Does Air Pollution Impair Work Safety? The Impact of PM2.5 on Severe Workplace Injuries

Zhanhan Yu *

Preliminary Draft. Comments Welcome.

Please do not circulate or cite without permission of the author

January, 2023

Abstract

I investigate the causal effect of air pollution on workplace safety using novel data on work-related severe injuries and PM2.5 pollution in the United States from 2015 through 2018. I focus on fine particulate matter, known as PM2.5, a primary air pollutant found to adversely impact human cognitive abilities and potentially affect workplace safety via biological channels. Credibly pinning down the causal effect of air pollution is challenging because air pollutants are not randomly assigned across space or workplaces. To deal with the endogeneity of air pollution, I employ a quasi-experimental design, exploiting exogenous variation in PM2.5 driven by two different instruments — rainfall and wind direction. I start by testing the validity of these instruments and show that they violate the assumptions for point identification of interest. Then, I leverage partial identification strategies using the same instruments to estimate bounds on the effect of air pollution. The estimated bounds on the effect of PM2.5 pollution suggest that air pollution increases the workplace accident rate by between 7 and 39% relative to the sample average accident rate. The effect appears to be more prominent for industries that require outdoor work, such as agriculture, mining, and construction, than non-manufacturing indoor industries. A back-of-the-envelope calculation indicates that increasing the annual number of days with PM2.5 pollution by two days is estimated to raise annual total costs of workers' compensation by at least 0.9 billion dollars and up to 5.1 billion, equivalent to about 1 to 8% of total workers' compensation paid in 2018.

Keywords: Air Pollution; Work Safety; Bounds; Imperfect IV
JEL Classification: Q53, Q51, J28, J24

*Department of Economics and Center for Policy Research, Syracuse University; 426 Eggers Hall, Syracuse NY 13244-1020; Email: zyu127@syr.edu. I am grateful to Alfonso Flores-Lagunes, Hugo Jales, and Maria Zhu for advice and guidance. I thank Gary Engelhardt, David Popp, Alex Rothenberg, Ying Shi, Panle Jia Barwick, and Analisa Packham for their insightful comments. Useful feedback has been provided by participants at the Syracuse University's Applied Micro Seminar, Southern Economic Association 92nd Annual Meeting, and the Syracuse Labor Group. All errors are my own.

1 Introduction

The analysis of workplace safety is a perennial topic of interest in labor economics. In particular, workplace injuries are costly to workers, employers, and society. Workers' compensation benefits, along with productivity loss and medical expenses incurred because of work-related deaths and injuries, cost more than $234 billion in 2018 (National Safety Council; Weiss, Murphy, and Boden 2020).

Among all aspects that impact workplace safety, air pollution is a factor that is surprisingly understudied. In addition to affecting visibility, air pollution, especially fine particulate matter, can affect work safety via biological channels. Fine particulate matter is inhalable and can penetrate deep into human bodies. Recent studies have found links between fine particulate matter and the decline of cognitive functions that play an important role in avoiding workplace accidents: e.g., memory, attention, and fluid reasoning (La Nauze and Severnini 2021; Bedi et al. 2021; Shehab and Pope 2019). More broadly, air pollution is a global environmental problem with far-reaching impacts on human health and economic development in both developing and developed countries. While low- and middle-income countries like India and China bear a disproportionate burden of such pollution, in the era of global warming, extreme weather events, such as wildfires and droughts, are posing new challenges to air pollution control in high-income countries and places like California and Utah.

This paper studies the impact of air pollution on workplace safety in the United States. I focus on PM2.5, which is fine particulate matter with a nominal mean aerodynamic diameter of up to 2.5 micrometers, and investigate its causal effect on severe workplace injuries, defined by the Occupational Safety and Health Administration (OSHA) as injuries involving amputation, in-patient hospitalization, or loss of an eye. To do this, I create a novel panel dataset for the contiguous United States from 2015 to 2018, compiling administrative data on severe workplace injuries from OSHA, high-resolution data on daily PM2.5 levels from the U.S. Environmental Protection Agency (EPA), and weather data from the North American Regional Reanalysis (NARR).

Pinning down the causal effect of air pollution is challenging, because it is not randomly assigned across space or workplaces. Many economic activities are correlated with both pollution and workplace accidents, raising concerns of omitted variable bias. In addition, workers and employers can adjust the extent of exposure to pollution by, for example, changing the work schedule. Such responses affect the incidence of workplace injuries, assuming the likelihood of workplace accidents is correlated with the length of time workers spend at work. Moreover, if the responses of workers and employers differ across locations due to unobserved factors, e.g., workers are more cautious in polluted places and thus their labor supply is more sensitive to air quality deterioration, then standard estimates may suffer from selection bias.

A typical solution is to apply the instrumental variable (IV) method. I employ two different and widely-used instruments to estimate the causal effect of PM2.5 pollution. One is wind direction, a popular instrument for air pollution in the environmental literature. The other is rainfall, a determining factor for removing particulate matter from the atmosphere. I estimate a two-stage least squares model of workplace accident rates as a function of PM2.5 using the constructed panel data. The IV estimate using wind direction is essentially zero (0.003 percentage points with a standard error of 0.0051), indicating no impact of air pollution on workplace safety. Furthermore, the IV estimate using rainfall is negative (-0.09 percentage points with a standard error of 0.0041), but not statistically different from zero. In contrast, the corresponding Ordinary Least Squares (OLS) estimate is 0.11 percentage points and statistically significant at the 5% level. These results suggest OLS estimation is biased upward.

Among other assumptions, a valid IV must satisfy the exclusion restriction: the instrument must be uncorrelated with workplace safety except through its effect on air pollution. Violation of this restriction will render the instrumental variable invalid, cause the traditional IV estimator to be biased, and make it infeasible to point identify the causal effect of interest. In my empirical setting, I find statistical evidence that these two proposed instruments are invalid by applying the testing procedures of Mourifié and Wan (2017).

Consequently, I adopt the partial identification methods of Nevo and Rosen (2012) and Ban and Kédagni (2022) to estimate bounds on the causal effect of air pollution. These methods allow me to exploit the exogenous variability induced by the instrumental variables, while accounting for their failure of excludability. These partial identification methods substitute the exclusion restriction with three weaker assumptions: (1) the same direction correlation, which states that the correlation of the instrumental variable with the potential outcomes is of the same direction as the correlation of the treatment variable with the potential outcomes; (2) the less endogenous instrument, which asserts the instrumental variable is less correlated with the potential outcomes than the treatment variable; and (3) the monotone treatment response, which imposes the restriction that worse air quality does not decrease the incidence of workplace accidents.

For wind direction, the bounds estimation suggests that experiencing PM2.5 pollution for more than five days in a month increases the workplace accident rate by at least 7% and up to 39% relative to the sample mean accident rate. Results using rainfall are of similar magnitude (at least 5% and up to 39%). Although the estimated bounds cannot statistically rule out a zero effect at conventional significance levels, they rule out the effects greater than 3.9 percentage points. Under the monotone treatment response assumption, the 95% confidence region of estimated bounds contains the OLS estimate and the 2SLS estimate employing wind direction, but excludes the 2SLS estimate using rainfall. In addition, the estimated effects are comparable to point estimates of the effect of other air pollutants (8.6% for a ten-unit increase in the air quality index (AQI), Depalo and Palma 2020) and the effect of other environmental factors (4.8% to 6.6% for working on a higher temperature day, Park, Pankratz, and Behrer 2021). Moreover, I examine the heterogeneous impacts for different industries and find that the effects appear to be more prominent in industries that require outdoor work, e.g., agriculture, mining, and construction.

Utilizing the estimated bounds, I quantify the impact of air pollution on workers' compensation costs. A back-of-the-envelope calculation implies that reducing the number of PM2.5-polluted days by two days per year saves annual workers' compensation costs by at least 0.9 billion dollars and up to 5.1 billion. The midpoint reduction in this range is about 3 billion dollars, equivalent to 4.8% of total workers' compensation paid in 2018.

This paper contributes to the literature by exploring an additional channel through which air pollution affects worker productivity: increasing severe workplace injuries. In this way, I complement studies that investigate the causal link of environmental factors with worker productivity (Chang et al. 2016; Graff Zivin and Neidell 2012; Park, Pankratz, and Behrer 2021, among others).

Most of the existing literature that examines the causal effect of air pollution on labor productivity restricts the sample to a single firm or small geographical areas. This paper leverages granular national-level data covering 48 states in the contiguous United States and offers evidence beyond selective firms and geographical areas. In addition, this paper formally tests the validity of widely used instruments for air pollution. The validity of instruments depends on the outcome variable under consideration. For the outcome of workplace injuries, I provide statistical evidence that the rainfall and wind direction instruments are invalid. Moreover, my exercise of bounds estimation illustrates that the partial identification strategy is a useful approach to investigate causal effects when valid instruments are unavailable. Lastly, this paper has policy implications. It suggests that improving air quality benefits both employers and employees and ignoring the effect of air pollution on workplace safety likely underestimates these benefits of air quality improvement.

The rest of the paper is organized as follows. In Section 2, I introduce PM2.5 and discuss the biological channels through which PM2.5 affects work safety. I describe the data in Section 3. In Section 4, I test the validity of two proposed instruments. I discuss in Section 5 the partial identification methods and in Section 6 the estimated bounds. Based on these results, I evaluate the impact of PM2.5 pollution on workers' compensation costs in Section 7. Section 8 concludes.

2 Background and Previous Studies

2.1 Particulate Matter

Particulate matter (PM, also called particle pollution) is a mixture of solid particles and liquid droplets found in the ambient air, which varies greatly in size and composition. Fine PM (PM2.5) is the particulate matter with a nominal mean aerodynamic diameter of up to 2.5 micrometers (μm), whereas coarse PM (PM10-2.5) is the particulate matter with a nominal mean aerodynamic diameter of 2.5-10 micrometers. Coarse particles PM2.5-10 is generally comprised of dust, dirt, soot, or smoke, are large or dark enough to be seen with the naked eye. Fine particles PM2.5 can be as small as 1/70 of the average human hair (about 70 micrometers in diameter) (EPA 2019). Particles of different sizes also differ in their atmospheric lifetimes, transport distances, and removal processes

(EPA 2009). Specifically, coarse PM (PM10-2.5) has substantially greater spatial variability than fine particles (PM2.5). The transport distances for PM10-2.5 range from less than 1 km to 1000 km, but its time in the atmosphere only lasts for hours, while PM2.5 can transport from 10 km to 100 km and remain suspended in air for days to weeks. The major processes for removing both PM10-2.5 and PM2.5 from the atmosphere are incorporation of particles into cloud droplets that fall as rain (rainout) and collisions with falling rain (washout), as well as fallout (EPA 2009).[1] The removal process for PM motivates the choice of the instrumental variable of PM2.5. As shown in Section 4, there is a significant and sizable negative relationship between precipitation and the level of PM2.5 in an area.

PM is comprised of components that are directly emitted (primary PM) as well as formed through atmospheric chemical reactions (secondary PM). Within an urban environment, most primary PM2.5 emissions are from anthropogenic sources and include some combination of industrial activities, motor vehicles, and fuel combustion, including biomass burning. However, in many locations, secondary PM formed from the precursors sulfur dioxide (SO2), oxides of nitrogen (NOx), ammonia (NH3), and volatile organic compounds (VOCs), accounts for the majority of PM2.5 mass (EPA 2019). Hence, particle pollution is highly correlated to human activities.

Exposure to PM is not limited to outdoor activities. In indoor environments without indoor sources of pollution, pollutants from outdoors are the main cause of indoor air pollution (WHO 2021). Outdoor air pollutants can affect the indoor air quality via the processes of ventilation and penetration through the building envelope. PM2.5 is found to have the highest ventilation rate and penetration rate (WHO 2021). It also implies that not only outdoor workers but also indoor workers, especially for those who work in poorly insulated, well-ventilated workplaces, are affected by PM2.5 (Chang et al. 2016). Thus, presumably, PM2.5 exposure and its effect on workplace accidents should not be limited to industries with more outdoor jobs, though one may expect a stronger impact for these industries.

[1]Other hydrometeors (snow, ice) can also serve the same purpose as rain in the rainout process.

2.2 Mechanisms

PM2.5 is inhalable and is found to have serious adverse health effects due to its ability to penetrate deep into the lungs, blood streams and brain, causing health problems including heart attacks, respiratory diseases, and premature death (for a comprehensive review on the adverse health effects of PM2.5, see EPA 2009, 2019). In addition, exposure to PM2.5 also leads to more subtle effects, such as changes in blood pressure, irritation in the eyes, ear, nose, throat, and skin, and mild headaches (Pope 2000; Auchincloss et al. 2008). These milder effects, which arise from exposure to lower levels of PM2.5, typically do not lead to health care encounters, but may be associated with inattention and distraction (Chang et al. 2016).

