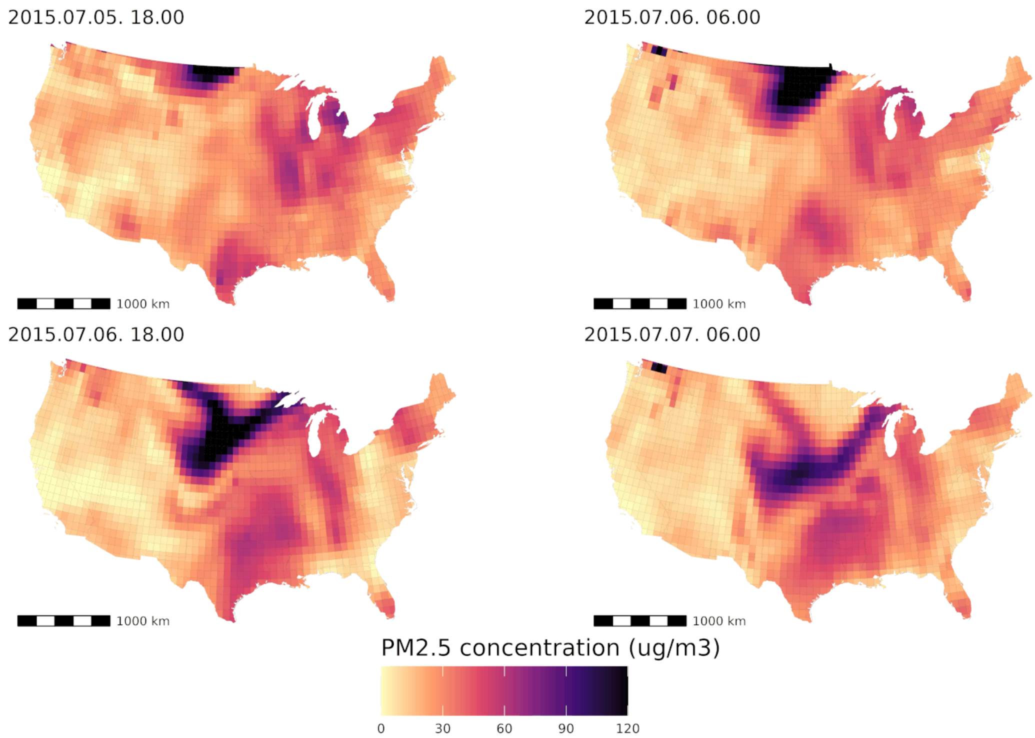

Figure 3: A cloud of wildfire smoke (purple-black) passing over U.S. at the beginning of July 2015.

The first row of Table 1 shows descriptive statistics of exposure to $PM_{2.5}$, aggregated on a county-date level. Figure 4 then shows cross-sectional variation in pollution exposure during the whole month of July 2015. Notably, only 15 out of 3,108 counties present in the dataset had mean $PM_{2.5}$ concentrations compliant with the WHO air quality guidelines for long-term exposure (WHO 2021).

Prior research also relates other environmental variables, such as temperature and precipitation, with sentiment on Twitter (Baylis et al. 2018; Zheng et al. 2019; Baylis 2020). These studies find significant relationships between temperature, precipitation and multiple sentiment measures. We therefore include both temperature and precipitation in our control variables, as well as wind strength, which has also been shown to have an effect on mood (Denissen et al. 2008).

Moreover, we also include visibility—defined here as the distance into the environment one can see unimpeded—in our controls, as visibility is another environmental factor that could potentially influence sentiment. Poor visibility, often associated with bad weather conditions such as fog, haze, or heavy pollution, may directly impact mood and sentiment by altering outdoor experiences and perceptions of safety or comfort. By controlling for visibility, we aim to isolate the specific impact of $PM_{2.5}$ on sentiment through biological mechanisms, ensuring that the observed effects are not conflated with those of visibility-related factors. Correlation between visibility and $PM_{2.5}$ exposure is only weak (-0.11).

2.4 Wildfire smoke data

To complement our main analysis, we employ a direct measure of wildfire smoke exposure to examine the link between mood and elevated air pollution levels.

This measure is derived from a dataset produced by the National Oceanic and Atmospheric Administration's (NOAA) Hazard Mapping System (HMS). The HMS compiles observations from multiple satellites operated by NOAA and the National Aeronautics and Space Administration (NASA), which generate imagery to detect wildfire smoke emissions across the United States (Ruminski et al. 2006). Each satellite typically provides imagery twice each day. HMS analysts then validate and process these satellite data, resulting in georeferenced polygons that represent the spatial extent of smoke. These polygons are then classified into three thickness categories: light, medium, and heavy.

To construct our measure of smoke exposure, shown in Figure 5, we assign numerical values to these categories (1 for light, 2 for medium, and 3 for heavy) and calculate the median smoke thickness for each county on each date.

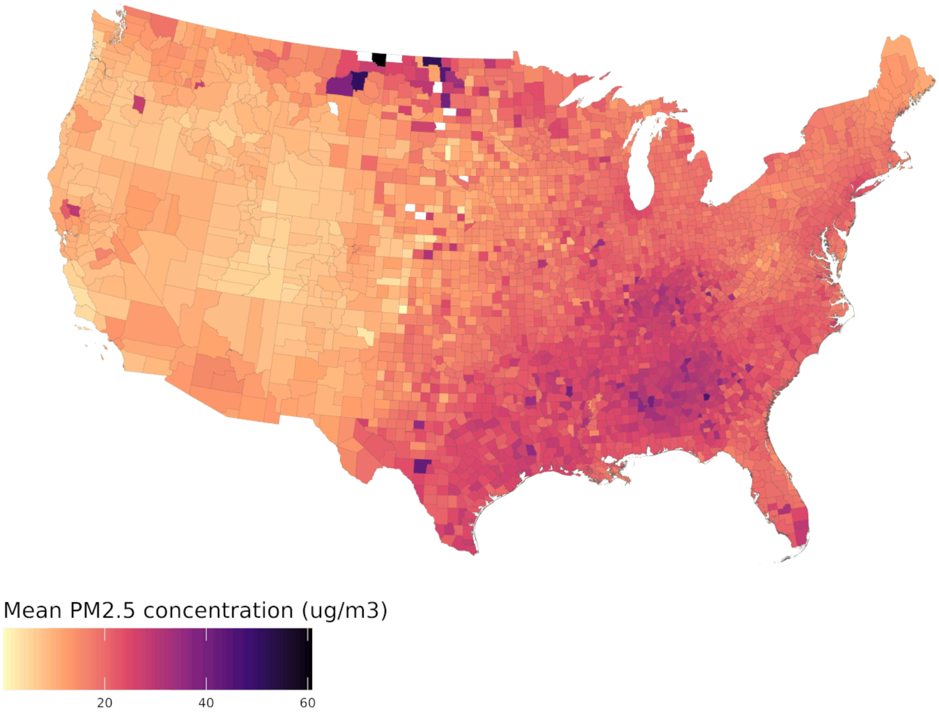

Figure 4: Mean $PM_{2.5}$ exposure by county during July 2015.

Figure 5: Wildfire smoke exposure derived from the HMS data.

3 Empirical strategy

We identify the effect of air pollution on sentiment using a panel fixed effects model. The baseline specification, with the unit of observation being a county (i) on a given date (t), is as follows:

$$\overline{S}_{it} = \alpha \overline{PM}_{2.5it} + \beta \overline{X}_{it} + \theta_i + \theta_t + \varepsilon_{it} \tag{1}$$

where $\overline{S}_{it}$ is the county-date median of a specific sentiment measure, $\overline{PM}_{2.5\text{it}}$ is the median pollution exposure in a given county on a given date, $\overline{X}_{it}$ represents weather controls including temperature and precipitation, θ_i represents county fixed effects, θ_t represents date fixed effects, and ε_{it} is the idiosyncratic error term, clustered by county and date. The regression is weighted by the number of tweets from each county. We use median as the aggregation function because of extreme values skewing the corresponding distributions.[6]

Both Zheng et al. (2019) and Baylis (2020) have documented an inverted U-shape relationship between temperature and sentiment, suggesting an "ideal" temperature, which is why we include temperature in our model as a quadratic polynomial. Given that the relationship between pollution and sentiment could also be nonlinear, we also estimate the following nonparametric specification

$$\overline{S}_{it} = f(\overline{PM}_{2.5it}) + \beta \overline{X}_{it} + \theta_i + \theta_t + \varepsilon_{it} \tag{2}$$

where $f()$ is an unknown function of air pollution. We implement $f()$ as a binned specification, where $f(\overline{PM}_{2.5it}) = \sum_{j=1}^{k} \alpha_j \overline{PM}^{j}_{2.5it}$, where $\overline{PM}^{j}_{2.5it} = 1$ if median pollution exposure in county i on date t falls in the corresponding bin j.

While our primary analysis is conducted at the county-date level, mainly due to the large grid and granularity of pollution and weather data, we are also able to estimate the effects using individual posts at a given hour as unit of observation. Since our data include username associated with each tweet, we are able to include fixed effects for the individual user on top of county and date fixed effects. This allows us to control for potential compositional effects, such as when certain types of users tweet more on days with higher pollution. We therefore also estimate the following equation:

6. As a sensitivity check in Section 4.1, we use mean instead of median as an aggregation function and show that the results hold.

$$S_{it} = \alpha PM_{2.5it} + \beta X_{it} + \theta_i + \theta_t + \theta_{\text{user}} + \varepsilon_{it} \tag{3}$$

with the unit of observation now being an individual tweet (i) at a given hour (t), and where θ_i now includes county and user fixed effects. Moreover, when using individual-level data aggregated by hour, there is a possibility that mood fluctuations throughout the day may confound the results. For instance, individuals may exhibit lower mood levels in the morning before work and higher levels in the evening, which could coincide with variations in pollution levels, particularly during morning rush hours when air quality tends to worsen. To account for this potential confounder, θ_t now also includes hour-of-day fixed effects, thereby controlling for systematic mood variations over the course of the day and ensuring that the observed effects are not driven by time-of-day fluctuations.

The TweetNLP library (Camacho-Collados et al. 2022), which provides the RoBERTa-base model API, also provides a pre-trained model for classification of tweets into 19 available topics.[7] This further enables us to control for topic fixed effects and thus for variations in sentiment that are driven by the specific content or subject matter of the tweet.