A number of studies have documented the impacts of PM2.5 on human cognitive ability. PM2.5 is found to impair human cognitive functions (Guxens et al. 2018; Peters et al. 2000; La Nauze and Severnini 2021; Bedi et al. 2021) and affect mental well-being (Weuve et al. 2012; Prado Bert et al. 2018). Even short-term exposure to PM2.5 can result in the decline of adults' cognitive performance (Shehab and Pope 2019). The cognitive functions affected by PM2.5 include memory, attention, and fluid reasoning. For example, La Nauze and Severnini (2021) found exposure to PM2.5 impairs adults' cognitive performance, with the largest effect for those in prime working age and the strongest effect on memory. Bedi et al. (2021) found that exposure to high levels of PM2.5 reduces college students' performance on a fluid reasoning test.

PM2.5 pollution can affect work safety through its link with the decline of cognitive functions. Inattention, memory change, and the decline of the fluid reasoning ability are potentially linked to workplace accidents. Although, there is no direct evidence of the association between the lack of attention and work accident occurrence, inattention and distraction are found to be an important factor leading to road accidents (Rezaei and Klette 2014; Talbot, Fagerlind, and Morris 2013). Moreover, memory change is found to be correlated to the reduced ability to ignore irrelevant information and slowed processing speed (Harada, Love, and Triebel 2013), which are potential factors increasing the incidence of workplace accidents. Lastly, fluid reasoning is the ability to think flexibly and solve problems in novel situations. It help human beings to deal with uncertainty and fast-changing situation. The decline of the fluid reasoning ability may reduce workers' ability to avoid accidents in the workplace or to reduce the extent of injuries caused by the accident.

2.3 Previous Studies

The adverse impact of air pollution on health outcomes is well-documented (Deryugina et al. 2019; Anderson 2020, among others). There is a growing interest among economists in the impact of air pollution on non-health outcomes (e.g., Bharadwaj et al. 2017; Hanna and Oliva 2015; Graff Zivin and Neidell 2012). Early studies, such as Hausman, Ostro, and Wise (1984), Ostro (1983), Pönkä (1990), and Hansen and Selte (2000), have found positive correlation between various air pollutants (i.e., ozone, SO2, PM10) and sick leave and work loss. Most of the latest studies on the relationship of air pollution with labor productivity look at the agriculture and manufacturing sectors. For example, Graff Zivin and Neidell (2012) examined the causal effect of ambient ozone concentrations on worker productivity on a large farm in California. They found that a 10 parts per billion (ppb) decrease in ozone concentrations increases worker productivity by 4.2%. Chang et al. (2016) provided evidence from indoor workers in a pear-packing factory. They found an increase in PM2.5 pollution of 10 micrograms per cubic meter reduces the productivity of these workers by $0.41 per hour, approximately 6% of average hourly earnings. For developing countries, Chang et al. (2019) collected productivity data of call center workers from a service firm in China and found that a 10-unit increase in the air pollution index (API) decreases the number of daily calls handled per worker by 0.35%. Besides being affected by contemporaneous pollution, labor productivity may also respond to accumulated pollution exposure. Using worker output data from two Chinese manufacturing sites, He, Liu, and Salvo (2019) found that a 10 micrograms per cubic meter increase in the 25-day accumulated PM2.5 exposure leads to a reduction in the mean daily output by 0.5 to 3%.

Depalo and Palma (2020) investigate the impact of air pollution on work safety in Italy. They found that a ten-unit increase in the air quality index (AQI) results in a 8.6% increase in the number of accidents, while it has no effect on accident-related disabilities. My paper differs from Depalo and Palma (2020) in several ways. First, I use different air quality measures and provide evidence from the United States. More specifically, this paper focuses on the fine particulate matters (PM2.5), a major pollutant that is found to have the most adverse effect on human cognitive functions and have the potential to affect workplace safety via biological channels. In contrast, the air quality index used in Depalo and Palma (2020) does not contain PM2.5. Moreover, although both papers use instrumental variables, the instruments applied are different. Depalo and Palma (2020) uses winter heating rules as the instrument for the air quality in an instrumental variable framework. I employ the rainfall and wind direction as two different instruments for PM2.5, and use partial identification techniques given statistical evidence of the invalidity of the instruments.

3 Data

To estimate the impact of air pollution on work safety, I create a novel panel data combining administrative data on severe workplace injuries and high-resolution data on daily PM2.5 levels and meteorological variables of the contiguous United States (48 states excluding Alaska and Hawaii, and District of Columbia) from 2015 to 2018. The geographical unit of analysis is a raster with 4426 grids at 0.5 degree scale (approximately 55 kilometers). As shown in Figure 1a, each grid cell is of about 3025 (55×55) square kilometers (1168 square miles), and equivalent to areas combined Houston and Phoenix.[2] To visually illustrate the size of a grid, I plot two grids that cover the New York city in Figure 1b.

3.1 Data Sources

Workplace Accidents

The daily count of workplace accidents with severe injuries comes from two data sources from OSHA: the Severe Injury Reports and Integrated Management Information System (IMIS) Enforcement Data.[3] Starting from January 1st, 2015, the Occupational Safety and Health Administration (OSHA) requires employers to report all severe work-related injuries, defined as an amputation,

[2] I exclude grids with population count of zero (and thus missing population density values) and grids missing air quality and meteorological measures from the analysis sample. The excluded grids (1656 grids) are shown as gray shaded areas in Figure 1a. Most of them are in remote areas in the West and Midwest.

[3] Both data sets are updated periodically. The Severe Injury Reports data used in this paper was downloaded on February 3rd, 2021 and IMIS Enforcement Data was downloaded on January 2nd, 2021.

in-patient hospitalization, or loss of an eye. The reported severe injury records are compiled into a public-available data set named the Severe Injury Reports by OSHA. The Severe Injury Reports (SIR) Data lists all employer-reported injuries with information on the date, address, and geographic coordinates of the accident. The IMIS Enforcement Data consists of closed OSHA inspection cases dated back to 1970s. Linking the inspection table, accident table, and injury table from the IMIS Enforcement Data allows researchers to (indirectly) observe work-related injuries. Like the Severe Injury Reports, the IMIS Enforcement Data contains the exact date and address of the accident and information about the injury associated with the accident.[4]

To construct the grid-level measure of workplace accidents, I map accidents into the 0.5 degree scale raster using their geographic coordinates.[5] To obtain the workplace accident measure by industry, I classify accidents into 11 industry sectors by 2-digit NAICS codes. The classification of industries follows BLS which groups the 21 NAICS industry categories into two broad categories: goods-producing industries and service-providing industries, and several sub-groups within the two groups.[6]

Unlike the self-report Workers' Compensation claims which are commonly used in studies on work safety, the administrative data used in this paper contains only workplace injuries associated with physical illness (compared to psychiatric or mental illness) that are usually severe, involving hospitalization and surgery.

Air Quality

Daily PM2.5 data is retrieved from EPA's Fused Air Quality Surface Using Downscaling (FAQSD) data set.[7] FAQSD is based on a scientific model (Bayesian space-time downscaler model) that

[4]Even though all employers are obligated to report any severe injuries, employers have the incentive to under-report the accidents. To alleviate the measurement error driven by under-reporting, I use the IMIS inspection data as the complement of the SIR data. Moreover, I employ the instrumental variable method and instruments that are uncorrelated to the measurement error.

[5]Since the IMIS data does not contains information of geographic coordinates, I use the Google Maps API to search and acquire the longitude and latitude of each accident's address.

[6]BLS's classification can be found here: https://www.bls.gov/iag/tgs/iag_index_naics.htm. Table B.1 lists a crosswalk of industry categories according to the 2-digit NAICS codes and the classification by BLS.

[7]FAQSD output files are downloaded via EPA RSIG website https://www.epa.gov/hesc/rsig-related-downloadable-data-files.

projects the station-based atmosphere variables (e.g., PM2.5 and Ozone) to outputs at a finer scale.[8] Specifically, data from FAQSD provide a daily measure of ambient PM2.5 level at the 12 km scale. To merge the PM2.5 data with the workplace accidents data, I aggregate the daily PM2.5 data to the 55 km gridded raster by taking the average over all points within each 55 km grid cell.

Weather

Daily (24 hours) precipitation and wind direction data come from the North American Regional Reanalysis (NARR, Mesinger et al. 2006), which is a long-term, consistent, high-resolution climate data set for North America.[9] The raw data is measured at about 32 km scale. Like the PM2.5 data, the precipitation and wind data is integrated into the 55 km $\times$ 55 km grid cells. The propensity of workplace accidents may be affected by local weather conditions. To control for time-varying near-surface meteorological conditions, I collect weather parameters such as daily average air temperature (in degrees Celsius) and wind speed (m/s) from NARR.

Population

I collect demographic data from the Gridded Population of World (GPW).[10] GPW provides estimation of population density (number of persons per square kilometer) and population count based on counts consistent with national censuses and population registers. I use the data of year 2015 which is the starting year of the sample. GPW data is available at various resolutions, e.g., 0.5 degree scale, which is the same scale of the unit of analysis. Thus, I directly merge the population data with the workplace accidents data.

[8]The model uses both the daily fine particulate air monitoring data from the national, state, and local air monitoring stations, and emission and meteorological data from the Community Multiscale Air Quality (CMAQ) model.

[9]NARR data used in this study is downloaded from the NOAA Physical Sciences Laboratory (PSL). For more information, see https://psl.noaa.gov.

[10]Center for International Earth Science Information Network - CIESIN - Columbia University (2018). For more information about the data, check its website https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-rev11.

3.2 Data Construction and Summary Statistics

The final data used for analysis is at the monthly level. I aggregate all the daily data to the monthly level for two reasons. First, the monthly aggregated data provide more variation in the count of workplace accidents. Workplace accidents involving severe injuries are scarce in the United States and thus more than 95% of the daily observations do not have workplace accidents. Second, unlike using the daily data which investigates the spontaneous effect of PM2.5 exposure on workplace accidents on the same day, using the monthly data better captures the potentially delayed effect of PM2.5 exposure and the cumulative effect of prolonged exposure to PM2.5 pollution.

Grids that miss observations in the PM2.5 variable, population variable, or weather controls are dropped.[11] The final data used in the analysis consists of 2770 grids and 48 months (2015-2018), with total sample size of $132,960$ ($= 48 \times 2770$).

Work safety is measured by the prevalence rate of workplace accidents that involve severe injuries. Specifically, the outcome variable Y_{im} measures the monthly average accident rate per ten thousand population in grid i and month m.[12]

To measure the PM2.5 pollution at the monthly level, I construct the treatment variable following a two-stage procedure. First, I define two binary PM2.5 indicators for each grid i and day j:

$$D1_{ij} = \begin{cases} 1 & \text{if} \quad PM2.5_{ij} > 12\mu g/m^3 \\ 0 & \text{otherwise} \end{cases} \quad ; \quad D2_{ij} = \begin{cases} 1 & \text{if} \quad PM2.5_{ij} > 15\mu g/m^3 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The first indicator $D1_{ij}$ signifies days with the level of ambient PM2.5 greater than $12\mu g/m^3$. The cutoff $12\mu g/m^3$ is taken from the upper end of the range for the “Good” Air Quality Index (AQI) defined by EPA (2012). In other words, I classify PM2.5 levels greater than $12\mu g/m^3$ as “Not Good”, including EPA categories of “Moderate”, “Unhealthy for Sensitive Groups”, and “Unhealthy”. The threshold of PM2.5 level for $D2_{ij}$ is set at $15\mu g/m^3$, which is the World Health Organization’s air quality guideline value of 24-hour exposure to PM2.5 (WHO 2021). WHO

[11] There were 4426 grids constructed for the contiguous U.S. and 1656 grids are dropped due to missing observations in variables measuring the daily PM2.5 level, population, and weather conditions. As shown in Panel (a) of Figure 1, most of them are in remote areas in the West and Midwest.

[12] To remove outliers and to ensure the common support assumption for the nonparametric bounds estimation (See Section 5.2), I top-code the accident rate at the 0.05th percentile in the right tail. All empirical analyses use the same top-coded outcome variable.

provides recommended air quality guideline level after assessing the causal effect of air pollutants on health outcomes.

Figure 2 plots the geographical distribution of the average monthly number of days with PM2.5 pollution (based on the EPA's standard) over years 2015-2018. I mark the top 4 Core Based Statistical Areas (CBSAs) with the highest average monthly number of PM2.5-polluted days in blue. All of them are located in California: Fresno, Visalia, Bakersfield, and Los Angeles-Long Beach. Figure 2 shows that there are nontrivial variations in the monthly number of PM2.5-polluted days in the sample.

Then, I count the number of days whose PM2.5 levels are greater than the selected threshold at the month-year level and define two binary treatment variables as

$$T1_{imy} = \begin{cases} 1 & \text{if} \quad \sum_j D1_{ij} > k1_m \\ 0 & \text{otherwise} \end{cases} \quad ; \quad T2_{imy} = \begin{cases} 1 & \text{if} \quad \sum_j D2_{ij} > k2_m \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $k1_m$ and $k2_m$ are cutoffs of the number of air-polluted days for month m. I set the cutoff at the 75th percentile of the pooled distribution of air-polluted days for a particular month over years 2015-2018. For example, $k1_1$ is the 75th percentile of the distribution of days with PM2.5 greater than $12\mu g/m^3$ for the whole sample in January.[13] In other words, the binary indicators $T1$ and $T2$ mark the most polluted grid-month cells where the number of air-polluted days are above the 75th percentile of the air polluted days in the sample.

I present in Table 1 the summary statistics of the outcome variable (the workplace accident rate), the treatment variable (PM2.5 measures), and the covariates.[14] The sample average workplace accident rate per ten thousand population is about 0.10. For an average grid with population size at the sample mean $(112,291)$, an accident rate at 0.1 is equivalent to a monthly count of workplace accidents (with severe injuries) of about 1.12. The monthly average PM2.5 level in the contiguous United States is about 7.28 micrograms per cubic meter.

[13] In the robustness check, I use other thresholds, such as the median. Appendix Figure A.2 plots the cutoffs by month, i.e., the 75th percentile or the median of the sample distribution of the number of air polluted days.

[14] Full summary statistics can be found in the Appendix Table B.2.

4 Testing Point Identification Assumptions

In this section, I first discuss the endogeneity concerns of air pollution. To obtain consistent estimates with an endogenous treatment variable, I propose two commonly-used instruments for air pollution — rainfall and wind direction. For the instrumental variable to point identify the causal effect, it is assumed to be independent of all potential outcomes and potential treatments, thus uncorrelated to the observed outcomes except through the observed treatment (i.e., exogeneity). The instrumental variable must also affect the treatment status in a monotone way (i.e., the monotonicity condition in Imbens and Angrist (1994)). I will first present OLS and IV estimates of the effect of PM2.5 on workplace injuries. Then, I test the joint validity of the exogeneity and monotonicity assumptions for the two proposed instrumental variables, adopting the testing procedures of Mourifié and Wan (2017). I show that these assumptions are not jointly satisfied for the two proposed IVs, rendering them invalid.

4.1 OLS and Endogeneity Concerns

I present in Table 2 the OLS estimation results, regressing the workplace accident rate on each of the two PM2.5 pollution indicators. All models include grid and month by year fixed effects to control for time-specific and unit-specific confounders.[15] In addition, time-varying weather conditions, such as air temperature and wind speed, could also play a role in affecting the workplace accident rate. I control for the contemporaneous weather conditions such as monthly average air temperature and monthly average wind speed. Models are weighted by the population density in 2015.

The OLS results suggest that PM2.5 pollution is positively associated with the workplace accident rate. For example, Column (1) shows that PM2.5 pollution (as classified by the EPA criteria) is associated with an increase in the workplace accident rate by about 0.11 percentage points, which is equivalent to a percentage change of about 1.1% relative to the sample average accident rate over all grids. The estimate is statistically significant at the 5% level. The estimate is

[15]PM2.5 also varies seasonally, Figure A.1 shows the evolution of the 24-Hour average PM2.5 over all grids in the sample by date during 2015-2018. The level of average ambient PM2.5 is generally high in the summer and relatively lower in the spring and fall, which motivates the inclusion of month by year fixed effects.

smaller for PM2.5 pollution defined by the WHO criteria, as shown in Column (2), at about 0.03 percentage points and is not statistically significant at conventional levels.

These OLS estimates can be biased because PM2.5 pollution is still endogenous. Human and economic activities are highly correlated to both PM2.5 pollution and workplace accidents because emissions from industrial activities, motor vehicles, and fuel combustion are the major sources of PM2.5 pollution, while workplace safety is associated with economic development, industrial composition, and population density. In addition, workers and employers can respond to changes in air quality and adjust the extent of exposure to PM2.5 pollution in multiple ways. For instance, workers may reduce the time exposed to air pollutants by changing their work schedule. Changes in working hours affect not only workers' exposure to PM2.5 pollution but also their exposure to workplace accidents.[16] The problem of selection bias arises if the extent to which workers and firms respond to the changes of air quality differs across areas with different levels of PM2.5 pollution due to unobserved factors. For example, if workers and firms are more cautious in places with poorer air quality and thus more responsive to changes in air quality, then the unobserved preference will be associated with both workplace accidents and PM2.5 pollution, introducing selection bias.

4.2 Instrumental Variable Method

A common way to deal with the endogeneity of air pollution exposure is adopting the instrumental variable (IV) method. Wind direction and rainfall (precipitation) are two potential instrumental variables. Wind direction is an IV commonly used in air pollution literature (see, for example, Deryugina et al. 2019; Heyes and Zhu 2019; Anderson 2020; Balietti, Datta, and Veljanoska 2022). In addition, studies in environmental science have found that precipitation can clear some of the pollutants in the air via coagulation, i.e., the rainfall droplets would attract tiny aerosol particles during the process of falling to the ground, and thus "washout" the air (see, Tai, Mickley, and Jacob 2010; Guo et al. 2016). This "washout" effects are found to be most effective for heavy rain.

[16]By definition, workplace accidents are incidents occurred at workplaces. Given the probability of occurring a workplace accident per unit of time, the longer workers spend at the workplace, the higher the workplace accident rate. Another way to think of this relationship is that the longer the work hours, the more fatigued are workers (workers become less focused), and the more likely for workers to make mistakes, hence the higher risk of workplace accidents.

Therefore, I use wind direction and rainfall as two different instruments for PM2.5 pollution. The construction of these two instrumental variables follows a two-stage procedure similar to the construction of the treatment variable. For the rainfall IV, I first define a binary indicator of heavy rainfall which equals one if the daily (24 hours) accumulated precipitation is greater or equal to $15mm$.[17] Then, I count the number of days with heavy rainfall at the month-year level and define a binary instrumental variable $Z1$ as:

$$Z1_{imy} = \begin{cases} 1 & \text{if} \quad \sum_j \mathbb{1}(Prcp_{ij} \geq 15) > c_m^{Z1} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where $Prcp_{ij}$ denotes the daily accumulated precipitation of grid i on day j. c_m^{Z1} denotes the cutoff of the rainfall IV for month m. In the main analysis, the cutoff is set at the 75th percentile of the pooled distribution of heavy rainfall days for a particular month over the years 2015-2018, which is consistent with the cutoff for the treatment variable. The average number of c_m^{Z1} is 2 (i.e., two days) in the sample.

To construct the wind direction IV, I first categorize the daily average wind direction into 45-degree bins and denote the wind direction categorical variable as $WndDir_{ij}$. Then, for each grid in the sample, I run a regression of the daily PM2.5 level on the 45-degree wind direction dummies by month, and save the estimates.[18] Next, for each grid, I find the wind direction range(s) that yield the lowest and the second lowest PM2.5 pollution for each month and denote them as $WndDir\_MinPM25_{im}$ and $WndDir\_2ndMinPM25_{im}$. I define a prevailing wind direction indicator as $\mathbb{1}(WndDir_{ij} = WndDir\_MinPM25_{im}$ or $WndDir\_2ndMinPM25_{im})$. The prevailing wind direction indicator, $\mathbb{1}(WndDir)$, flags the day when the daily average wind direction of grid i falls in the 45-degree intervals that are associated with a lower level of PM2.5.[19] Then, the wind direction IV, denoted as $Z2$, is defined as

$$Z2_{imy} = \begin{cases} 1 & \text{if} \quad \sum_j \mathbb{1}(WndDir_{ij}) > c_m^{Z2} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

[17]Having $15mm$ daily accumulated precipitation is equivalent to experiencing heavy rainfall for two hours (Glickman and Zenk 2000).

[18]The sample size is about 30 days $\times\, 4$ years $= 120$ for each regression.

[19]La Nauze and Severnini (2021) used a similar process to construct the wind direction indicator suggesting wind directions that brings the highest level of PM2.5 at 60-degree intervals.

Similarly, c_m^{Z2} denotes the cutoff of the wind direction IV for month m. Like in the construction of the rainfall IV, the cutoff is set at the 75th percentile of the pooled distribution of days with wind blowing from the prevailing wind direction that associated with cleaner air for a particular month over the years 2015-2018. The average number of c_m^{Z2} is 9 (i.e., nine days) in the sample.

I show in Panel A of Table 3 the results of the first-stage estimation and find that both the rainfall and the wind direction IVs are negatively associated with the two PM2.5 pollution indicators. For example, the effect of rainfall on the PM2.5 pollution as defined by the EPA criteria is estimated at -0.107. It implies that on average, having more than two days with heavy rainfall (measured by daily accumulated precipitation greater than 15 *mm*) in a month reduces the probability of being the most air-polluted grid-month cell (the cell with the number of air-polluted days greater than the 75th percentile of the pooled distribution of air-polluted days) by about 11%. The effect of wind direction on the PM2.5 pollution (based on the EPA standard) is estimated at about -0.083, indicating that on average, the probability of being the most air-polluted grid-month cell decreases by about 8% when the wind blowing from the prevailing wind direction associated with cleaner air for over nine days.

To evaluate the instrument strength, I report the effective first-stage F statistics for a robust weak instrument test described in Montiel Olea and Pflueger (2013).[20] The Montiel-Pflueger effective F statistics are used to test the null hypothesis of weak instrument under heteroscedasticity and serial correlation. The null hypothesis is rejected if the 2SLS Nagar bias, the bias of an approximating distribution, is greater than a "worst-case" benchmark value. The benchmark value is then set at the τ% of the OLS bias at the 5% significance level where τ can take values of 5 or 10. In my sample, the 5% critical values for testing the null hypothesis that the 2SLS Nagar bias exceeds 10% and 5% of the OLS bias are 23.11 and 37.42, respectively. Comparing the reported effective first-stage F statistics with the critical values, the effective F statistics for both the rainfall IV and the wind direction IV exceed the 10% "worse-case" benchmark bias. That is to say, if the critical value is conservatively set at the 10% "worst-case" benchmark, then the null hypothesis of weak

[20] Andrews, Stock, and Sun (2019) recommend adopting the testing procedure of Montiel Olea and Pflueger (2013) and using the effective F-statistics to gauge instrument strength for a single endogenous regressor case. The Montiel-Pflueger effective F-statistics have the advantages over other methods for detecting weak instruments since they are robust to heteroscedasticity, serial correlation, and panel data clustering.

instrument is rejected at the 5% level for both the rainfall IV and the wind direction IV and both treatment variables. Alternatively, if the critical value is set at the 5% "worst-case" benchmark, the null hypothesis of weak instrument is still rejected for the rainfall IV for both treatment variables, but not rejected for the wind direction IV as the instrument for the PM2.5 pollution classified by the WHO standard.

4.3 2SLS and Identification Assumptions

Panel B of Table 3 presents the results of the second-stage estimation. The 2SLS estimates employing the rainfall IV are negative, while the estimates employing the wind direction IV are positive but close to zero. Specifically, for the PM2.5 pollution classified by the EPA criteria, the effect is estimated at -0.09 percentage points using the rainfall IV and 0.003 percentage points using the wind direction IV. The estimated effect is of similar magnitude for the pollution defined by the WHO standard (-0.12 percentage points using the rainfall IV and 0.004 using the wind direction IV). These estimates are not statistically significant at conventional levels. The 2SLS estimates are smaller than the corresponding OLS estimates, which is consistent with the conjecture that the OLS estimator is positively biased.

For the IV estimator to identify the causal effect of interest, the instrumental variable must satisfy the exclusion restriction and the monotonicity assumption. Formally, consider the potential outcome model (Rubin 1974) where $Y = Y_1 \cdot T + Y_0 \cdot (1-T) \in \Upsilon \subset \mathbb{R}$ and Y_1 and Y_0 are potential outcomes indicating the outcome with or without treatment. Here, Y denotes the prevalence rate of workplace accidents. Let T denote the binary treatment variable measuring the PM2.5 pollution. Denote Z as the binary instrumental variable, where Z could be either the rainfall IV ($Z1$) or the wind direction IV ($Z2$). Let T_z be the potential treatment if the instrumental variable had been exogenously set to $z = 0,1$. Imbens and Angrist (1994) shows that the instrumental variable method identifies the local average treatment effect under the following two assumptions.

Assumption LI. *(LATE Independence): (a) the quadruple (Y_0, Y_1, T_0, T_1) is jointly independent of Z and (b) $P[T=1|Z=0] \neq P[T=1|Z=1]$.*

Assumption LI states that the instrumental variable is correlated to the treatment variable but independent of the potential outcomes and the potential treatments. It implies that the instrumental variable Z can only affect the outcome variable Y through its effect on the treatment variable T, which is the exclusion restriction. Empirically, one can verify the correlation between the treatment variable and the instruments via a first-stage regression.

Assumption LM. *(LATE Monotonicity): Either $T_0 \leq T_1$ almost surely or $T_0 \geq T_1$ almost surely.*

Assumption LM implies the absence of defiers. In my case, it asserts that having more days with heavy rainfall ($Z1 = 1$) or wind blowing from the pre-specified wind directions ($Z2 = 1$) should lead to fewer PM2.5-polluted days ($T = 0$) and vice versa for each grid-and-month-year cell.