We standardized all variables prior to analysis to have mean zero and standard deviation equal to one. In the baseline specification (1), estimate of coefficient α therefore represent expected changes in conditional mean of the given sentiment measure, measured in standard deviations, as a result of one standard deviation (approx. 14 μg/m^3) increase in pollution exposure. In specification (2), coefficients α_j represent expected changes in conditional mean of the given sentiment measure, also measured in standard deviations, as a result of replacing a day with the lowest median possible pollution exposure (0–5 μg/m^3) with a day of median pollution exposure in bin j.

3.1 Instrumental variables estimation

As an exogenous source of air pollution that is independent of human activities, drifting wildfire smoke can serve as an instrument for air pollution levels. In our baseline equation (1), if $PM_{2.5}$ is endogenous, the parameter of interest, α, which represents the effect of air pollution on sentiment, will be estimated with bias. To address this potential endogeneity

7. The 19 topics are: "arts & culture", "business & enterpreneurs", "celebrity & pop culture", "diaries & daily life", "family", "fashion & style" "film tv & video", "fitness & health", "food & dining", "gaming", "learning & educational", "music", "news & social concern", "other hobbies", "relationships", "science & technology", "sports", "travel & adventure", and "youth & student life", or in case of no suitable topic, a separate category labeled "no topic", see Antypas et al. (2022).

and provide further evidence for a causal relationship, we complement the main analysis with an instrumental variable (IV) estimation strategy, exploiting the quasiexperimental variation in air pollution generated by wildfire smoke.

The first stage of this IV estimation is represented by the following equation:

$$\overline{PM}_{2.5it} = \alpha\overline{Smoke}_{it} + \beta\overline{X}_{it} + \theta_i + \theta_t + \varepsilon_{it} \tag{4}$$

Here, we use our constructed measure of wildfire smoke exposure, $\overline{Smoke}_{it}$, as an instrument for air pollution exposure, estimating the effect with the standard two-stage least squares (2SLS) method. A key identifying assumption for the IV approach is that the exclusion restriction holds. Specifically, we assume that, conditional on control variables and fixed effects, wildfire smoke affects both positive and negative sentiment only through its impact on air pollution levels.

A similar IV approach has been used to estimate the effect of air pollution on labor market outcomes (Borgschulte, Molitor, and Zou 2022), suicides (Molitor, Mullins, and White 2023), and health (Miller, Molitor, and Zou 2024).

4 Results

The first two columns of Table 2 show the results of our baseline specification (1). We find statistically significant decline in positive sentiment probability and a corresponding increase in negative sentiment probability resulting from higher exposure to particulate matter. One standard deviation (approximately $14\mu g/m^3$) increase in $PM_{2.5}$ results in a 0.05 standard deviations (6 %) decrease in positive sentiment probability and a 0.06 standard deviation (14.4 %) increase in negative sentiment probability.

The second panel of Table 2 shows that the individual effects are substantially smaller in magnitude, but their direction and significance generally still hold, even after including user, hour-of-day and topic fixed effects. For an individual, a one standard deviation increase in $PM_{2.5}$ concentration results in a 0.002 standard deviations decrease in positive sentiment probability and a slightly smaller increase in negative sentiment probability.

Figure 6 documents the results of estimating Equation (2). The red-colored line shows the negative sentiment probability response, while the green-colored line shows the posi-

Table 2: Baseline results.

	County-date (1)		Individual posts (3)			
	BERTweet (positive)	BERTweet (negative)	BERTweet (positive)	BERTweet (negative)	BERTweet (positive)	BERTweet (negative)
$PM_{2.5}$	-0.0451^{***}	0.0536^{***}	-0.0020^{**}	0.0017^{**}	-0.0020^{**}	0.0011^{*}
	(0.0116)	(0.0147)	(0.0009)	(0.0006)	(0.0008)	(0.0006)
Temperature	−0.0299	0.0116	-0.0040^{**}	-0.0046^{***}	−0.0018	0.0018
	(0.0252)	(0.0208)	(0.0015)	(0.0010)	(0.0014)	(0.0012)
Temperature^2	0.0007	0.0072	-0.0013^{**}	0.0030^{***}	−0.0001	0.0020^{***}
	(0.0103)	(0.0121)	(0.0006)	(0.0004)	(0.0008)	(0.0004)
Precipitation	-0.0372^{**}	0.0307^{*}	-0.0017^{**}	0.0006	-0.0011^{**}	0.0008^{*}
	(0.0164)	(0.0165)	(0.0007)	(0.0004)	(0.0005)	(0.0004)
Visibility	0.0559^{***}	-0.0559^{***}	0.0025^{***}	-0.0017^{**}	0.0026^{***}	-0.0023^{***}
	(0.0100)	(0.0112)	(0.0008)	(0.0007)	(0.0008)	(0.0006)
Wind strength	0.0226	−0.0282	0.0021^{*}	-0.0012^{**}	0.0016	−0.0004
	(0.0177)	(0.0228)	(0.0011)	(0.0005)	(0.0011)	(0.0005)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE					Yes	Yes
User FE			Yes	Yes	Yes	Yes
Topic FE					Yes	Yes
Observations	81,909	81,909	6,724,301	6,724,301	6,724,301	6,724,301
R^2	0.32	0.34	0.37	0.30	0.38	0.31

Notes: Estimates of regressions (1) and (3), with standard errors clustered by county and date reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

tive sentiment probability response. The relationship between both positive and negative sentiment measures and $PM_{2.5}$ appears to be approximately linear.[8]

Table 3 shows the results of the 2SLS estimation. The first column of Panel A reports the first stage estimate, which shows that increasing smoke exposure indeed also increases $PM_{2.5}$ levels at the county-date level, despite being a very noisy measure of actual particulate matter concentrations. Panel B shows that the effect of $PM_{2.5}$—instrumented by smoke exposure—remains discernible and negative for positive sentiment probability, confirming that elevated pollution levels reduce expressions of positive mood. The effect on negative sentiment probability is positive, but no longer statistically significant in this setting.

Table 3: First stage, reduced form, and IV estimates.

	$PM_{2.5}$	BERTweet (positive)	BERTweet (negative)
A. First stage and reduced form estimates			
Smoke exposure	0.1702***	−0.0403**	0.0181
	(0.0433)	(0.0152)	(0.0166)
Observations	73,580	73,580	73,580
F(1, 73,573)	2,407.8		
B. IV estimates			
$\hat{PM}_{2.5}$	-	−0.2367**	0.1060
	-	(0.0977)	(0.0940)
Observations	-	73,580	73,580

Notes: IV estimates of the effect of air pollution on sentiment, using wildfire smoke exposure measure described in Section 2.4 as an instrument. The unit of observation is a county on a given date. The regression includes the same controls and fixed effects as (1) and is weighted by the number of posts in each county. Standard errors are clustered by county and date and reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

To put the sizes of these effects in context, we employ a similar approach to Baylis (2020), where we compare the changes in sentiment caused by elevated pollution levels to the average changes in sentiment on different days of the week. A significant variation in sentiment on different days of the week has been documented before by Dodds et al. (2011) and Baylis

8. Figure 10 in Appendix A.2 documents the same responses when using RoBERTa and VADER as sentiment measures. Table 10 in the Appendix A.2 shows the respective estimated coefficients for all three sentiment measures, along with their standard errors.

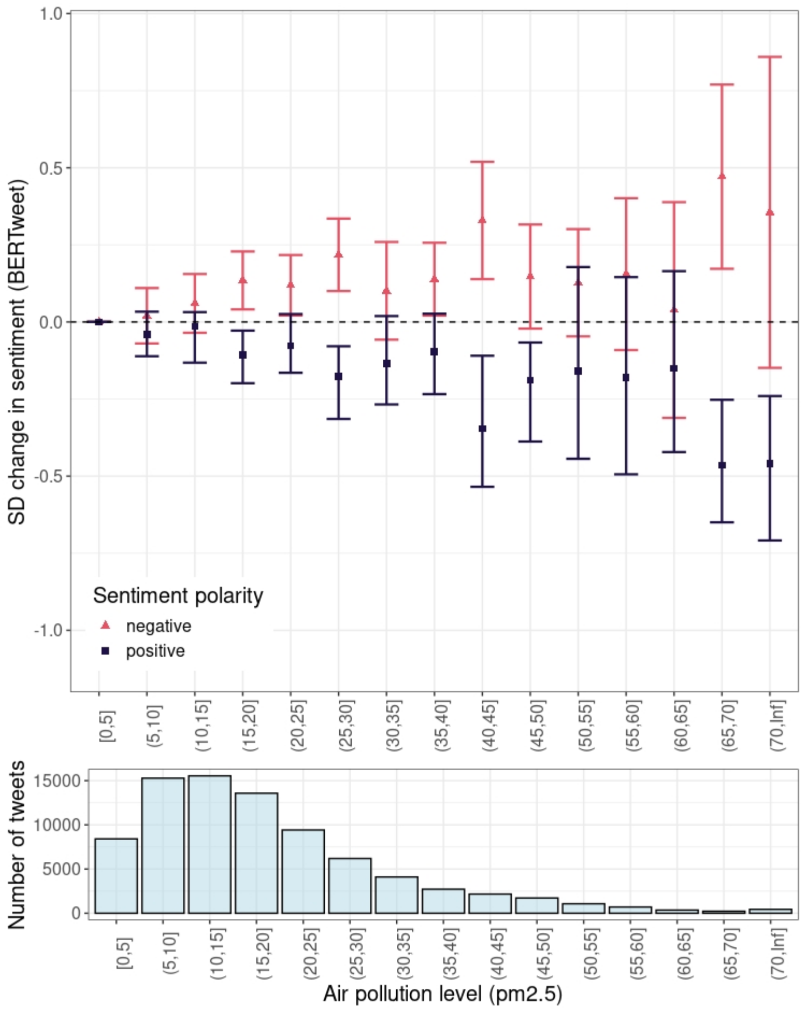

Figure 6: Visualization of results of estimating Equation (2) on sentiment at the county-date level with 95 % confidence intervals. Red color: Negative sentiment probability response. Blue color: Positive sentiment probability response. Standard errors are clustered by county and date.