4.4 Testing the Validity of Instrumental Variables

The testing of the joint validity of assumptions LI and LM follows Mourifié and Wan (2017). Without loss of generality, assuming $T_0 \leq T_1$, Mourifié and Wan (2017) shows that the testable sharp characterization of the LI and LM assumptions can be written as:

$$\theta(y,1) = E[c_1 T(1-Z) - c_0 TZ | Y = y] \leq 0 \tag{5}$$

$$\theta(y,0) = E[c_0(1-T)Z - c_1(1-T)(1-Z) | Y = y] \leq 0 \tag{6}$$

where $c_k = P(Z = k), k = 0, 1$.[21] In other words, if Assumptions LI and LM hold, equations (5) and (6) hold, and there exists a joint distribution of $(\tilde{T}_0, \tilde{T}_1, \tilde{Y}_0, \tilde{Y}_1, Z)$ such that LM and LI hold and $(\tilde{Y}, \tilde{T}, Z)$ has the same distribution as (Y, T, Z). The null hypothesis that the LI and LM assumptions hold can be formulated as

$$H_0 : \theta_0 = \sup_{v \in V} \theta(v) \leq 0, \quad H_1 : \theta_0 > 0 \tag{7}$$

where $v \in \Upsilon \times \{0,1\}$. The test statistics θ_0 implies the joint validity of the LI and LM assumptions. Moreover, by testing whether θ_0 less or equal to zero, one can test whether the assumptions LI and LM jointly hold for the instrumental variable under testing. Mourifié and Wan propose to adopt the intersection bounds framework of Chernozhukov, Lee, and Rosen (2013) to estimate

[21] In my case, the correlation between Z (the instrumental variables, either wind direction or rainfall) with T (the PM2.5 indicator) is negative, i.e., $T_0 \geq T_1$. The testing procedures of Mourifié and Wan (2017) can be readily applied to this case. See Appendix C for more details.

nonparametric bounds on θ_0 and perform an intersection bounds test.[22] The null hypothesis is rejected if a collection of lower intersection bounds of θ_0 is greater than zero.

Several practical problems emerge when applying this test. First, this paper uses a large panel data. Computing nonparametric bounds using this data requires a significant amount of computing power, which makes it empirically infeasible to conduct the testing using the whole sample. A practical solution to deal with this issue, as suggested by the applications in Mourifié and Wan (2017), is splitting the sample and separately conducting the test on each sub-sample. Specifically, I divide the sample into $S = 2770$ sub-samples and each sub-sample s consists of data spanning the 48 month-year for each grid. The null hypothesis H_0 then can be defined as $H_0 = H_0^{(1)} \cap H_0^{(2)} \cap \cdots \cap H_0^{(S)}$, where $H_0^{(s)}$ suggests that the moment inequalities hold for sub-sample s. Rejection of any $H_0^{(s)}$ implies the rejection of H_0. Second, because of the seasonality pattern of the workplace accidents and the existence of confounders (either time-specific or grid-specific), it is more plausible to assume that the instrumental variable is independent of the potential outcomes and potential treatments conditional on covariates. That is, for covariates X with support χ, $(Y_0, Y_1, T_0, T_1) \perp Z | X = x$ and $P[T = 1 | Z = 0, X = x] \neq P[T = 1 | Z = 1, X = x], \forall x \in \chi$. As shown in Mourifié and Wan (2017), under conditional independence, for all $(x, y) \in \chi \times \Upsilon$, inequalities (5) and (6) become:

$$\tilde{\theta}(x, y, 1) = E[c_1(x)T(1 - Z) - c_0(x)TZ | Y = y, X = x] \leq 0 \tag{8}$$

$$\tilde{\theta}(x, y, 0) = E[c_0(x)(1 - T)Z - c_1(x)(1 - T)(1 - Z) | Y = y, X = x] \leq 0 \tag{9}$$

where $c_k(x) = P(Z = k | X = x)$. Covariates X can consist of multiple variables. For continuous X, they can be incorporated as additional conditioning variables. To accommodate discrete or categorical covariates in the testing, it involves dividing the sample according to the value of covariates and evaluating inequalities (8) and (9) on each sub-sample. For the setting of this paper, the desired covariates X should at least include weather conditions and month by year fixed effects to control for the seasonality pattern or time-specific confounders. However, it is empirically infeasible to simultaneously consider the weather controls and the grid-specific and time-specific confounders. Because to do this, the sample needs to be divided by grid, time, and

[22] Chernozhukov et al. 2015 provides a Stata package for estimation and inference on intersection bounds. Specifically, the intersection bounds test is conducted by the *clrtest* command.

covariates dummies, and the size of the sub-sample may become so small that the distribution of the instrumental variable Z or the conditioning variable Y degenerates to a masspoint, making the sub-sample untestable. Since the null hypothesis for the whole sample H_0 is defined as the intersection of the null hypothesis for the sub-samples H_0^s, one would fail to test the null hypothesis for the whole sample if the vast majority of sub-samples is untestable.

To implement the testing procedure in the scenario of multiple categorical covariates, I propose a pre-testing estimation. In the first stage, I run a fixed-effects regression of the accident rate Y with grid and month by year fixed effects, and obtain the residuals $\hat{u}$.[23] Residuals $\hat{u}$ can be viewed as the adjusted accident rate where grid-specific and time-specific factors are partialled out.[24] Instead of conditioning on Y and multiple covariates X, I use $\hat{u}$ as the conditioning variable and compute the intersection bounds. Conditioning on a single variable $\hat{u}$ accelerates the testing process and provides a flexible way to accommodate the categorical covariates in testing.[25]

Because the testing procedure is applied to a large number of sub-samples, it is important to control the family-wise error rate (FWER) at targeted levels. Following Mourifié and Wan (2017) and Holm (1979), I adopt a multiple testing procedure. For each test on the sub-sample s, I set the marginal significance level at which $H_0^{(s)}$ is rejected. Denote α^* as the lowest significance level at which $H_0^{(s)}$ is rejected, then H_0 is rejected with the FWER controlled by no greater than $\alpha^* \times S$.[26]

I summarize the testing results for the treatment variable $T2$ which measures PM2.5 pollution classified by the EPA criteria in Figure 3. For each sub-sample, the testing is conducted at the significance level of 0.003%, which gives the FWER at 8.3%.[27] Figures 3a and 3b plot the testing results in each grid for the rainfall IV and the wind direction IV, respectively. Green flags grids

[23] I run the following fixed-effects regression using the grid-day panel data:

$$Y_{im} = \gamma_i + X_{im} + \tau_{my} + u_{im}$$

where Y_{im} denotes the accident rate and X_{im} are a vector of weather controls including monthly average temperature, wind speed, etc. γ_i and τ_{my} represent the grid and month by year fixed effects, respectively.

[24] I demean the estimated residuals by subtracting their sample mean.

[25] See Appendix C for more technical details on the implementation of this test.

[26] For example, if the sample is divided into $S = 2770$ grid, and α^* is set at 0.003%, then H_0 is rejected with the FWER controlled by no more than $0.003\% \times 2770 = 8.3\%$.

[27] It is an *ad hoc* choice to set the significance level at 0.003%. I choose 0.003% for the 2770 sub-samples to have a FWER around a conventional 10% level. I show in the Appendix Figure A.5 that controlling the FWER at around 5% does not change the conclusion that the null hypothesis of satisfying the joint assumptions LI and LM is rejected for both instruments.

where the null hypothesis is not rejected, while red flags grids where the null hypothesis is rejected. Hatched grids are those not in the sample due to missing values, and white grids indicate that no test is conducted, either because the outcome distribution reduces to a masspoint of zero or because the distribution of the IV indicator is degenerate. Figure 3 shows that for the treatment variable $T2$, the point identification assumptions are violated in most of grids for either the rainfall or the wind direction IVs (or both).[28]

In summary, the rejection of $H_0^{(s)}$ for some s leads to the rejection of H_0, i.e., the null hypothesis that the point identification assumptions are satisfied for the tested IV for the whole sample. The violation of the joint assumptions LI and LM for the two instrumental variables renders them invalid IVs.

4.5 Conjecture of What Causes the Violation

The point identification assumptions may fail to hold for either the rainfall IV or the wind direction IV for several reasons. The exclusion restriction may be violated for the rainfall IV because the rainfall itself can cause workplace accidents independently of PM2.5 pollution. Additionally, workers may actively adjust their intertemporal labor supply on rainy days to avoid difficulties of commuting to work during the bad weather. The weather-related labor shortage could be correlated to the probability of workplace accidents. Besides, outdoor jobs are more likely to be rescheduled on days with rainy weather conditions and thus outdoor workers may be asked by employers to work shorter hours or even to not work on days with bad weather, which passively adjusts their labor supply. Such adjustments are more salient for workers who have flexible work schedule, who work outdoor, and in extreme weather days.[29]

[28]As shown in Figures A.4, I find similar testing results that the joint assumptions LI and LM are violated for the treatment variable $T1$.

[29]Because the data used in this study is not at the individual level, workers' intertemporal labor supply adjustment with respect to weather is unobserved. In Appendix D, I investigate whether workers' intertemporal labor supply changes with the level of precipitation, using the American Time Use Survey (ATUS) 2005-2019. As shown in Table B.5, while I do not find statistically significant effect of rainfall on the time spent in work-related activities for employed individuals, I find that outdoor workers, especially male workers, work about 20 minutes less on rainy days. Moreover, I find suggestive evidence that the rainfall is positively associated with work absenteeism: having heavy rainfall in the past week is associated with an increase in the probability of work absence in the same week of about 0.9 percentage points for outdoor workers. See Appendix D for more details.

For the wind direction IV, it may appear intuitive that compared to the rainfall, the wind direction itself is less correlated with the workplace accidents and the labor supply. However, the direction of the impact of wind on air quality may be ambiguous: wind can either bring air pollutants to a certain area and increase the pollution level, or blow away pollutants and improve the air quality. In this case, the assumption LM that is jointly tested with LI may fail for some grid-month-year cells. More specifically, recall the two-stage process of constructing the wind direction IV: It relies on a pre-estimation procedure that finds the prevailing wind directions that give the smallest monthly average PM2.5 for each month-grid cell. Then, on the second-stage, the wind direction IV is defined as an indicator variable that signifies a greater number (e.g., greater than the 75th percentile) of days with wind blowing from the wind directions found in the first-stage. Therefore, the assumption LM states that at each grid-month-year cell, having a greater number of days with wind blowing from the pre-specified directions must lead to a smaller number of air-polluted days and vice versa. The LM assumption is violated when there exists some cells which have a great (small) number of wind-blowing days while their numbers of air-polluted days still pass (do not pass) the pre-specified threshold of being treated. These groups are the so-called "defiers". Because the first-stage estimation detects the wind directions associated with a lower level of monthly average PM2.5 pooling over days and years, it is possible that on some days, the wind blowing from the directions specified in the first-stage does not blow away PM2.5 but, at the extreme case, bring pollutants to the place. In this circumstance, a greater (smaller) number of days in which winds blowing from the directions specified in the first-stage would lead to a greater (smaller) rather than a smaller (greater) number of air-polluted days, which generates defiers.

5 Partial Identification Strategy

In Section 4.4, I found statistical evidence that the point identification assumptions fail to hold for both the rainfall IV and the wind direction IV, making it infeasible to point identify the causal impact of air pollution on workplace accidents. Therefore, I turn to a partial identification strategy which exploits the exogenous variability of instrumental variables, accounting for their invalidity. This method replaces the exclusion restriction for point identification with weaker assumptions and instead of obtaining a point estimate, it estimates bounds within which the true effect lies. I start with the approach by Nevo and Rosen (2012), a method based on a linear parametric model assuming constant treatment effects. The constant treatment effect assumption is a strong assumption in my empirical setting. For this reason, I turn to a recent nonparametric strategy proposed by Ban and Kédagni (2022) which extends Nevo and Rosen (2012) to allow for heterogeneous treatment effects.

5.1 Parametric Bounds: Nevo and Rosen (2012)

The treatment effect of interest in the Nevo and Rosen (2012) method is the average treatment effect ATE expressed as ATE $= E[Y_1] - E[Y_0]$. In the circumstance that the instrumental variable Z is invalid, Nevo and Rosen (2012) establish sharp bounds on ATE under the following assumptions.

Assumption A1. *(SDC - Same Direction of Correlation):* $Cov(Y_t, T)Cov(Y_t, Z) \geq 0,\ \forall\, t \in \{0,1\}$.

where t denotes the treatment status. Assumption SDC states that the direction of the correlations between the treatment T and the potential outcomes Y_1 and Y_0 are the same as the correlation between the invalid instrument Z and Y_t.[30]

Assumption A2. *(LEI - Less Endogenous Instrument):* $|\rho_{Y_t,T}| \geq |\rho_{Y_t,Z}|,\ \forall\, t \in \{0,1\}$.

where ρ denotes the correlation. Assumption LEI asserts that the invalid instrument Z is weakly less correlated with the potential outcomes than is the endogenous treatment. Intuitively, it argues that, although the instrumental variable Z is invalid, it is less endogenous than the treatment.