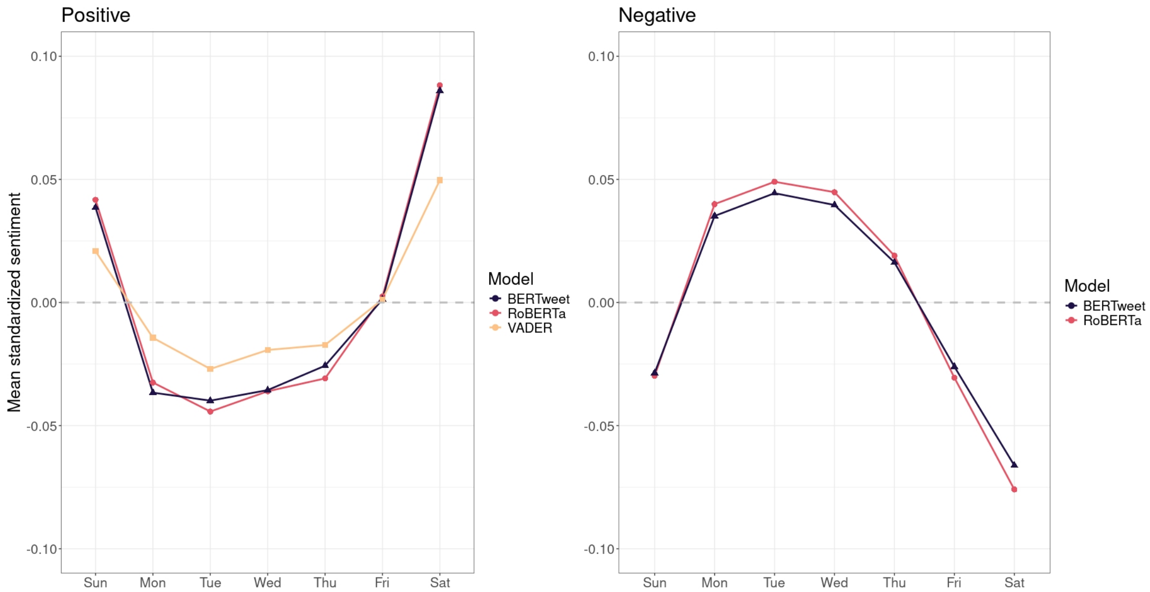

Figure 7: Mean standardized sentiment probability during different days of week. Left panel: VADER, RoBERTa and BERTweet positive measures. Right panel: RoBERTa and BERTweet negative measures.

(2020), showing a higher positive sentiment on weekends than during weekdays, with Saturday being the most positive day.

Figure 7 shows the average standardized sentiment measures by day of week. The left panel shows positive measures together with VADER, while the right panel shows negative measures. Our findings clearly match previous work. The average difference in the probability of positive sentiment between Sunday and Monday is between 0.04 standard deviations and 0.07 standard deviations in all three measures, which is roughly the same as the difference caused by one standard deviation increase (14 $\mu g/m^3$) in air pollution levels. The average difference in the probability of negative sentiment between Sunday and Monday is approximately 0.06 standard deviations, which is again roughly the same effect as one standard deviation increase in pollution exposure.

4.1 Robustness and sensitivity analysis

In this section, we conduct a series of robustness and sensitivity checks to support our main findings.

While smoke from Canadian wildfires provides an exogenous source of variation in air pollution, domestic wildfires within the United States may introduce confounding factors. Such events can impact public mood through channels beyond $PM_{2.5}$, including direct threats to safety, displacement, and emotional distress. In July 2015, several wildfire incidents occurred in the United States, with the majority concentrated in California and Washington.[9] To address this concern, we perform a robustness check by excluding all counties in California and Washington from our sample and re-estimating Equation (1). The first panel of Table 4 shows that the effects are only slightly smaller in magnitude, but in the baseline specification, their overall direction and significance hold, despite the fact that tweets from California and Washington together comprise almost 20 % of our overall sample. When we move to individual-level data, however, we start lacking the power to detect such modest effects.

In our baseline specification, we used median as the aggregation function. We find median to be a more accurate representation of all variables, be it environmental variables, or sentiment probability, as it better captures the "typical" situation. However, as the second panel of Table 4 shows, the results hold even when using mean instead of median as the aggregation function. We also re-estimate the baseline equation after excluding tweets that contain pollution-related terms. Specifically, we remove any tweets that mention words such as *pollution, air quality, CO_2, particulate matter, dust, ash, haze*, and other similar terms[10] that could indicate an explicit discussion of pollution or environmental conditions. This ensures that our results are not driven by individuals tweeting directly about pollution or its effects. The third panel of Table 4 shows that excluding these tweets has virtually no effect on our baseline results.

Furthermore, we evaluate sentiment using two additional models: the Valence Aware Dictionary for sEntiment Reasoning (VADER), a rule-based and lexicon-based model developed by Hutto and Gilbert (2014) and a RoBERTa-base model. The RoBERTa-base model is based on the same architecture as BERTweet and is trained on approximately 124 millions of tweets, fine-tuned for sentiment analysis (Liu et al. 2019; Loureiro et al. 2022). Although BERTweet is our primary model of interest due to its more sophisticated architecture and

9. See https://www.fire.ca.gov/incidents/2015 and https://www.ncei.noaa.gov/access/monitoring/monthly-report/fire/201507.

10. For a full list of the excluded terms, see Appendix A.5.

Table 4: Robustness checks.

	Excluding CA and WA		Mean instead of median		Excl. pollution and wildfire terms	
	BERTweet (positive)	BERTweet (negative)	BERTweet (positive)	BERTweet (negative)	BERTweet (positive)	BERTweet (negative)
$PM_{2.5}$	-0.0435^{***}	0.0466^{**}	-0.0428^{***}	0.0506^{***}	-0.0477^{***}	0.0534^{***}
	(0.0137)	(0.0174)	(0.0111)	(0.0143)	(0.0122)	(0.0151)
Temperature	−0.0395	0.0152	−0.0357	0.0289	−0.0316	0.0122
	(0.0246)	(0.0237)	(0.0249)	(0.0211)	(0.0248)	(0.0199)
Temperature^2	−0.0066	0.0192	−0.0030	0.0025	0.0006	0.0070
	(0.0132)	(0.0132)	(0.0093)	(0.0129)	(0.0101)	(0.0121)
Precipitation	−0.0123	0.0075	-0.0463^{**}	0.0478^{***}	-0.0371^{**}	0.0319^{*}
	(0.0141)	(0.0125)	(0.0185)	(0.0172)	(0.0169)	(0.0178)
Visibility	0.0623^{***}	-0.0539^{***}	0.0491^{***}	-0.0496^{***}	0.0553^{***}	-0.0571^{***}
	(0.0158)	(0.0157)	(0.0112)	(0.0129)	(0.0102)	(0.0114)
Wind strength	−0.0036	0.0072	0.0228	−0.0273	0.0083	−0.0138
	(0.0121)	(0.0133)	(0.0172)	(0.0220)	(0.0138)	(0.0164)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	79,027	79,027	81,913	81,913	81,626	81,626
R^2	0.30	0.31	0.32	0.34	0.32	0.34

Notes: Estimates of regression (1). First panel: All California (CA) and Washington (WA) counties are excluded. Second panel: Using mean instead of median as aggregation function. Third panel: Estimates using a subsample of data excluding pollution-related terms. Standard errors are clustered by county and date reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

specific tuning for Twitter data, these alternative models provide further validation of our results. Detailed descriptions of these models, along with their correlations with BERTweet measures (Table 8), as well as descriptive statistics (9) can be found in Appendix A.1.

As shown in Table 5, the results using the RoBERTa-base model remain consistent across all specifications. For VADER, being a much noisier and simpler sentiment measure, only the baseline specification remains significant on a 10 % level (Table 6).

As the last part of our robustness analysis, we conduct a placebo test using a sample of automated bot accounts (see Section 2.1.1 and Table 12 in the Appendix). Bots typically generate content that is not influenced by external factors such as air pollution, as their posts are either pre-programmed or algorithmically generated without reflecting genuine emotions.[11] Even though our strategy for identifying bot accounts was limited and the resulting sample includes false positives, we would not anticipate to see a significant relationship between air pollution and the sentiment expressed in these posts. The results displayed in Table 7 show no significant effect of air pollution on bot-generated sentiment, confirming that the observed relationship in the primary analysis is unlikely to be driven by spurious correlations or other unobserved factors unrelated to genuine emotional responses. Visibility, which was a highly significant factor in the primary analysis, remains significant for positive sentiment probability. Apart from our bot-identifying strategy not being perfect, this could be attributed to the fact that weather bots announcing sunny weather are assigned more positive sentiment than those announcing bad weather.

11. There is, however, evidence for the number of traffic accidents increasing with higher levels of $PM_{2.5}$ (Shi et al. 2022; Ahmadi, Khorsandi, and Mesbah 2021) Tweets by local police bots could therefore have higher negative sentiment probability in areas with higher levels of air pollution.

Table 5: Baseline results using RoBERTa as sentiment measure.

	County-date (1)		Individual posts (3)			
	RoBERTa (positive)	RoBERTa (negative)	RoBERTa (positive)	RoBERTa (negative)	RoBERTa (positive)	RoBERTa (negative)
$PM_{2.5}$	-0.0490^{***}	0.0400^{**}	-0.0018^{**}	0.0023^{***}	-0.0019^{**}	0.0016^{**}
	(0.0141)	(0.0153)	(0.0008)	(0.0006)	(0.0008)	(0.0006)
Temperature	−0.0275	0.0050	-0.0038^{**}	-0.0053^{***}	−0.0022	−0.0018
	(0.0192)	(0.0194)	(0.0016)	(0.0010)	(0.0016)	(0.0011)
Temperature^2	−0.0160	0.0196^{*}	-0.0015^{**}	0.0031^{***}	−0.0004	0.0019^{***}
	(0.0102)	(0.0096)	(0.0006)	(0.0004)	(0.0007)	(0.0005)
Precipitation	-0.0359^{**}	0.0169	-0.0013^{**}	0.0009^{**}	-0.0007^{*}	0.0011^{**}
	(0.0143)	(0.0121)	(0.0006)	(0.0004)	(0.0004)	(0.0004)
Visibility	0.0589^{***}	-0.0475^{***}	0.0035^{***}	-0.0028^{***}	0.0038^{***}	-0.0036^{***}
	(0.0146)	(0.0131)	(0.0008)	(0.0007)	(0.0008)	(0.0006)
Wind strength	0.0185	−0.0189	-0.0020^{*}	−0.0006	−0.0016	0.0003
	(0.0173)	(0.0188)	(0.0010)	(0.0006)	(0.0010)	(0.0006)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Hour FE					Yes	Yes
User FE			Yes	Yes	Yes	Yes
Topic FE					Yes	Yes
Observations	81,909	81,909	6,724,301	6,724,301	6,724,301	6,724,301
R^2	0.36	0.39	0.35	0.35	0.36	0.36

Notes: Estimates of regressions (1) and (3) using a RoBERTa-base model, with standard errors clustered by county and date reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6: Baseline results using VADER as sentiment measure.

	County-date	Individual posts	
	VADER (compound)	VADER (compound)	VADER (compound)
$PM_{2.5}$	-0.0432^{*}	-0.0006	-0.0008
	(0.0215)	(0.0008)	(0.0008)
Temperature	-0.0213	-0.0042^{***}	-0.0015
	(0.0175)	(0.0014)	(0.0012)
Temperature^2	0.0081	-0.0011^{*}	-0.0005
	(0.0153)	(0.0006)	(0.0007)
Precipitation	-0.0051	-0.0004	-0.0003
	(0.0152)	(0.0006)	(0.0005)
Visibility	0.0226	0.0029^{***}	0.0025^{***}
	(0.0133)	(0.0007)	(0.0007)
Wind strength	0.0099	-0.0020^{**}	-0.0015^{*}
	(0.0097)	(0.0009)	(0.0009)
County FE	Yes	Yes	Yes
Date FE	Yes	Yes	Yes
Hour FE			Yes
User FE		Yes	Yes
Topic FE			Yes
Observations	80,097	6,724,301	6,724,301
R^2	0.08	0.26	0.26

Notes: Estimates of regressions (1) and (3) using VADER, with standard errors clustered by county and date reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 7: Placebo test.

	Placebo test	
	BERTweet (positive)	BERTweet (negative)
$PM_{2.5}$	-0.0186	0.0093
	(0.0175)	(0.0215)
Temperature	-0.0289	-0.0797^{*}
	(0.0284)	(0.0324)
Temperature^2	0.0089	0.0217
	(0.0149)	(0.0167)
Precip	-0.00175	-0.0215
	(0.0184)	(0.0240)
Visibility	0.0479^{**}	-0.0146
	(0.0209)	(0.0401)
Wind strength	0.0148	-0.0381
	(0.0281)	(0.0288)
County FE	Yes	Yes
Date FE	Yes	Yes
Observations	67,590	67,590
R^2	0.08	0.06

Notes: Estimates of regression (1) using a sample of bot-generated tweets. Standard errors are clustered by county and date reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5 Conclusion

This study demonstrates that elevated levels of fine particulate matter have a statistically significant and also economically meaningful effect on public mood, reducing positive sentiment and increasing negative sentiment across U.S. counties. These effects are robust across various sentiment analysis measures, model specifications, and hold when controlling for weather variables, time trends, and county-specific factors. The findings align with those of Zheng et al. 2019 based on Chinese data, demonstrating that the observed effects of $PM_{2.5}$ are also significant in countries like the United States, where average pollution levels are significantly lower than in China.

Despite the robustness of our results, the study is not without its limitations. The first potential limitation is the representativeness of the population in the Twitter data, and thus external validity of the results. Social media users, particularly Twitter users, may not be fully representative of the general population, as certain groups, such as older adults, children, low-income individuals, and those without internet access, are less likely to use these platforms. Some of these groups, however, are also among those most vulnerable to air pollution. As a result, the observed effects may underestimate the true impact of air pollution on mood.

Additionally, the granularity of the environmental data presents another limitation. The environmental variables were available on an 80×80 km grid, and hourly values used in the individual-level specification were approximated using a linear interpolation algorithm. This level of granularity may obscure localized variations in pollution exposure, which could affect precision of the estimated relationships. Furthermore, sentiment analysis models, though widely used and more and more sophisticated, are not perfect, as they sometimes struggle to accurately capture language nuances such as sarcasm or irony, which could lead to misclassification of sentiment in certain cases.

Despite these limitations, the results suggest that air pollution can have broader societal implications through mood as a mediating channel, such as influencing consumer behavior, reducing labor productivity, and affecting decision-making processes. We contribute to the growing body of literature concerning psychological and societal effects of air pollution, which should be incorporated in policy discussions, as they can contribute to broader economic challenges.

References

Aguilera, Rosana, Thomas Corringham, Alexander Gershunov, and Tarik Benmarhnia. 2021. "Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California." *Nature communications* 12 (1): 1493.

Ahmadi, Maryam, Babak Khorsandi, and Mahmoud Mesbah. 2021. "The effect of air pollution on drivers' safety performance." *Environmental Science and Pollution Research* 28 (13): 15768–15781.

Alaparthi, Shivaji, and Manit Mishra. 2020. "Bidirectional Encoder Representations from Transformers (BERT): A sentiment analysis odyssey." *arXiv preprint arXiv:2007.01127.*

Amabile, Teresa M, Sigal G Barsade, Jennifer S Mueller, and Barry M Staw. 2005. "Affect and creativity at work." *Administrative science quarterly* 50 (3): 367–403.

Antypas, Dimosthenis, Asahi Ushio, Jose Camacho-Collados, Leonardo Neves, Vitor Silva, and Francesco Barbieri. 2022. "Twitter topic classification." *arXiv preprint arXiv:2209.09824.*

Babadjouni, Robin M, Drew M Hodis, Ryan Radwanski, Ramon Durazo, Arati Patel, Qinghai Liu, and William J Mack. 2017. "Clinical effects of air pollution on the central nervous system; a review." *Journal of Clinical Neuroscience* 43:16–24.

Baylis, Patrick. 2020. "Temperature and temperament: Evidence from Twitter." *Journal of Public Economics* 184:104161.

Baylis, Patrick, Nick Obradovich, Yury Kryvasheyeu, Haohui Chen, Lorenzo Coviello, Esteban Moro, Manuel Cebrian, and James H Fowler. 2018. "Weather impacts expressed sentiment." *PloS one* 13 (4): e0195750.

Bellet, Clement S, Jan-Emmanuel De Neve, and George Ward. 2024. "Does employee happiness have an impact on productivity?" *Management science* 70 (3): 1656–1679.

Block, Michelle L, and Lilian Calderón-Garcidueñas. 2009. "Air pollution: mechanisms of neuroinflammation and CNS disease." *Trends in neurosciences* 32 (9): 506–516.

Bollen, Johan, Huina Mao, and Xiaojun Zeng. 2011. "Twitter mood predicts the stock market." *Journal of computational science* 2 (1): 1–8.

Bondy, Malvina, Sefi Roth, and Lutz Sager. 2020. "Crime is in the air: The contemporaneous relationship between air pollution and crime." *Journal of the Association of Environmental and Resource Economists* 7 (3): 555–585.

Borgschulte, Mark, David Molitor, and Eric Yongchen Zou. 2022. "Air pollution and the labor market: Evidence from wildfire smoke." *Review of Economics and Statistics,* 1–46.

Brunekreef, Bert, and Stephen T Holgate. 2002. "Air pollution and health." *The lancet* 360 (9341): 1233–1242.

Burkhardt, Jesse, Jude Bayham, Ander Wilson, Ellison Carter, Jesse D Berman, Katelyn O'Dell, Bonne Ford, Emily V Fischer, and Jeffrey R Pierce. 2019. "The effect of pollution on crime: Evidence from data on particulate matter and ozone." *Journal of Environmental Economics and Management* 98:102267.

Cabral, Marika, and Marcus Dillender. 2024. *Air Pollution, Wildfire Smoke, and Worker Health.* Technical report. National Bureau of Economic Research.

Camacho-Collados, Jose, Kiamehr Rezaee, Talayeh Riahi, Asahi Ushio, Daniel Loureiro, Dimosthenis Antypas, Joanne Boisson, Luis Espinosa-Anke, Fangyu Liu, Eugenio Martínez-Cámara, et al. 2022. "TweetNLP: Cutting-edge natural language processing for social media." *arXiv preprint arXiv:2206.14774.*

Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2016. "Particulate Pollution and the Productivity of Pear Packers." *American Economic Journal: Economic Policy* 8, no. 3 (August): 141–69. https://doi.org/10.1257/pol.20150085.

Coviello, Decio, Erika Deserranno, Nicola Persico, and Paola Sapienza. 2024. "Effect of mood and worker incentives on workplace productivity." *The Journal of Law, Economics, and Organization* 40 (2): 362–393.

Currie, Janet, Matthew Neidell, and Johannes F. Schmieder. 2009. "Air pollution and infant health: Lessons from New Jersey." *Journal of Health Economics* 28, no. 3 (May): 688–703.

Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell. 2014. "What Do We Know About Short- and Long-Term Effects of Early-Life Exposure to Pollution?" *Annual Review of Resource Economics* 6 (1): 217–247. https://doi.org/10.1146/annurev-resource-100913-012610. eprint: https://doi.org/10.1146/annurev-resource-100913-012610.

Denissen, Jaap JA, Ligaya Butalid, Lars Penke, and Marcel AG Van Aken. 2008. "The effects of weather on daily mood: a multilevel approach." *Emotion* 8 (5): 662.

Dockery, Douglas W, and C Arden Pope. 1994. "Acute respiratory effects of particulate air pollution." *Annual review of public health* 15 (1): 107–132.

Dodds, Peter Sheridan, Kameron Decker Harris, Isabel M Kloumann, Catherine A Bliss, and Christopher M Danforth. 2011. "Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter." *PloS one* 6 (12): e26752.

Fuller, Richard, Philip J Landrigan, Kalpana Balakrishnan, Glynda Bathan, Stephan Bose-O'Reilly, Michael Brauer, Jack Caravanos, Tom Chiles, Aaron Cohen, Lilian Corra, et al. 2022. "Pollution and health: a progress update." *The Lancet Planetary Health* 6 (6): e535–e547.

Graff Zivin, Joshua, and Matthew Neidell. 2012. "The Impact of Pollution on Worker Productivity." *American Economic Review* 102, no. 7 (December): 3652–73. https://doi.org/10.1257/aer.102.7.3652.