The main idea of Nevo and Rosen (2012) is to construct an instrument, denoted as V, that satisfies the moment condition $E[V \cdot Y_t] = 0$ using the invalid IV and the treatment variable. In other words, the constructed V is uncorrelated to the potential outcome. Since the constructed IV is empirically unobserved, the IV estimator employing V as the instrumental variable cannot point identify the average treatment effect of interest. But as long as V can be bounded by some finite values, Nevo and Rosen (2012) shows that it is feasible to partially identify the ATE. Specifically,

[30] Ban and Kédagni (2022) shows that Assumption SDC can be seen as a weaker version of the monotone IV (MIV) and monotone treatment selection (MTS) assumptions employed by Manski and Pepper (2000).

consider the ratio of the correlation of the potential outcome Y_t with Z and T, defined as

$$\lambda^* \equiv \begin{cases} \rho_{Y_t,Z}/\rho_{Y_t,T} & \text{if} \quad \rho_{Y_t,T} \neq 0 \\ 0 & \text{if} \quad \rho_{Y_t,T} = \rho_{Y_t,Z} = 0 \end{cases}$$

Under Assumptions SDC and LEI, it can be shown that

$$E[(\sigma_T Z - \lambda^* \sigma_Z T)Y_t] = 0 \tag{10}$$

where σ_T and σ_Z are the standard deviation of T and Z, respectively. Equation (10) suggests that there exists a λ^* to construct a weighted average of T and Z that is uncorrelated to the potential outcome. Define a function $V(\lambda)$

$$V(\lambda) = \sigma_T Z - \lambda \sigma_Z T$$

such that $E[V(\lambda^*) \cdot Y_t] = 0$. $V(\lambda^*)$ is the constructed valid IV (V) that satisfies the moment condition $E[V \cdot Y_t] = 0$. If λ^* were known, $V(\lambda^*)$ could be used to point identify the ATE. While λ^* is empirically unknown, Assumptions SDC and LEI imply that λ^* is bounded between 0 and 1, i.e., $\lambda^* \in [0,1]$. Bounding λ^* is equivalent to bounding the constructed valid IV — $V(\lambda^*)$ — and hence the ATE. Plugging the boundary values of λ, i.e., $\lambda = 0$ and $\lambda = 1$, into Equation (10) yields $V(0) = \sigma_T Z$ and $V(1) = \sigma_T Z - \sigma_Z T$. Nevo and Rosen (2012) shows that the ATE is bounded by the two IV estimators that separately employ $V(0)$ and $V(1)$ as the instrumental variable. I denote the IV estimators θ_{IV}^{Z}, $\theta_{IV}^{V(0)}$, and $\theta_{IV}^{V(1)}$ as the IV estimators employing the instrumental variable Z, $V(0)$, and $V(1)$, respectively. It is intuitive to see that $\theta_{IV}^{V(0)} = \theta_{IV}^{Z}$. Whether $\theta_{IV}^{V(0)}$ and $\theta_{IV}^{V(1)}$ are the upper or the lower bound is determined by the sign of the correlation of the treatment variable with the invalid IV and the sign of the correlation of the potential outcome Y_t with T and Z. Proposition 1 shows the bounds derived by Nevo and Rosen (2012):

Proposition 1. *(Nevo and Rosen, 2012) Given assumptions SDC and LEI, if $Cov(T,Z) < 0$, then there exists a two-sided bound where $\theta_{IV}^{Z} \leq ATE \leq \theta_{IV}^{V(1)}$ if $Cov(Y_t,T), Cov(Y_t,Z) \geq 0$ or $\theta_{IV}^{V(1)} \leq ATE \leq \theta_{IV}^{Z}$ if $Cov(Y_t,T), Cov(Y_t,Z) \leq 0$. Instead, if $Cov(T,Z) > 0$, then there exists a one-sided bound where $ATE \leq \min\{\theta_{IV}^{Z}, \theta_{IV}^{V(1)}\}$ if $Cov(Y_t,T), Cov(Y_t,Z) \geq 0$ or $ATE \geq \max\{\theta_{IV}^{Z}, \theta_{IV}^{V(1)}\}$ if $Cov(Y_t,T), Cov(Y_t,Z) \leq 0$. These bounds are sharp.*

Proposition 1 states that when the endogenous treatment variable T is negatively correlated with the invalid IV, the average treatment effect can be bounded between the invalid IV estimator θ_{IV}^{Z} and the constructed IV estimator $\theta_{IV}^{V(1)}$.[31] When the treatment variable and the invalid IV are positively correlated, the method of Nevo and Rosen (2012) yields a one-sided bound. The direction of the one-sided bound is determined by the sign of $Cov(Y_t, T)$ and $Cov(Y_t, Z)$.

I will discuss the validity of the partial identification assumptions SDC and LEI before presenting the corresponding estimated bounds in Section 6. Empirically, Nevo and Rosen's method is easy to implement and can be readily applied to the panel data, accommodating covariates and (high dimensional) fixed effects. However, these advantages are at the expense of imposing the constant treatment effect assumption. Moreover, this method assumes a linear parametric specification.

5.2 Nonparametric Bounds: Ban and Kédagni (2022)

Noticing the limitations of Nevo and Rosen (2012), I then turn to the more flexible method of Ban and Kédagni (2022), which extends the method of Nevo and Rosen (2012) to a nonparametric setting. The nonparametric bounds method of Ban and Kédagni (2022) also assumes SDC and LEI as in Nevo and Rosen (2012). Besides, they impose two additional assumptions:

Assumption A3. *(BOS - Bounded Support):* $Supp(Y_t|T \neq t) = Supp(Y_t|T = t) = [\underline{y}_t, \bar{y}_t]$, $\forall\ t \in \{0,1\}$.

The BOS assumption ensures the same support of the counterfactual outcome and the factual outcome. It is a standard assumption when estimating nonparametric bounds (Manski 1990; Manski and Sims 1994). By imposing the bounded support assumption, Ban and Kédagni (2022) relaxes the parametric linear assumption in Nevo and Rosen (2012) and obtains two-sided bounds regardless of whether the correlation of treatment T with the invalid IV is negative or not.[32] Moreover, they show that if Assumption BOS holds for the outcome variable, the SDC assumption has a testable implication: SDC is rejected if the nonparametric bounds defined under BOS and SDC are empty. In practice, the rejection of SDC suggests the data using for analysis is not compatible with the identification assumptions for the partial identification strategy in use.

[31] Proposition 1 can be generalized to the model allowing multiple regressors or multiple invalid IVs. For the proof of Proposition 1 and bounds in other cases, see Nevo and Rosen (2012).

[32] Recall that Nevo and Rosen (2012) provides two-sided bounds when $Cov(T,Z) < 0$ and one-sided bound when $Cov(T,Z) > 0$.

Assumption A4. *(MTR - Monotone Treatment Response):* $Y_t \geq Y_{t'}$, $\forall\, t > t'$.

Assumption MTR claims that the potential outcome weakly increases with the level of the treatment (Manski 1997). That is, the treatment effect for each sample unit is no less than zero. As shown below in Section 6.2, adding the MTR assumption helps tighten the bounds. This assumption will be justified in the context of my empirical setting in Section 6.

To derive nonparametric bounds on the ATE, Ban and Kédagni (2022) first identify bounds on the potential outcome expectation $\theta_t = E[Y_t]$ for $t \in \{0,1\}$ such that $LB_t \leq \theta_t \leq UB_t$, and then obtain the bounds on the average treatment effect ATE as $LB_1 - UB_0 \leq \text{ATE} \leq UB_1 - LB_0$. The derivation of bounds on θ_t shares a similar intuition with Nevo and Rosen (2012), that is, leveraging a weighted average of the treatment variable and the invalid instrument to identify θ_t, and then bounding the constructed weighted average to obtain bounds on θ_t. For example, the derivation of nonparametric bounds under the BOS and SDC assumptions involves moment inequalities $E[\delta(\alpha,\beta)Y_t] \geq \theta_t$ where $\delta(\alpha,\beta) = 1 + \alpha(\beta(T - E[T]) + (1-\beta)(Z - E[Z]))$. Like $V(\lambda)$ of the parametric method, $\delta(\alpha,\beta)$ is an nonparametric analogue of the weighted average of T and Z with parameters α and β. Additional steps are needed to formally derive the upper bound estimator of θ_t, but the main idea is bounding the moment function of the weighted average $\delta(\alpha,\beta)$.[33]

Lastly, the nonparametric bounds in Ban and Kédagni (2022) involve maximum and minimum operators. The regular inference procedure to estimate the confidence regions for bounds containing maximum and minimum operators suffers from bias in finite samples. Because the concavity and convexity, respectively, of the minimum and maximum operators cause the sample analog estimates of bounds narrower relative to the corresponding true bounds. To obtain the correct confidence regions, I follow Ban and Kédagni (2022) and implement the estimation and the valid-inference procedure proposed by Chernozhukov, Lee, and Rosen (2013). The bounds estimators in Chernozhukov, Lee, and Rosen (2013) are half-median unbiased in a sense that the estimated lower (upper) bound falls below (above) the true lower (upper) bound with a probability of at least one-half asymptotically. The resulting confidence intervals are valid for the true parameter of interest.

[33]See Ban and Kédagni (2022) for a formal derivation.

Implementation

To apply Ban and Kédagni's method to my panel data, I adopt the following steps:

(i) First, I run a panel regression and obtain the residuals partialling out the fixed effects and covariates (the time-varying weather controls). Then, I demean the estimated residuals by subtracting their sample mean.

(ii) Second, to facilitate computation, I divide the panel data by grid and obtain 2770 sub-samples.[34]

(iii) For each grid j, I follow Ban and Kédagni (2022) and use the Stata command *clr3bound* (Chernozhukov et al. 2015) to estimate the nonparametric bounds and their 95% two-sided confidence set on the demeaned residual. Let's denote the estimated bounds under different assumptions as $LB_j^s(t) \leq \text{ATE} \leq UB_j^s(t)$, for each treatment $t \in \{0,1\}$, sub-sample j, and assumption set $s \in \{\text{LEI, MTR}\}$.[35] Then the bounds on the average treatment effect of the sub-sample can be obtained as $LB_j^s(1) - UB_j^s(0) \leq \text{ATE}_j^s \leq UB_j^s(1) - LB_j^s(0)$.

(iv) ATE_j suggests the heterogeneous effect of PM2.5 on the workplace accident rate for each grid. One can obtain bounds on the average treatment effect for the whole sample by taking the average of the estimated bounds over all grids.

6 Bounds Estimation

6.1 Evidence From Parametric Bounds

Assumptions

Before proceeding to the results using the parametric bounds, I discuss the validity of the SDC and LEI assumptions. Both assumptions impose restrictions on the correlations of the potential

[34]This is the same sample division process as in the testing for the IV validity.
[35]For brevity, I use LEI and MTR to denote the assumption set of BOS + SDC + LEI and BOS + SDC + LEI + MTR.

outcomes with the treatment variable and the instrumental variables, i.e., $Cov(Y_t, T)$ and $Cov(Y_t, Z)$, respectively. I argue that SDC and LEI likely hold for both the rainfall IV and the wind direction IV.

First of all, it is likely that the correlation of PM2.5 pollution with the latent workplace accident rate is weakly positive and hence $Cov(Y_t, T) \geq 0$. Better air quality (less air pollution) is naturally linked with safer workplaces (fewer workplace accidents), because places (grids) that value clean air and are willing to invest money in air quality improvement (usually places with higher income level) tend to also value work safety and are willing to invest money in creating safer workplaces. At the same time, places with more intensive industrial activities are more likely to simultaneously experience a higher level of air pollution and a higher rate of workplace accidents, assuming the probability of workplace accidents is positively correlated to the intensity of production activities.

Moreover, places with poorer air quality would have a higher rate of workplace accidents if workers and employers in these places are more cautious and thus more responsive in adjusting their intertemporal labor supply and work schedule with changes in air quality. Existing literature has documented that workers reduce their intertemporal labor supply on air-polluted days (for example, Hansen and Selte 2000; Currie et al. 2009; Hanna and Oliva 2015).[36] Although being absent from work reduces the probability of workplace accidents for the worker on leave, work absenteeism is likely positively correlated to the rate of accident at the grid-month level for two reasons: First, on the day(s) the worker being absent from work, the workload of coworkers increases. Second, on days after the absent-from-work day(s), the worker may be asked by the employer to work longer time in order to compensate the productivity loss on previous days, which may weakly increase the likelihood of accidents on the succeeding days and hence the overall monthly accident rate. Given all these factors, it is likely $Cov(Y_t, T) \geq 0$.

Next, I consider the correlation between the latent workplace accident rate and the instrumental variables. For the rainfall IV, heavy rainfall itself can increase the likelihood of workplace accidents.

[36]Several studies document a negative relationship between air pollution and labor supply. An increase in air pollution is found to be associated with an increase in work loss days (Hausman, Ostro, and Wise 1984; Ostro 1983), sick-leaves (Hansen and Selte 2000), and sick-related absenteeism (Pönkä 1990), while the decline in air pollution leads to an increase in working hours (Hanna and Oliva 2015). Moreover, air pollution can be associated with school absences (Currie et al. 2009; Ransom and Pope III 1992), implying that pollution may also have real effects on work hours if it induces absenteeism among workers who are responsible for child care (Hanna and Oliva 2015).