Graff Zivin, Joshua, Matthew Neidell, Nicholas J Sanders, and Gregor Singer. 2023. "When externalities collide: Influenza and pollution." *American Economic Journal: Applied Economics* 15 (2): 320–351.

Greenstone, Michael, and Rema Hanna. 2014. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." *American Economic Review* 104, no. 10 (October): 3038–72. https://doi.org/10.1257/aer.104.10.3038.

Gu, Hejun, Weiran Yan, Ehsan Elahi, and Yuxia Cao. 2020. "Air pollution risks human mental health: an implication of two-stages least squares estimation of interaction effects." *Environmental Science and Pollution Research* 27:2036–2043.

Herrnstadt, Evan, Anthony Heyes, Erich Muehlegger, and Soodeh Saberian. 2021. "Air pollution and criminal activity: Microgeographic evidence from Chicago." *American Economic Journal: Applied Economics* 13 (4): 70–100.

Heyes, Anthony, Matthew Neidell, and Soodeh Saberian. 2016. *The effect of air pollution on investor behavior: Evidence from the S&P 500.* Technical report. National Bureau of Economic Research.

Hoek, Gerard, Ranjini M Krishnan, Rob Beelen, Annette Peters, Bart Ostro, Bert Brunekreef, and Joel D Kaufman. 2013. "Long-term air pollution exposure and cardio-respiratory mortality: a review." *Environmental health* 12 (1): 1–16.

Huang, Jiekun, Nianhang Xu, and Honghai Yu. 2020. "Pollution and performance: Do investors make worse trades on hazy days?" *Management Science* 66 (10): 4455–4476.

Hutto, Clayton, and Eric Gilbert. 2014. "Vader: A parsimonious rule-based model for sentiment analysis of social media text." In *Proceedings of the international AAAI conference on web and social media,* 8:216–225. 1.

Inness, Antje, M. Ades, A. Agustí-Panareda, J. Barré, A. Benedictow, A.-M. Blechschmidt, J.J. Dominguez, et al. 2019. "CAMS global reanalysis (EAC4). Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store (ADS)." Accessed: 2023-11-13. https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-reanalysis-eac4-monthly?tab=overview.

Isen, Alice M, Kimberly A Daubman, and Gary P Nowicki. 1987. "Positive affect facilitates creative problem solving." *Journal of personality and social psychology* 52 (6): 1122.

Isham, Amy, Simon Mair, and Tim Jackson. 2020. "Wellbeing and productivity: a review of the literature." *CUSP Working Paper Series.*

Jayachandran, Seema. 2009. "Air quality and early-life mortality: Evidence from Indonesia's wildfires." *Journal of Human resources* 44 (4): 916–954.

Jayaraj, Richard L, Eric A Rodriguez, Yi Wang, and Michelle L Block. 2017. "Outdoor ambient air pollution and neurodegenerative diseases: the neuroinflammation hypothesis." *Current environmental health reports* 4:166–179.

Kamstra, Mark J, Lisa A Kramer, and Maurice D Levi. 2003. "Winter blues: A SAD stock market cycle." *American economic review* 93 (1): 324–343.

Kim, Changsoo, Sang Hyuk Jung, Dae Ryong Kang, Hyeon Chang Kim, Ki Tae Moon, Nam Wook Hur, Dong Chun Shin, and Il Suh. 2010. "Ambient particulate matter as a risk factor for suicide." *American journal of psychiatry* 167 (9): 1100–1107.

Al-Kindi, Sadeer G, Robert D Brook, Shyam Biswal, and Sanjay Rajagopalan. 2020. "Environmental determinants of cardiovascular disease: lessons learned from air pollution." *Nature Reviews Cardiology* 17 (10): 656–672.

Kirchsteiger, Georg, Luca Rigotti, and Aldo Rustichini. 2006. "Your morals might be your moods." *Journal of Economic Behavior & Organization* 59 (2): 155–172.

Krekel, Christian, George Ward, and Jan-Emmanuel De Neve. 2019. "Employee wellbeing, productivity, and firm performance." *Saïd Business School WP* 4.

Landrigan, Philip J. 2017. "Air pollution and health." *The Lancet Public Health* 2 (1): e4–e5.

Lerner, Jennifer S, Deborah A Small, and George Loewenstein. 2004. "Heart strings and purse strings: Carryover effects of emotions on economic decisions." *Psychological science* 15 (5): 337–341.

Levinson, Arik. 2012. "Valuing public goods using happiness data: The case of air quality." *Journal of Public Economics* 96 (9-10): 869–880.

Liu, Yinhan, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. "Roberta: A robustly optimized bert pretraining approach." *arXiv preprint arXiv:1907.11692.*

Loureiro, Daniel, Francesco Barbieri, Leonardo Neves, Luis Espinosa Anke, and Jose Camacho-Collados. 2022. "TimeLMs: Diachronic language models from Twitter." *arXiv preprint arXiv:2202.03829.*

Lu, Jackson G., Julia J. Lee, Francesca Gino, and Adam D. Galinsky. 2020. "Air Pollution, State Anxiety, and Unethical Behavior: A Meta-Analytic Review." PMID: 32453616, *Psychological Science* 31 (6): 748–755. https://doi.org/10.1177/0956797620924765. eprint: https://doi.org/10.1177/0956797620924765.

Miller, Nolan H, David Molitor, and Eric Zou. 2024. *The Nonlinear Effects of Air Pollution on Health: Evidence from Wildfire Smoke.* Technical report. National Bureau of Economic Research.

Mittli, Dániel. 2023. "Inflammatory processes in the prefrontal cortex induced by systemic immune challenge: focusing on neurons." *Brain, Behavior, & Immunity-Health* 34:100703.

Molitor, David, Jamie T Mullins, and Corey White. 2023. "Air pollution and suicide in rural and urban America: Evidence from wildfire smoke." *Proceedings of the National Academy of Sciences* 120 (38): e2221621120.

Nattero, Giovanni, and Annalisa Enrico. 1996. "Outdoor pollution and headache." *Headache: The Journal of Head and Face Pain* 36 (4): 243–245.

Neidell, Matthew. 2017. "Air pollution and worker productivity." *IZA World of Labor* (January). https://doi.org/10.15185/izawol.363.

Newell, Katherine, Christiana Kartsonaki, Kin Bong Hubert Lam, and Om Kurmi. 2018. "Cardiorespiratory health effects of gaseous ambient air pollution exposure in low and middle income countries: A systematic review and meta-analysis." *Environmental Health* 17 (April). https://doi.org/10.1186/s12940-018-0380-3.

Nguyen, Dat Quoc, Thanh Vu, and Anh Tuan Nguyen. 2020. "BERTweet: A pre-trained language model for English Tweets." In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations,* edited by Qun Liu and David Schlangen, 9–14. Online: Association for Computational Linguistics, October. https://doi.org/10.18653/v1/2020.emnlp-demos.2.

Nofsinger, John R. 2005. "Social mood and financial economics." *The Journal of Behavioral Finance* 6 (3): 144–160.

Oswald, Andrew J, Eugenio Proto, and Daniel Sgroi. 2015. "Happiness and productivity." *Journal of labor economics* 33 (4): 789–822.

Pakdamanian, Erfan, Niroshni Shiyamsunthar, and David Claudio. 2016. "Simulating the effect of workers' mood on the productivity of assembly lines." In *2016 Winter Simulation Conference (WSC),* 3440–3451. IEEE.

Pérez, Juan Manuel, Mariela Rajngewerc, Juan Carlos Giudici, Damián A Furman, Franco Luque, Laura Alonso Alemany, and María Vanina Martínez. 2021. "pysentimiento: A python toolkit for opinion mining and social nlp tasks." *arXiv preprint arXiv:2106.09462.*

Pfeffer, Jürgen, and Fred Morstatter. 2016. *Geotagged Twitter posts from the United States: A tweet collection to investigate representativeness.* Data File Version 1.0.0. https://doi.org/10.7802/1166.

Rangel, Marcos A, and Tom S Vogl. 2019. "Agricultural fires and health at birth." *Review of Economics and Statistics* 101 (4): 616–630.

Rothbard, Nancy P, and Steffanie L Wilk. 2011. "Waking up on the right or wrong side of the bed: Start-of-workday mood, work events, employee affect, and performance." *Academy of Management Journal* 54 (5): 959–980.

Rotton, James, and James Frey. 1984. "Psychological costs of air pollution: Atmospheric conditions, seasonal trends, and psychiatric emergencies." *Population and Environment* 7 (1): 3–16.

Ruminski, Mark, Shobha Kondragunta, Roland Draxler, and Jian Zeng. 2006. "Recent changes to the hazard mapping system." In *Proceedings of the 15th International Emission Inventory Conference,* 15:18. National Oceanic and Atmospheric Administration.

Shi, Daqian, Ziwei Liu, Jie Fu, and Hongwei Yu. 2022. "The impact of drivers' short-term exposure to air pollution on traffic deaths." *Environmental Science and Pollution Research* 29 (40): 61323–61333.

Stanton, Steven J, Crystal Reeck, Scott A Huettel, and Kevin S LaBar. 2014. "Effects of induced moods on economic choices." *Judgment and Decision making* 9 (2): 167–175.

Staw, Barry M, Robert I Sutton, and Lisa H Pelled. 1994. "Employee positive emotion and favorable outcomes at the workplace." *Organization science* 5 (1): 51–71.

Szyszkowicz, Mieczysław, Termeh Kousha, Mila Kingsbury, and Ian Colman. 2016. "Air pollution and emergency department visits for depression: a multicity case-crossover study." *Environmental health insights* 10:EHI–S40493.

Szyszkowicz, Mieczysław, Errol M Thomson, Ian Colman, and Brian H Rowe. 2018. "Ambient air pollution exposure and emergency department visits for substance abuse." *PLoS One* 13 (6): e0199826.

Szyszkowicz, Mieczysław, Jeff B Willey, Eric Grafstein, Brian H Rowe, and Ian Colman. 2010. "Air pollution and emergency department visits for suicide attempts in Vancouver, Canada." *Environmental health insights* 4:EHI–S5662.

Thomson, Errol M. 2019. "Air pollution, stress, and allostatic load: linking systemic and central nervous system impacts." *Journal of Alzheimer's Disease* 69 (3): 597–614.