Moreover, workers and firms may respond to heavy rainfalls in terms of their labor supply, in a similar way with the change of air quality. Particularly, workers may reduce their labor supply on rainy days.[37] Therefore, the weather-driven work absenteeism would also be positively correlated to the grid-month level workplace accident rate for similar reasons. It implies that the correlation between the workplace accidents and the rainfall IV would share the same sign with the correlation between the workplace accidents and the PM2.5 pollution. Furthermore, because it is less likely for wind direction to directly affect workplace accidents (as discussed in Section 4.5), the correlation between wind direction and potential outcomes can only through its impact on PM2.5. Hence, it is intuitive that $Cov(Y_t, Z2)$ and $Cov(Y_t, T)$ share the same sign and the SDC assumption is satisfied for the wind direction IV. In summary, the same direction of correlation (SDC) assumption likely holds for both the rainfall and wind direction IVs.

As for the less endogenous instrument (LEI) assumption, it is plausible that, compared to wind direction, PM2.5 has a stronger correlation with the workplace accidents because, as previously discussed, wind direction is likely directly uncorrelated with workplace accidents. For the rainfall instrument, although rainfall (especially heavy rainfall) may directly affect workplace accident by affecting, for example, the visual condition at the workplace, its link with workplace accidents is likely not as strong as the link between PM2.5 and workplace accidents. First of all, the effect of rainfall on workplace accidents is likely limited to (semi-)outdoor workers who work on rainy days, while PM2.5 may have a broader effect for both outdoor and indoor workers as PM2.5 can easily penetrate indoors. In addition, it is unlikely for the rainfall to affect workers' cognitive performance, a biological channel through which PM2.5 pollution affects workplace accidents. Furthermore, air pollution is highly correlated to economic activities which simultaneously affect the incidence of workplace accidents, while the link between economic activities and rainfall is in general much weaker. Given all of these factors, it is plausible to assume that the rainfall is less correlated with the workplace accidents than is the PM2.5 pollution. Therefore, the less endogenous instrument (LEI) assumption likely holds for both the rainfall and wind direction IVs.

[37]The preliminary analysis in the Appendix D provides empirical evidence that workers adjust their work schedule on rainy days. Specifically, workers are found to reduce working hours and are more likely to be absent from work on days with heavy rainfall.

Results

Given that Assumptions SDC and LEI are likely satisfied, I proceed to estimate the Nevo and Rosen (2012) bounds for the effect of PM2.5. Panel A of Table 3 suggests that the treatment variables are negatively associated with both the rainfall and the wind direction IVs, hence $Cov(T,Z1) < 0$ and $Cov(T,Z2) < 0$. Thus, Proposition 1 implies two-sided bounds from both IVs. Because $Cov(Y_t,T), Cov(Y_t,Z1), Cov(Y_t,Z2) > 0$, the ATE is bounded between β_{IV}^{Z1} (the lower bound) and $\beta_{IV}^{V_1(1)}$ (the upper bound) employing the rainfall IV, and between β_{IV}^{Z2} (the lower bound) and $\beta_{IV}^{V_2(1)}$ employing the wind direction IV. $V_1(1)$ and $V_2(1)$ denote the upper bound of the constructed valid IV from the imperfect instruments $Z1$ and $Z2$, respectively.

I present in Table 4 the estimated bounds on the average treatment effect of PM2.5 pollution and the 95% and 90% confidence intervals of the bounds for the rainfall IV and wind direction IV.[38] I show in Panel A the estimated bounds employing the rainfall IV and Panel B the estimated bounds employing the wind direction IV. Columns (1) and (2) report the lower bound and the upper bound on the average treatment effect of PM2.5 pollution classified by the EPA standard, while Columns (3) and (4) report the estimated bounds on the average treatment effect of PM2.5 pollution defined by the WHO standard.

Columns (1) and (2) of Panel A shows that the estimated bounds for PM2.5 pollution defined by the EPA standard are between -0.09 and 0.09 percentage points, employing the rainfall IV. The estimated bounds include zero but do rule out negative effects of less than -0.09 percentage points (equivalent to a -0.9% decrease relative to the sample mean workplace accident rate), and positive effects of greater than 0.09 percentage points (equivalent to an increase of 0.9% relative to the sample mean workplace accident rate). As shown in Columns (1) and (2) of Panel B, the estimated PM2.5 effect are bounded between 0.003 and 0.09 percentage points employing the wind direction IV, suggesting the effect of PM2.5 pollution (classified by the EPA standard) is an increase of at least 0.03% and up to 0.9% relative to the sample mean accident rate. However, the estimated bounds cannot statistically rule out a zero effect in that their 90% confidence intervals include zero.

Columns (3) and (4) present the estimated bounds for PM2.5 pollution based on the WHO

[38] I present alternative results of treatment variables defined by a different threshold (i.e., the median) in the Appendix Table B.3. The choice of the cutoff does not significantly change the estimation results.

standard. Similar to the results in Columns (1) and (2) of Panel A, employing the rainfall IV, the estimated bounds in Columns (3) and (4) of Panel A contain zero. The lower bound is estimated at -0.12 percentage points while the upper bound is estimated at 0.02 percentage points. They rule out negative effects of less than -1.2% and positive effects greater than 0.2%. Using the wind direction IV, as shown in Columns (3) and (4) of Panel B, the estimated PM2.5 effect is bounded between 0.004 and 0.02 percentage points, suggesting PM2.5 increases the rate of workplace accidents by at least 0.04% and up to 0.2% relative to the sample mean accident rate. These bounds are not precisely estimated because their 90% confidence intervals do not exclude zero.

The method of Nevo and Rosen (2012) imposes a strong assumption of constant treatment effect, which is likely violated in the empirical setting of this paper. One of the reasons why the effect of air pollution on workplace accidents may be heterogeneous is that the magnitude of effect is likely to be affected by the level of exposure, and workers' exposure to air pollution may vary by their work environment. In addition, the estimation builds on a linear parametric model. The consistency of the estimator can be sensitive to model misspecification. In the next section, I provide evidence from a more flexible method which relaxes the constant treatment effect assumption and estimates nonparametric bounds on the treatment effect of interest.

6.2 Evidence From Nonparametric Bounds

In this section, I present and discuss results of nonparametric bounds estimated by the method of Ban and Kédagni (2022). I estimate bounds under two sets of assumptions: a) BOS, SDC and LEI, discussed in Section 6.1, and b) BOS, SDC, LEI, and MTR.

Assumptions

The BOS assumption is satisfied because intrinsically the number of workplace accidents is zero or a positive integer and thus the outcomes are bounded in the right tail. Researchers can conservatively top-code the sample in the right tail to further ensure the outcome support is bounded. Assumption MTR states that the potential outcome weakly increases with the level of the treatment. In the context of this paper, because the analysis is at the grid-and-month-year level, MTR states that exposure to a higher number of PM2.5-polluted days in a month does not *decrease* workplace accidents. It is plausible to assume MTR at the grid-month level because the adverse impact of PM2.5 on human cognitive functions provide a channel through which PM2.5 may increase the workplace accidents. Even though workers can respond to the air quality deterioration by calling in sick on the air polluted days, it has a spillover effect on coworkers. Coworkers may be subjected to a higher hazard of workplace accidents because work absenteeism may increase the workload of coworkers, making them exposed to extra stress and fatigue. The probability of accidents at the workplace is the same or higher for workers who are fatigued, discomforted, or experience cognitive decline.[39] For these reasons, the overall effect of PM2.5 pollution at the grid-and-month-year level is unlikely be negative, making the MTR assumption is plausible.

Results

I summarize the estimated nonparametric bounds employing either the rainfall IV or the wind direction IV for the whole sample in Table 5. I report the estimated lower and upper bounds and their 95% confidence intervals. Panels A and B report the nonparametric bounds on the average treatment effect of PM2.5 pollution under two sets of assumptions: BOS + SDC + LEI and BOS + SDC + LEI + MTR, respectively. I present the estimated bounds for PM2.5 pollution defined according to the EPA criteria in Columns (1) and (2), and the estimated bounds for PM2.5 pollution defined by the WHO standard in Columns (3) and (4).

As shown in Panel A, under assumption BOS + SDC + LEI, the estimated bounds include zero for both treatment variables. Nevertheless, the estimated bounds still provide information about the interval in which the true effect falls. For example, the estimated bounds in Columns (3) and (4) of Panel A rule out negative effects of less than -14 percentage points and positive effects of greater than 18 percentage points, employing the rainfall IV for the PM2.5 pollution according to the WHO standard. Employing the wind direction IV, the bounded effect is estimated between -14 and 17 percentage points. Similar results are found for the PM2.5 pollution defined by the EPA standard. These estimated nonparametric bounds are fairly wide compared to parametric bounds shown in

[39]Studies on the risk factors of workplace injuries have found that fatigued workers are exposed to a higher risk of injuries at workplaces (Swaen et al. 2003; Tucker 2003).

Table 4. It is likely due to relaxing the parametric assumption which tightens the previous bounds.

Imposing the MTR assumption helps considerably in tightening the nonparametric bounds. Panel B presents the estimated bounds under assumptions BOS + SDC + LEI + MTR. Results in Columns (1) and (2) of Panel B imply that the true effect of PM2.5 pollution (defined by the EPA standard) is between 0.54 and 3.93 percentage points employing the rainfall IV and between 0.68 and 3.92 employing the wind direction IV. It suggests that the estimated effect of PM2.5 pollution is at least 5% (7%) and up to 39% (39%) relative to the sample mean accident rate, employing the rainfall (wind direction) IV. The estimated bounds are of similar magnitude for PM2.5 pollution defined by the WHO standard. As shown in Columns (3) and (4) of Panel B, the lower bound is estimated at 0.56 percentage points employing the rainfall IV, equivalent to a minimal effect of about 6% relative to the sample mean accident rate, and 0.49 percentage points employing the wind direction IV, equivalent to a minimal effect of about 5% relative to the sample mean accident rate. The upper bound is estimated at about 4.02 percentage points using the rainfall IV, equivalent to a maximal percentage effect of about 40% relative to the sample mean accident rate, and 4.59 percentage points using the wind direction IV, equivalent to a maximal percentage effect of about 46% relative to the sample mean accident rate. However, these bounds are not precisely estimated, as the 95% confidence intervals of the estimated bounds include zero. Thus, I cannot statistically exclude a zero effect.

Under the monotone treatment response assumption, the 95% confidence intervals of the estimated bounds contain the OLS estimates using the same treatment variable (those in Table 2), although the estimated bounds exclude and lie at the right of the OLS estimate. Similar patterns apply to the 2SLS estimates employing the wind direction IV (see Columns (2) and (4) in the Panel B of Table 3): the point estimates of 2SLS are included in the 95% confidence region of the estimated bounds but lie to the left of the estimated bounds. On the other hand, the estimated bounds and their 95% confidence region exclude the 2SLS point estimates employing the rainfall IV.

Furthermore, the estimated bounds are comparable to point estimates of the effect of other air pollutants and the effect of other environmental factors. For example, Depalo and Palma (2020) finds evidence from Italy that a ten-unit increase in the air quality index (AQI), which captures air pollutants PM10, CO, NO_2, and SO_2, increases the number of workplace accidents by 8.6%, equivalent to an increase of 0.22 accidents. To compare the estimated bounds in Table 5 to the estimated effect of Depalo and Palma (2020), I transfer the effect on the accident rate to changes in the number of accidents. For an average grid with population size at the sample mean ($112,291$), the maximal percentage effect of 46% is equivalent to about an additional 0.52 accidents, while the minimal percentage effect (5%) is equivalent to an increase of 0.06 accidents.[40] I find the estimated bounds contain the effect found in Depalo and Palma (2020). In addition, using data from California, Park, Pankratz, and Behrer (2021) finds that a day with maximum temperature between 85F and 90F increases the injury claims by 0.026 (about 4.8% relative to the baseline mean). The effect appears to be larger for maximum temperature ranging from 100F to 105F, which is estimated at 0.036 (equivalent to a relative increase of 6.6%). Under the assumption that one accident is associated with one injury, the estimated least effect of PM2.5 pollution is twice as large as the effect of higher temperature found in Park, Pankratz, and Behrer (2021). However, because the estimated bounds in this paper capture the cumulative impacts of PM2.5 pollution on workplace accidents at the monthly level, it is plausible that the estimated lower bound of PM2.5 effect is greater than the effect of one day exposure to higher temperatures. The average number of days with maximum temperature between 85F and 90F per month in the sample of Park, Pankratz, and Behrer (2021) is about 2.4. Therefore, transforming their estimates of the daily effect to the monthly level ($= 0.026 \times 2.4$)), I obtain a point estimate of about 0.062, which is close to my estimated lower bound on the PM2.5 effect.

Heterogeneous Effects

In this subsection, I explore the heterogeneous effects of air pollution on the workplace accident rate. Since the estimated bounds are comparable for both the rainfall and the wind direction IVs, I only present the results of bounds employing the wind direction IV. Additional results can be

[40]These numbers are calculated by multiplying the maximal/minimal percentage effect 46%/5% by the sample mean population 112,291 and the sample mean workplace accident rate 0.1, and dividing by the scale 10,000, i.e., $[46\% * 0.1 * 112,291]/10,000 \approx 0.52$; $[5\% * 0.1 * 112,291]/10,000 \approx 0.06$

found in the Appendix Figures A.6-A.7. I start by exploring how the effects vary across industry groups. I classify accidents into 11 industry sectors by 2-digit NAICS codes following the BLS's classification. The bounds estimation is then conducted separately for each industry category. Because the frequency of accidents in certain industries, such as financial activities, education and health, and leisure and hospitality, is very low in the sample (see Appendix Table B2), it is infeasible to estimate nonparametric bounds separately for these industries. Since most of jobs in those industries are conducted indoor, I aggregate the workplace accidents in those industries and denote it as "Aggregated Indoor". Moreover, I combine the other service and public administration industries into one group, denoted as "Other Services & Public Administration".