Tymstra, Cordy, Brian J Stocks, Xinli Cai, and Mike D Flannigan. 2020. "Wildfire management in Canada: Review, challenges and opportunities." *Progress in Disaster Science* 5:100045.

Wen, Jeff, and Marshall Burke. 2022. "Lower test scores from wildfire smoke exposure." *Nature Sustainability* 5 (11): 947–955.

WHO. 2021. *WHO global air quality guidelines: particulate matter (PM2. 5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide.* World Health Organization.

Zhang, Xin, Xiaobo Zhang, and Xi Chen. 2017. "Happiness in the air: How does a dirty sky affect mental health and subjective well-being?" *Journal of environmental economics and management* 85:81–94.

Zheng, Siqi, Jianghao Wang, Cong Sun, Xiaonan Zhang, and Matthew E Kahn. 2019. "Air pollution lowers Chinese urbanites' expressed happiness on social media." *Nature human behaviour* 3 (3): 237–243.

A Appendix

A.1 Alternative sentiment measures

The first alternative sentiment analysis model is Valence Aware Dictionary for sEntiment Reasoning (VADER), a rule-based and lexicon-based model developed by Hutto and Gilbert (2014). It has a set of predefined words which are labeled by humans as either positive or negative. Each word in the lexicon is rated for its sentiment on a scale from -4 (extremely negative) to $+4$ (extremely positive). VADER is specifically attuned to sentiments expressed in social media and was developed to be sensitive both to polarity (positive/negative) and intensity (strength) of emotion. The result of applying VADER to a piece of text is a *compound score*, which is a metric that calculates the sum of all the lexicon ratings which have been normalized between -1 (most extreme negative) and $+1$ (most extreme positive). Figure 8 shows mean VADER scores for each county during July 2015.

The second alternative is a RoBERTa-base model trained on approximately 124 millions of tweets and fine-tuned for sentiment analysis (Liu et al. 2019; Loureiro et al. 2022). This model is, same as BERTweet, based on BERT transformer architecture, and fine-tuned for analyzing Twitter posts. The main difference between BERTweet and RoBERTa-base lies in the specific optimizations and data on which they were fine-tuned. While RoBERTa-base has been trained on more general data initially and then adapted for Twitter, BERTweet is trained directly on Twitter data from the start.

A.2 Nonlinear effects of $PM_{2.5}$ on sentiment

Table 10 shows estimates of Equation (2) corresponding to Figure 6.

Table 8: Correlation between sentiment measures. RoBERTa-base and BERTweet are strongly positively correlated in both positive and negative dimensions. The compound score of VADER is positively correlated with positive metrics of RoBERTa-base and BERTweet and negatively correlated with their negative counterparts, although the relationship is weaker. This can be attributed to the fundamental differences in how these models evaluate sentiment (VADER is rule-based and lexicon-based, while the transformer models can evaluate sentiment from the entire sentence structure), and also to the different metrics they produce. VADER's compound score is a normalized, weighted composite score of all tokens in the text, which can dilute the impact of any single sentiment expression. If a text contains words with mixed sentiments, it can lead to a moderate overall score, even if there are strong positive or negative cues that would be picked up by the more context-sensitive BERT-based models.

Measure	VADER (compound)	RoBERTa (positive)	RoBERTa (negative)	BERTweet (positive)	BERTweet (negative)
Vader (compound)	1.00	0.59	-0.38	0.53	-0.35
RoBERTa (positive)		1.00	-0.40	0.84	-0.31
RoBERTa (negative)			1.00	-0.27	0.86
BERTweet (positive)				1.00	-0.26
BERTweet (negative)					1.00

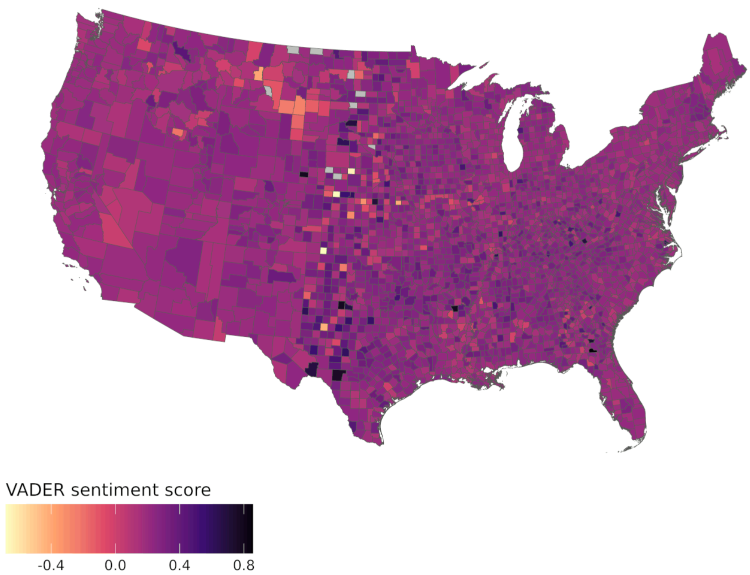

Figure 8: Mean VADER sentiment polarity by county during July 2015.

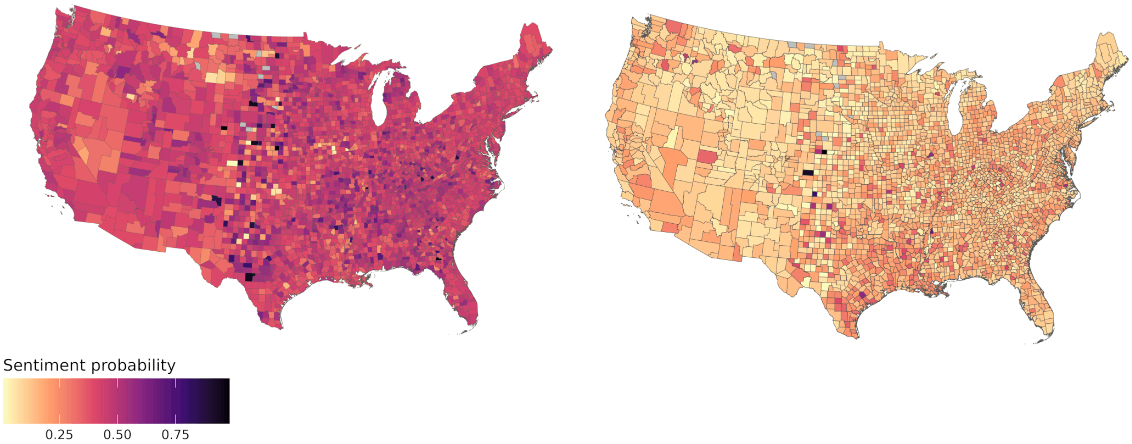

Figure 9: Mean RoBERTa positive (left) and negative (right) sentiment probability by county during July 2015.

Table 9: Summary statistics of the unstandardized measures of sentiment from BERTweet, RoBERTa and VADER on the level of individual posts.

Variable	Mean	Median	SD	Min	Max	Count
BERTweet (positive)	0.31	0.08	0.37	0.00	0.99	9,814,781
BERTweet (negative)	0.08	0.00	0.23	0.00	0.98	9,814,781
RoBERTa (positive)	0.39	0.21	0.35	0.00	0.99	9,814,781
RoBERTa (negative)	0.10	0.01	0.22	0.00	0.97	9,814,781
VADER	0.19	0.00	0.38	-1.00	1.00	9,814,781

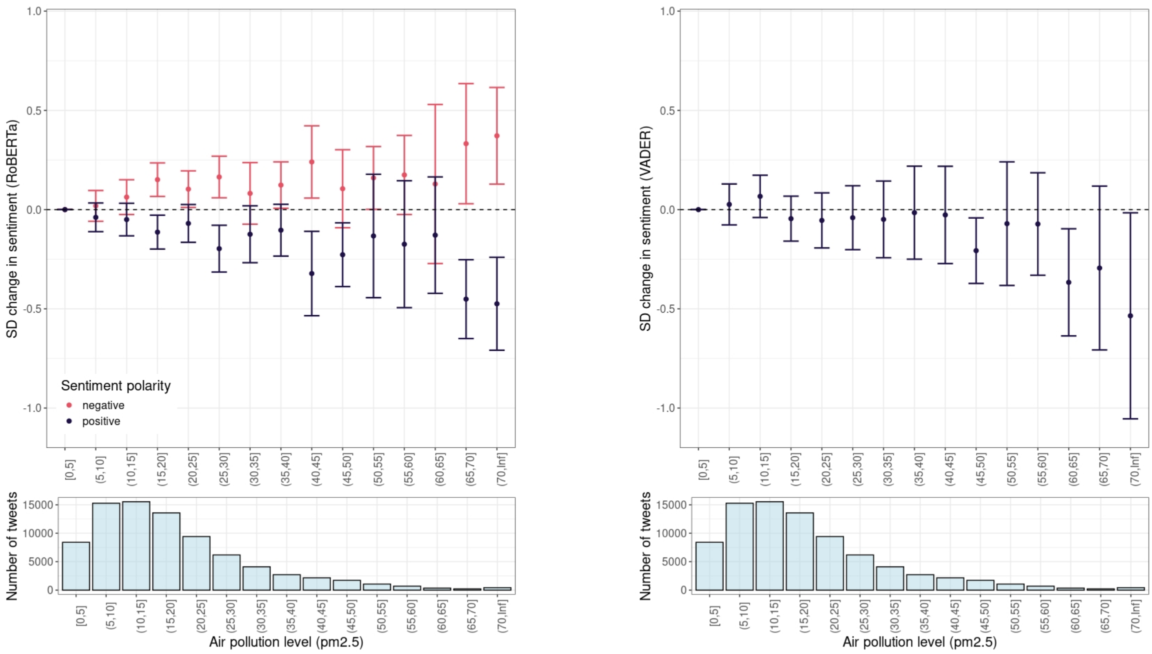

Figure 10: Visualization of results of estimating Equation (2) at the county-date level with 95 % confidence intervals. Left panel: RoBERTa as sentiment measure. Right panel: VADER as sentiment measure. Standard errors are clustered by county and date.

Table 10: Nonlinear specification.

	Sentiment measures				
	BERTweet (positive)	BERTweet (negative)	RoBERTa (positive)	RoBERTa (negative)	VADER (compound)
Independent Variable					
$PM_{2.5} \in (5, 10]$	−0.0409	0.0201	−0.0389	0.0187	0.0261
	(0.0546)	(0.0440)	(0.0354)	(0.0380)	(0.0505)
$PM_{2.5} \in (10, 15]$	−0.0146	0.0602	−0.0501	0.0630	0.0668
	(0.0534)	(0.0466)	(0.0402)	(0.0429)	(0.0522)
$PM_{2.5} \in (15, 20]$	-0.1069^{*}	0.1347^{***}	-0.1135^{**}	0.1509^{***}	−0.0456
	(0.0543)	(0.0460)	(0.0418)	(0.0413)	(0.0555)
$PM_{2.5} \in (20, 25]$	−0.0768	0.1189^{**}	−0.0696	0.1032^{**}	−0.0545
	(0.0621)	(0.0481)	(0.0466)	(0.0451)	(0.0680)
$PM_{2.5} \in (25, 30]$	-0.1767^{***}	0.2177^{***}	-0.1968^{***}	0.1643^{***}	−0.0408
	(0.0544)	(0.0575)	(0.0577)	(0.0513)	(0.0789)
$PM_{2.5} \in (30, 35]$	-0.1351^{*}	0.1009	-0.1243^{*}	0.0816	−0.0493
	(0.0722)	(0.0776)	(0.0702)	(0.0760)	(0.0945)
$PM_{2.5} \in (35, 40]$	−0.0958	0.1387^{**}	−0.1037	0.1236^{**}	−0.0155
	(0.0764)	(0.0578)	(0.0639)	(0.0572)	(0.1148)
$PM_{2.5} \in (40, 45]$	-0.3463^{***}	0.3292^{***}	-0.3221^{***}	0.2402^{**}	−0.0267
	(0.1082)	(0.0931)	(0.1041)	(0.0892)	(0.1200)
$PM_{2.5} \in (45, 50]$	-0.1895^{***}	0.1472^{*}	-0.2273^{***}	0.1052	-0.2069^{**}
	(0.0663)	(0.0827)	(0.0786)	(0.0963)	(0.0809)
$PM_{2.5} \in (50, 55]$	−0.1584	0.1271	−0.1329	0.1597^{**}	−0.0708
	(0.1226)	(0.0852)	(0.1522)	(0.0775)	(0.1525)
$PM_{2.5} \in (55, 60]$	−0.1810	0.1551	−0.1744	0.1746^{*}	−0.0726
	(0.1671)	(0.1206)	(0.1566)	(0.0976)	(0.1264)
$PM_{2.5} \in (60, 65]$	−0.1499	0.0388	−0.1287	0.1290	-0.3666^{***}
	(0.1437)	(0.1713)	(0.1436)	(0.1962)	(0.1322)
$PM_{2.5} \in (65, 70]$	-0.4640^{***}	0.4712^{***}	-0.4511^{***}	0.3323^{**}	−0.2943
	(0.1491)	(0.1463)	(0.0973)	(0.1482)	(0.2021)
$PM_{2.5} > 70$	-0.4583^{**}	0.3554	-0.4745^{***}	0.3719^{***}	-0.5352^{**}
	(0.1723)	(0.2469)	(0.1147)	(0.1192)	(0.2542)
County FE	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes
Observations	81,909	81,909	81,909	81,909	80,097
R-squared	0.33	0.34	0.36	0.39	0.08

Notes: Estimates of Equation (2) at the county-date level. Standard errors are clustered by county and date and reported in parentheses. Weather control variables are not shown. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

A.3 Tweet classification examples

Text	VADER	RoBERTa (positive)	RoBERTa (negative)	BERTweet (positive)	BERTweet (negative)
Loving my routine, beach before school ✌🏻🙌🏻🌴☀️ @ La Jolla, California	0.796	0.988	0.002	0.988	0.001
Future and meek been on some shit 🔥	-0.718	0.904	0.008	0.003	0.970
I'm thinking of a division at Mango Media dedicated to naming IPAs. I've gone from Moose Drool in…	0.459	0.140	0.013	0.048	0.003
Round 4 was a success 😛	0.572	0.976	0.002	0.950	0.001
Nothing changes.	0.000	0.136	0.274	0.007	0.903
Can't believe she did that	0.000	0.009	0.877	0.010	0.934
Ive never had a choice in my future college . I dont know why I even try .	0.000	0.007	0.916	0.003	0.973
First time in 11 years I haven't spent 4th of July with Jj 💔 today sucks	0.153	0.009	0.923	0.004	0.974
Good times, good people, good food. 😍	0.893	0.985	0.003	0.993	0.002

Table 11: Tweet classification examples.

A.4 Examples of tweets identified as bot-generated

Text	VADER	RoBERTa (positive)	RoBERTa (negative)	BERTweet (positive)	BERTweet (negative)
At 1:30 PM, Fort Morgan [Baldwin Co, AL] FIRE DEPT/RESCUE reports LIGHTNING #MOB	-0.482	0.018	0.045	0.014	0.009
We're #hiring! Read about our latest #job opening here: Recruiter (Entry Level Sales) #CityofIndustry, CA #Sales	0.000	0.594	0.004	0.211	0.002
Disabled vehicle on #TaconicStateParkway NB at Town of Putnam Valley; Town of Carmel Line	0.000	0.021	0.113	0.008	0.039
Sacramento man accused of purposely hitting cyclists pleads not guilty: A man accused of purposely hitting three…	-0.266	0.020	0.235	0.014	0.263
Opened Street or Sidewalk Cleaning request via iphone at 5317 Mission St. Garbage bags left on sidewalk.	0.000	0.025	0.454	0.008	0.032
See a video tour of my #listing 527 Beach Club Trail C410 #GulfShores #AL #realestate	0.000	0.091	0.006	0.055	0.003

Table 12: Examples of tweets identified as bot-generated.

A.5 Terms excluded in robustness analysis in Section 4.1

pollution, air quality, smog, particulate matter, PM2.5, PM10, ozone, carbon monoxide, CO_2, sulfur dioxide, nitrogen dioxide, emissions, exhaust, greenhouse gases, toxic air, hazardous air, fumes, environmental impact, contamination, pollutants, soot, dust, ash, industrial emissions, traffic pollution, atmospheric pollution, clean air, respiratory issues, asthma triggers, air purifier, air monitoring, pollution index, wildfire, wild fire, forest fire, bushfire, bush fire, wildland fire, firestorm, fire hazard, firefighting, firefighter, firefighters, controlled burn, prescribed burn, fire season, fire suppression, smoke plume, wildfire smoke, flames, blaze, burning, scorched, wildfire risk, fire evacuation, fire danger, fire warning, fire watch, smoke, ash, ember, dry lightning, flame retardant, fire retardant, fire containment, evacuation order, red flag warning, fire line, firebreak, brush fire, wildland, fire lookout, burn ban, burn zone, smoldering, charred, fire risk, air quality alert

MUNI Econ Working Paper Series (since 2018)

2025-05 Kecskésová, M., Mikula, Š. (2025). *The Effects of Air Pollution on Mood: Evidence from Twitter*. MUNI ECON Working Paper n. 2025-05. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2025-05

2025-04 Bucciol, A., Easaw, J., Trucchi, S. (2025). Household Income Expectations: The Role of Unexpected Income Changes and Aggregate Conditions. https://doi.org/10.5817/WP_MUN_ECON_2025-04

2025-03 Sarsenbayeva, A.; Alpysbayeva, D. (2025). Catastrophic Health Expenditure during Healthcare Financing Reform: Evidence from Kazakhstan. https://doi.org/10.5817/WP_MUN_ECON_2025-03

2025-02 Scervini, F., Trucchi, S. (2025). Alcohol Consumption in an Empty Nest. MUNI ECON Working Paper n. 2025-02. Brno: Masaryk University. https://doi.org/10.5817/WP_MUN_ECON_2025-02

2025-01 Soucek, C., Reggiani, T., Kairies-Schwarz, N. (2025). *Physicians' Responses to Time Pressure: Experimental Evidence on Treatment Quality and Documentation Behaviour*. MUNI ECON Working Paper n. 2025-01. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2025-01

2024-06 Fazio, A., Reggiani, T., Santori, P. (2024). *'Blessed are the Poor' The Weberian Spirit of Capitalism Under Experimental Scrutiny*. MUNI ECON Working Paper n. 2024-06. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2024-06

2024-05 Pulina, M., Salis, A. (2024). *Dynamics and trends of drug dealing: a local labour system perspective*. MUNI ECON Working Paper n. 2024-05. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2024-05

2024-04 Coufalová, L., Kecskésová, M., Mikula, Š., Ševčík, M. (2024). *Does Democracy Flourish in the Dark? Regional Development and Democracy Building*. MUNI ECON Working Paper n. 2024-04. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2024-04

2024-03 Coufalová, L., Dellinger, F. H., Huber, P., Mikula, Š. (2024). *Borders and Population Growth: Evidence from a Century of Border Regime Changes on the Austrian-Czech Border*. MUNI ECON Working Paper n. 2024-03. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2024-03

2024-02 Prochazka, J., Pandey, S., Castek, O., Firouzjaeiangalougah, M. (2024). *Replication of Changing Hearts and Minds? Why Media Messages Designed to Foster Empathy Often Fail (Gubler et al., 2022)*. MUNI ECON Working Paper n. 2024-02. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2024-02

2024-01 Marini, M. M., Ulivieri, G. 2024. *Meta-analyses in Economic Psychology: A sustainable approach to cross-cultural differences*. MUNI ECON Working Paper n. 2024-01. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2024-01

2023-09 Levi, E., Ramalingam, A. 2023. *Absolute vs. relative poverty and wealth: Cooperation in the presence of between-group inequality*. MUNI ECON Working Paper n. 2023-09. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2023-09

2023-08 Fumarco, L., Harrell, B., Button, P., Schwegman, D., Dils, E. 2023. *Gender Identity, Race, and Ethnicity-based Discrimination in Access to Mental Health Care: Evidence from an Audit Correspondence Field Experiment*. MUNI ECON Working Paper n. 2023-08. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2023-08

2023-07 Levi, E., Bayerlein, M., Grimalda, G., Reggiani, T. 2023. *Narratives on migration and political polarization: How the emphasis in narratives can drive us apart*. MUNI ECON Working Paper n. 2023-07. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2023-07