Figure 4 summarizes the estimated bounds and their 95% confidence intervals on the effect of PM2.5 pollution (defined by the WHO standard) by industry. There is suggestive evidence that the effects may be larger for industries where workers work in outdoor or semi-outdoor environment given that their estimated bounds are shifted towards the right. For example, the estimated upper bounds suggest the effect is at most 1.2 percentage points for natural resources and mining, 0.94 percentage points for construction, and 0.76 percentage points for transportation. In contrast, the suggested effect is at most 0.31 percentage points for public administration, 0.21 percentage points for aggregated indoor industries, and 0.38 percentage points for professional and business services. One exception is the manufacturing sector, where the effect is estimated to be at least 0.15 and up to 1.17 percentage points. The work context in natural resources and mining (including agriculture, forestry, and fishing) and construction sectors generally involves working in outdoor or semi-outdoor environments, while workers usually work indoor for the manufacturing sector. However, manufacturing plants may be located near where air pollution is the worse. In addition, the suggestive larger effects found for the manufacturing sector are consistent with recent studies that provide evidence of PM2.5-driven worker productivity loss from indoor factories (Chang et al. 2016; He, Liu, and Salvo 2019). Moreover, it is important to keep in mind that the true effect can be at any point within the estimated bounds. Therefore, it is difficult to formally compare the effects across different industries — a common feature when employing bounds, because the estimated bounds for each group often overlap substantially.

I also explore whether the PM2.5 effect differs among places of different population density. To do this, I split the whole sample into four sub-samples by population density: a) bottom 25% grids where population density is smaller than 2 people per square kilometer (pp/km^2); b) 25%-50% grids where the density is between $2pp/km^2$ and $11pp/km^2$; c) 50%-75% grids where the density is greater or equal to $11pp/km^2$ but below $34pp/km^2$; and d) top 25% grids where population density is above or equal to $34pp/km^2$.[41]

Figure 5 summarizes the estimated bounds and their 95% confidence intervals on the effect of PM2.5 pollution (defined by the WHO standard) by population density. The true effect is between 0.67 and 7.51 percentage points for the bottom 25% population density group, 0.52 and 4.8 percentage points for the 25%-50% population density group, while between 0.4 and 2.62 percentage points for the 50%-75% population density group and 0.46 and 2.4 for the top 25% population density groups. Although the estimated lower bounds are of similar magnitude for groups with different population density, the estimated upper bounds are smaller for densely populated areas.

The suggestive finding of a smaller effect for densely populated areas is consistent with the fact that industries found with a suggestive larger effect are not generally located in the city centers and densely populated areas. To check the geographic distribution of the PM2.5 effect, I plot in Figure 6 the estimated lower bounds on the percentage effect of PM2.5 pollution by grid. I flag in blue the top 200 Core Based Statistical Areas (CBSAs) in terms of the total population in the Census 2010 in the map. Figure 6 suggests that the estimated lower bounds are in general positive for populous areas, but relatively larger effects are found concentrated on places near the populous areas. This pattern helps explain why Figure 5 shows suggestively smaller effects in the top 50% population density groups.

[41] The population in a $55km \times 55km$ grid is approximately $6,000$ if population density $= 2pp/km^2$, $33,000$ if population density $= 11pp/km^2$, and $103,000$ if population density $= 34pp/km^2$.

7 The Impact on Workers' Compensation

Workplace injuries are costly to the injured workers and their families, the employers and the society. The direct costs of workplace accidents and injuries include workers' compensation payments and medical expenses. Besides the direct costs, employers also face indirect costs such as repairs of damaged equipment and property, training replacement employees, and accident investigation and implementation of corrective measures. Liberty Mutual Insurance estimates that serious, nonfatal workplace injuries cost U.S. businesses more than one billion dollars a week for medical and lost-wage expenses in 2018 (Workplace Safety Index 2021). Workers' compensation benefits paid, along with the productivity loss and medical expenses incurred because of work-related deaths and injuries cost the American society more than 234 billion dollars in 2018 (National Safety Council; Weiss, Murphy, and Boden 2020). For workers, severe injuries that involve hospitalization and amputation usually result in work loss that ranges from days to weeks and even cause disabilities that would limit workers' ability to work in the future. It also implies a potential long-term consequence of short-term exposure to PM2.5.

To quantify the monetary costs of the effect of air pollution on workplace accidents, I provide a back-of-the-envelope calculation of the increase in workers' compensation costs driven by air pollution based on the estimated nonparametric bounds. Consider an artificial location with 112,291 population (same as the size of the population at the sample average). Denote $\hat{\beta}(t), t = 1,2$ the estimated effect of PM2.5 pollution on the workplace accident rate for the PM2.5 pollution indicator $T1$ or $T2$. I use the estimated bounds of the PM2.5 effect employing the wind direction IV in Table 5. I conservatively assume each accident causes one injury. Let c be the estimated workers compensation costs per worker. The National Safety Council estimates the average cost of workers' compensation claims was \$42,008 in 2018-2019.[42] Thus, I set c at \$42,008. The air-pollution-related change of the annual workers' compensation costs is estimated by the following formula:

$$C_{\Delta(t)} = \hat{\beta}(t) \times 112,291 \times 2770 \times \sigma \times c \tag{11}$$

[42]Source: NSC Injury Facts https://injuryfacts.nsc.org/work/costs/workers-compensation-costs/, accessed Feb 5, 2022.

where $\Delta(t)$ denotes the number of monthly PM2.5 polluted days required for the artificial location with average number of PM2.5-polluted days to be treated, i.e., passing the 75th percentile cutoff. For the PM2.5 pollution measure $T1$, $\Delta(1) = 2$, while for the PM2.5 pollution measure $T2$, $\Delta(2) = 1$. σ denotes the pre-specified share of locations with average total population and PM2.5-polluted days that experience the change of PM2.5 pollution in this back-of-the-envelope calculation. I set σ at 1% and thus the back-of-the-envelope calculation is based on changing the annual number of PM2.5-polluted days for 1% of the total places in the sample, i.e., about 28 grids.

I summarize the back-of-the-envelope calculation in Table 6. Columns (3) and (4) show that for average locations with total population at the sample average and the average number of PM2.5-polluted days (according to the WHO standard), having an additional PM2.5-polluted day a year is estimated to increase annual costs of workers' compensation by at least 0.7 billion dollars, equivalent to about 1.03% of the total workers' compensation paid in 2018.[43] The estimated impact is at most 6.1 billion dollars, which is about 9.6% of the total workers' compensation paid in 2018. Similarly, as shown in Columns (1) and (2), for average locations, the estimated change in costs of workers' compensation related to an increase of two additional air pollution days (based on the EPA's standard) is at least 0.9 billion and at most 5.2 billion dollars, about 1.43% and 8.23% of the total workers compensation paid in 2018, respectively. This is a conservative estimation since the cost of workers' compensation are much greater for amputation and other severe injuries. For example, the cost is about $\$114,000$ per claim for amputation, around three times as much as the average cost.

A more straightforward way to see how air pollution-related workplace injuries affect workers' compensation costs is to consider a case study. Let's take Fresno, CA as an example — a place among the top 4 Core Based Statistical Areas in terms of the average monthly number of PM2.5-polluted days (based on the EPA standard). On average, the monthly number of PM2.5-polluted days (based on the EPA standard) is about ten days in Fresno. The bounded effect of PM2.5 pollution on the workplace accident rate is estimated between 2.14 and 14.05 percentage points for

[43]Weiss, Murphy, and Boden (2020) estimates the aggregate workers' compensation paid in 2018 is about 62.9 billion dollars.

Fresno.[44] The 75th percentile of the sample distribution of the monthly PM2.5-polluted days (based on the EPA standard) is about five days (see Figure A.2). That is to say, by reducing the number of PM2.5-polluted days by five days, the workplace accident rate in Fresno would be decreased by at most 14.05 and at least 2.14 percentage points.[45] According to the 2010 Census, Fresno has a population of about 934,450. Therefore, *ceteris paribus*, a five-day reduction in the number of PM2.5-polluted days for Fresno is estimated to save its workers' compensation costs by at least 0.84 billion and up to 5.52 billion dollars. The analysis in this section implies a sizable monetary cost of PM2.5-driven work-related accidents and potential benefits of environmental policies to improve air quality.

8 Conclusion

This paper investigates whether air pollution impairs workplace safety using novel data on workplace severe injuries in the United States. I focus on fine particulate matter (known as PM2.5), a major air pollutant found to have adverse impacts on human cognitive functions and thus potentially affecting workplace accidents via biological channels. Credibly pinning down the causal effect of air pollution on workplace accidents is challenging because air pollution is endogenous. The conventional solution that exploits exogenous variations driven by instrumental variables to point identify the causal effect, among other assumptions, requires a valid instrument that is uncorrelated with workplace accidents except through its effects on air pollution. I provide statistical evidence that commonly-used instruments for air pollution — wind direction and rainfall — violate the point identification assumptions for the outcome of workplace accidents. In the absence of valid instruments, I adopt a partial identification strategy to estimate bounds on the average treatment effect of air pollution, accounting for the invalidity of the instruments. In particular, the partial identification methods replace the exclusion restriction with weaker assumptions assuming a) the same direction of correlation of potential outcomes with the instrument and the treatment (SDC); b) the less endogenous instruments (LEI); c) the bounded support of outcome; and d) the monotonicity

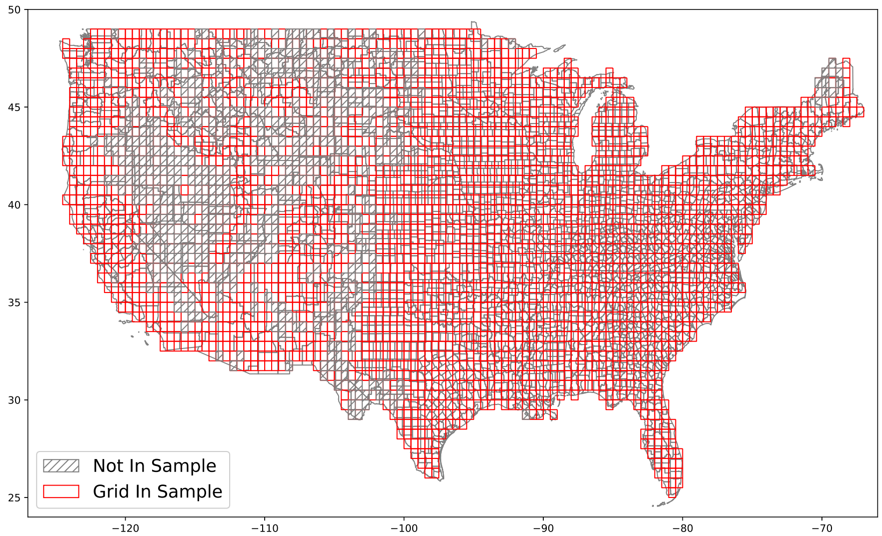
[44]The estimated bounds for Fresno are calculated by taking average over the bounded effects of all grids that lie within or intersect with the boundary of Fresno. The PM2.5 pollution is based on the EPA standard

[45]It is equivalent to reducing the number of PM2.5 polluted days in Fresno to a level in places like State College, PA.

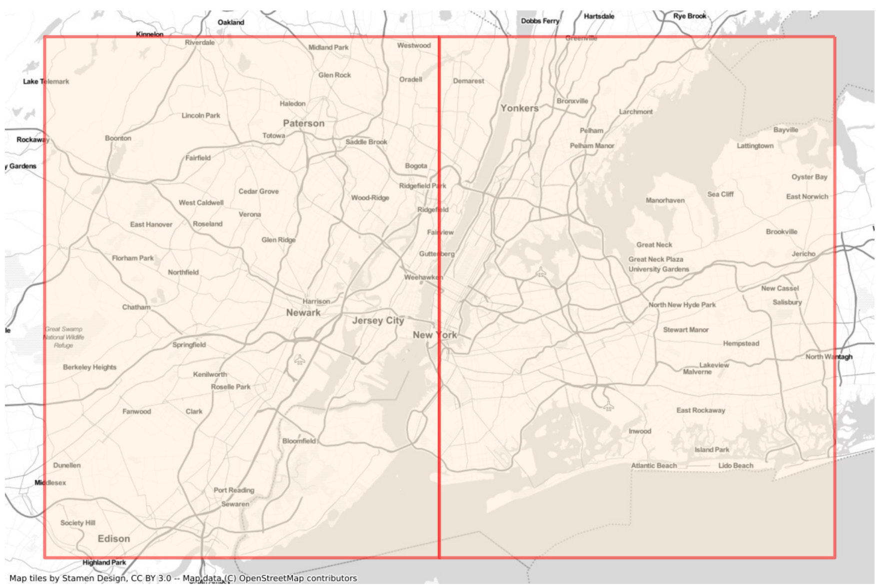
relationship on the treatment response (MTR).