2023-06 Fumarco, L., Gaddis, S. M., Sarracino, F., Snoddy, I. 2023. *sendemails: An automated email package with multiple applications*. MUNI ECON Working Paper n. 2023-06. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2023-06

2023-05 Harrell, B., Fumarco, L., Button, P., Schwegman, D., Denwood, K. 2023. *The Impact of COVID-19 on Access to Mental Healthcare Services*. MUNI ECON Working Paper n. 2023-05. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2023-05

2023-04 Friedhoff, T., Au, C., Krahnhof, P. 2023. *Analysis of the Impact of Orthogonalized Brent Oil Price Shocks on the Returns of Dependent Industries in Times of the Russian War*. MUNI ECON Working Paper n. 2023-04. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2023-04

2023-03 Mikula, Š., Reggiani, T., Sabatini, F. 2023. *The long-term impact of religion on social capital: lessons from post-war Czechoslovakia*. MUNI ECON Working Paper n. 2023-03. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2023-03

2023-02 Clò, S., Reggiani, T., Ruberto, S. 2023. *onsumption feedback and water saving: An experiment in the metropolitan area of Milan*. MUNI ECON Working Paper n. 2023-02. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2023-02

2023-01 Adamus, M., Grežo, M. 2023. *Attitudes towards migrants and preferences for asylum and refugee policies before and during Russian invasion of Ukraine: The case of Slovakia*. MUNI ECON Working Paper n. 2023-01. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2023-01

2022-12 Guzi, M., Kahanec, M., Mýtna Kureková, L. 2022. *The Impact of Immigration and Integration Policies On Immigrant-Native Labor Market Hierarchies*. MUNI ECON Working Paper n. 2022-12. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-12

2022-11 Antinyan, A., Corazzini, L., Fišar, M., Reggiani, T. 2022. *Mind the framing when studying social preferences in the domain of losses*. MUNI ECON Working Paper n. 2022-11. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-11

2022-10 Corazzini, L., Marini, M. 2022. *Focal points in multiple threshold public goods games: A single-project meta-analysis*. MUNI ECON Working Paper n. 2022-10. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-10

2022-09 Fazio, A., Scervini, F., Reggiani, T. 2022. *Social media charity campaigns and pro-social behavior. Evidence from the Ice Bucket Challenge.*. MUNI ECON Working Paper n. 2022-09. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-09

2022-08 Coufalová, L., Mikula, Š. 2022. *The Grass Is Not Greener on the Other Side: The Role of Attention in Voting Behaviour.*. MUNI ECON Working Paper n. 2022-08. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-08

2022-07 Fazio, A., Reggiani, T. 2022. *Minimum wage and tolerance for inequality.*. MUNI ECON Working Paper n. 2022-07. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-07

2022-06 Mikula, Š., Reggiani, T. 2022. *Residential-based discrimination in the labor market*. MUNI ECON Working Paper n. 2022-06. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-06

2022-05 Mikula, Š., Molnár, P. 2022. *Expected Transport Accessibility Improvement and House Prices: Evidence from the Construction of the World's Longest Undersea Road Tunnel*. MUNI ECON Working Paper n. 2022-05. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-05

2022-04 Coufalová, L., Mikula, Š., Ševčík, M. 2022. *Homophily in Voting Behavior: Evidence from Preferential Voting*. MUNI ECON Working Paper n. 2022-04. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-04

2022-03 Kecskésová, M., Mikula, Š. 2022. *Malaria and Economic Development in the Short-term: Plasmodium falciparum vs Plasmodium vivax*. MUNI ECON Working Paper n. 2022-03. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-03

2022-02 Mladenović, D., Rrustemi, V., Martin, S., Kalia, P., Chawdhary, R. 2022. *Effects of Sociodemographic Variables on Electronic Word of Mouth: Evidence from Emerging Economies*. MUNI ECON Working Paper n. 2022-02. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-02

2022-01 Mikula, Š., Montag, J. 2022. *Roma and Bureaucrats: A Field Experiment in the Czech Republic*. MUNI ECON Working Paper n. 2022-01. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2022-01

2021-14 Abraham, E. D., Corazzini, L., Fišar, M., Reggiani, T. 2021. *Delegation and Overhead Aversion with Multiple Threshold Public Goods*. MUNI ECON Working Paper n. 2021-14. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-14

2021-13 Corazzini, L., Cotton, C., Longo, E., Reggiani, T. 2021. *The Gates Effect in Public Goods Experiments: How Donations Flow to the Recipients Favored by the Wealthy*. MUNI ECON Working Paper n. 2021-13. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-13

2021-12 Staněk, R., Krčál, O., Mikula, Š. 2021. *Social Capital and Mobility: An Experimental Study*. MUNI ECON Working Paper n. 2021-12. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-12

2021-11 Staněk, R., Krčál, O., Čellárová, K. 2021. *Pull yourself up by your bootstraps: Identifying procedural preferences against helping others in the presence*. MUNI ECON Working Paper n. 2021-11. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-11

2021-10 Levi, E., Sin, I., Stillman, S. 2021. *Understanding the Origins of Populist Political Parties and the Role of External Shocks*. MUNI ECON Working Paper n. 2021-10. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-10

2021-09 Adamus, M., Grežo, M. 202. *Individual Differences in Behavioural Responses to the Financial Threat Posed by the COVID-19 Pandemic*. MUNI ECON Working Paper n. 2021-09. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-09

2021-08 Hargreaves Heap, S. P., Karadimitropoulou, A., Levi, E. 2021. *Narrative based information: is it the facts or their packaging that matters?*. MUNI ECON Working Paper n. 2021-08. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-08

2021-07 Hargreaves Heap, S. P., Levi, E., Ramalingam, A. 2021. *Group identification and giving: in-group love, out-group hate and their crowding out*. MUNI ECON Working Paper n. 2021-07. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-07

2021-06 Medda, T., Pelligra, V., Reggiani, T. 2021. *Lab-Sophistication: Does Repeated Participation in Laboratory Experiments Affect Pro-Social Behaviour?*. MUNI ECON Working Paper n. 2021-06. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-06

2021-05 Guzi, M., Kahanec, M., Ulceluse M., M. 2021. *Europe's migration experience and its effects on economic inequality*. MUNI ECON Working Paper n. 2021-05. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-05

2021-04 Fazio, A., Reggiani, T., Sabatini, F. 2021. *The political cost of lockdown´s enforcement*. MUNI ECON Working Paper n. 2021-04. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-04

2021-03 Peciar, V. *Empirical investigation into market power, markups and employment*. MUNI ECON Working Paper n. 2021-03. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-03

2021-02 Abraham, D., Greiner, B., Stephanides, M. 2021. *On the Internet you can be anyone: An experiment on strategic avatar choice in online marketplaces*. MUNI ECON Working Paper n. 2021-02. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-02

2021-01 Krčál, O., Peer, S., Staněk, R. 2021. *Can time-inconsistent preferences explain hypothetical biases?*. MUNI ECON Working Paper n. 2021-01. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2021-01

2020-04 Pelligra, V., Reggiani, T., Zizzo, D.J. 2020. *Responding to (Un)Reasonable Requests by an Authority*. MUNI ECON Working Paper n. 2020-04. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2020-04

2020-03 de Pedraza, P., Guzi, M., Tijdens, K. 2020. *Life Dissatisfaction and Anxiety in COVID-19 pandemic*. MUNI ECON Working Paper n. 2020-03. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2020-03

2020-02 de Pedraza, P., Guzi, M., Tijdens, K. 2020. *Life Satisfaction of Employees, Labour Market Tightness and Matching Efficiency*. MUNI ECON Working Paper n. 2020-02. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2020-02

2020-01 Fišar, M., Reggiani, T., Sabatini, F., Špalek, J. 2020. *a*. MUNI ECON Working Paper n. 2020-01. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2020-01

2019-08 Fišar, M., Krčál, O., Špalek, J., Staněk, R., Tremewan, J. 2019. *A Competitive Audit Selection Mechanism with Incomplete Information*. MUNI ECON Working Paper n. 2019-08. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2019-08

2019-07 Guzi, M., Huber, P., Mikula, M. 2019. *Old sins cast long shadows: The Long-term impact of the resettlement of the Sudetenland on residential migration*. MUNI ECON Working Paper n. 2019-07. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2019-07

2019-06 Mikula, M., Montag, J. 2019. *Does homeownership hinder labor market activity? Evidence from housing privatization and restitution in Brno.* MUNI ECON Working Paper n. 2019-06. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2019-06

2019-05 Krčál, O., Staněk, R., Slanicay, M. 2019. *Made for the job or by the job? A lab-in-the-field experiment with firefighters.* MUNI ECON Working Paper n. 2019-05. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2019-05

2019-04 Bruni, L., Pelligra, V., Reggiani, T., Rizzolli, M. 2019. *The Pied Piper: Prizes, Incentives, and Motivation Crowding-in*. MUNI ECON Working Paper n. 2019-04. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2019-04

2019-03 Krčál, O., Staněk, R., Karlínová, B., Peer, S. 2019. *Real consequences matters: why hypothetical biases in the valuation of time persist even in controlled lab experiments*. MUNI ECON Working Paper n. 2019-03. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2019-03

2019-02 Corazzini, L., Cotton, C., Reggiani, T., 2019. *Delegation And Coordination With Multiple Threshold Public Goods: Experimental Evidence*. MUNI ECON Working Paper n. 2019-02. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2019-02

2019-01 Fišar, M., Krčál, O., Staněk, R., Špalek, J. 2019. *The Effects of Staff-rotation in Public Administration on the Decision to Bribe or be Bribed*. MUNI ECON Working Paper n. 2019-01. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2019-01

2018-02 Guzi, M., Kahanec, M. 2018. *Income Inequality and the Size of Government: A Causal Analysis*. MUNI ECON Working Paper n. 2018-02. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2018-02

2018-01 Geraci, A., Nardotto, M., Reggiani, T., Sabatini, F. 2018. *Broadband Internet and Social Capital*. MUNI ECON Working Paper n. 2018-01. Brno: Masaryk University. https://doi.org/10.5817/WP_MUNI_ECON_2018-01

ISSN electronic edition 2571-130X
MUNI ECON Working Paper Series is indexed in RePEc:
https://ideas.repec.org/s/mub/wpaper.html