The nonparametric bounds are estimated between 0.5 and 3.9 percentage points employing the wind direction IV and of similar magnitude employing the rainfall IV (between 0.6 and 4.0 percentage points). The estimated bounds (based on the wind direction IV) indicate the effect of air pollution is at least 5% and at most 39% relative to the sample mean accident rate. However, I cannot statistically exclude a zero effect because the 95% confidence region of the estimated bounds includes zero. Given that, I choose to interpret the result as suggestive evidence that air pollution has a nonzero effect on workplace accidents. I further examine the heterogeneous effects. The bounded effects appear to be more prominent for industries that require (semi-)outdoor work, such as agriculture, mining, and construction, than indoor non-manufacturing industries and places near the densely populated areas. Using the estimated bounds to evaluate the impact of air pollution on workers' compensation costs, a back-of-the-envelope calculation indicates that increasing the annual number of PM2.5-polluted days by two additional days is estimated to raise the annual costs of workers' compensation by at least 0.9 billion dollars and up to 5.1 billion, equivalent to about 1 and 8% of total workers' compensation paid in 2018, respectively.

The above findings suggest that improving air quality benefits both employers and employees. It likely reduces labor productivity loss and both the direct and indirect costs of work-related accidents. Moreover, these findings call for caution in the benefit and cost evaluation of environmental policies. Without considering the likely gains from improving work safety, the benefits of environmental policies that intend to improve air quality, whether outdoor or indoor, are likely to be underestimated. The exercise of bounds estimation in this paper also shows the usefulness of the partial identification strategy in the scenario where valid instrumental variables are unavailable. Future work involves tightening the estimated nonparametric bounds, examining dynamic effects, and applying the empirical framework of this study to developing countries.



(a) Contiguous U.S.



(b) Example

Figure 1: Geographical Unit of Analysis.

Notes: Figure 1 presents the geographical unit of analysis. The contiguous US (excluding Alaska and Hawaii) is transformed into a raster with 4426 grids at 0.5 degree scale. Each cell is a square with width of about 55 kilometers, of about 3025 square kilometers (1168 square miles), and equivalent to areas combined Houston and Phoenix. Figure 1b shows an example of two grids covering the New York city. In the analysis, I further remove grids whose population count is zero (and thus missing population density values) and grids whose weather condition measures are missing. The gray shaded areas are grids that are excluded from the analysis sample.

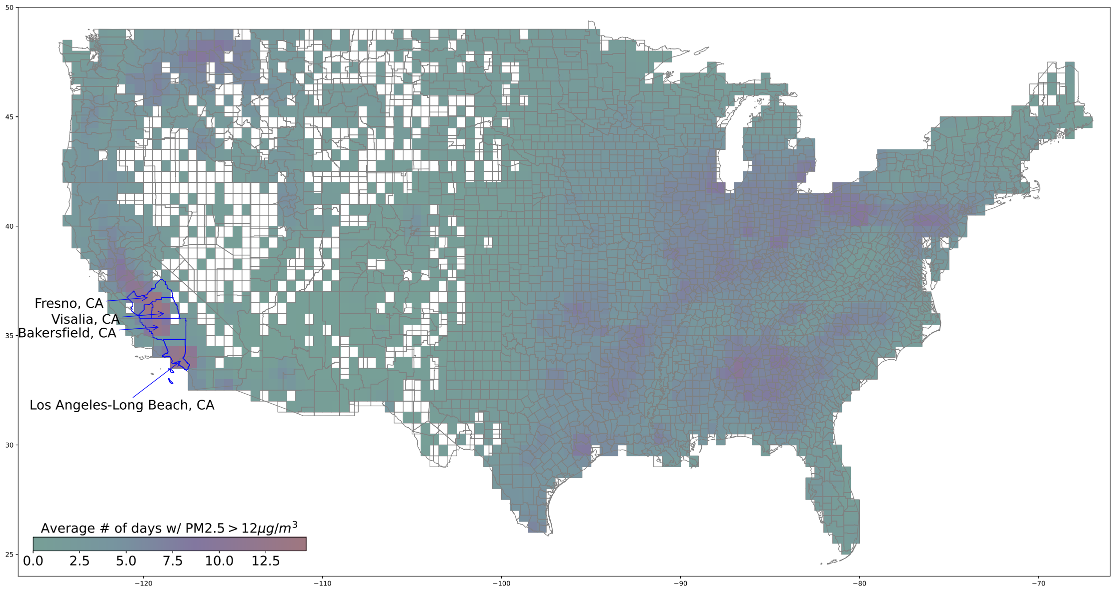
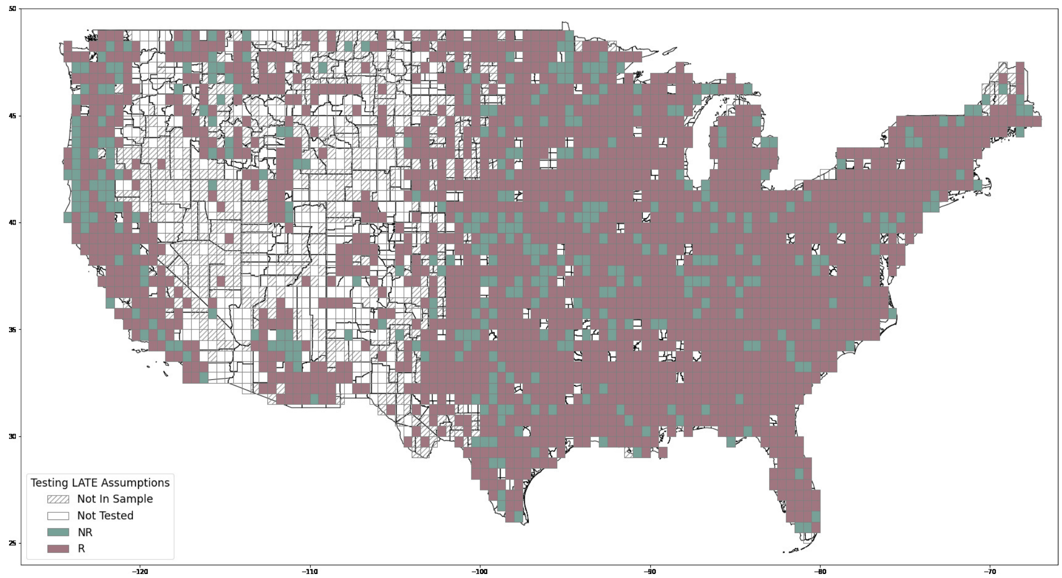
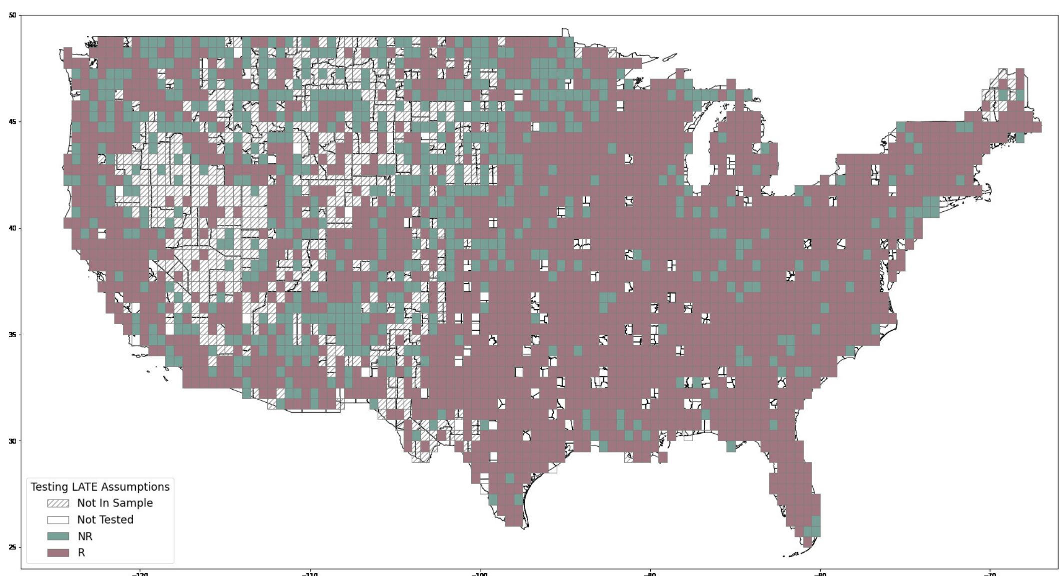


Figure 2: The Average Number of PM2.5 Pollution Days, EPA's Standard.

Notes: Figure 2 plots the geographical distribution of the average monthly number of days with PM2.5 pollution (based on the EPA standard). I mark the top 4 Core Based Statistical Areas (CBSAs) with the highest average monthly number of PM2.5-polluted days in blue. All of them are located in California: Fresno, Visalia, Bakersfield, and Los Angeles-Long Beach.



(a) $Z1$: Rainfall



(b) $Z2$: Wind Direction

Figure 3: Results of the IV Validity Test

Notes: Figure 3 shows the testing results for the IV validity using sub-samples that are split by grid. The marginal significance level is set at 0.003%. The outcome variable is the adjusted accident rate. The treatment variable is the binary indicator of the number of days with PM2.5 pollution according to the WHO standard (PM2.5$>15\mu g/m^3$). The instrumental variable is either the rainfall indicator (Panel a) or the wind direction indicator (Panel b). Green grids imply that the null hypothesis is not rejected, while red grids suggest that the null hypothesis is rejected. White grid indicates no test is conducted, either because the outcome distribution reduces to a masspoint of zero or because the distribution of the IV indicator is degenerated to zero. Grids that are not in the sample are hatched.

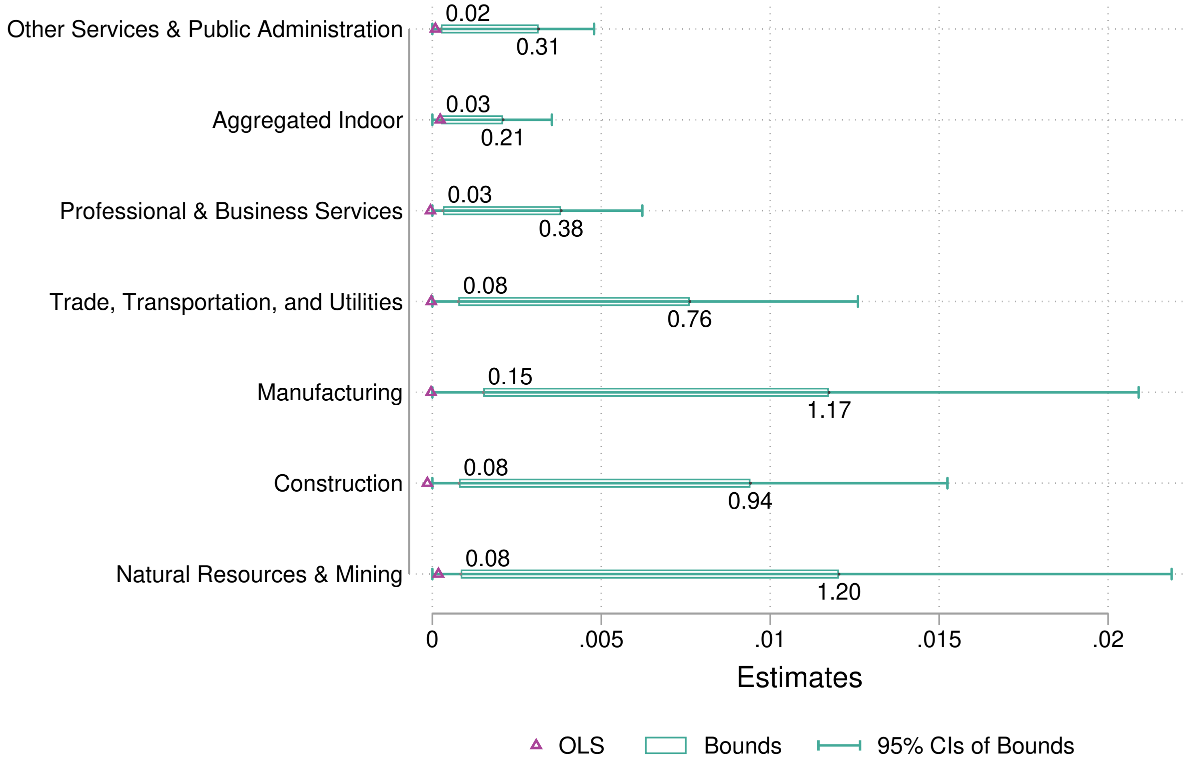


Figure 4: Point Estimates and the 95% Confidence Intervals of Bounds by Industry.

Notes: Figure 4 summarizes the estimated nonparametric bounds on the effect of PM2.5 pollution defined by the WHO standard ($T2$) by industry, employing the wind direction IV. The definition of the treatment variable follows Equation 2. Numbers aside the boxes presents the estimated bounds in percentage points.

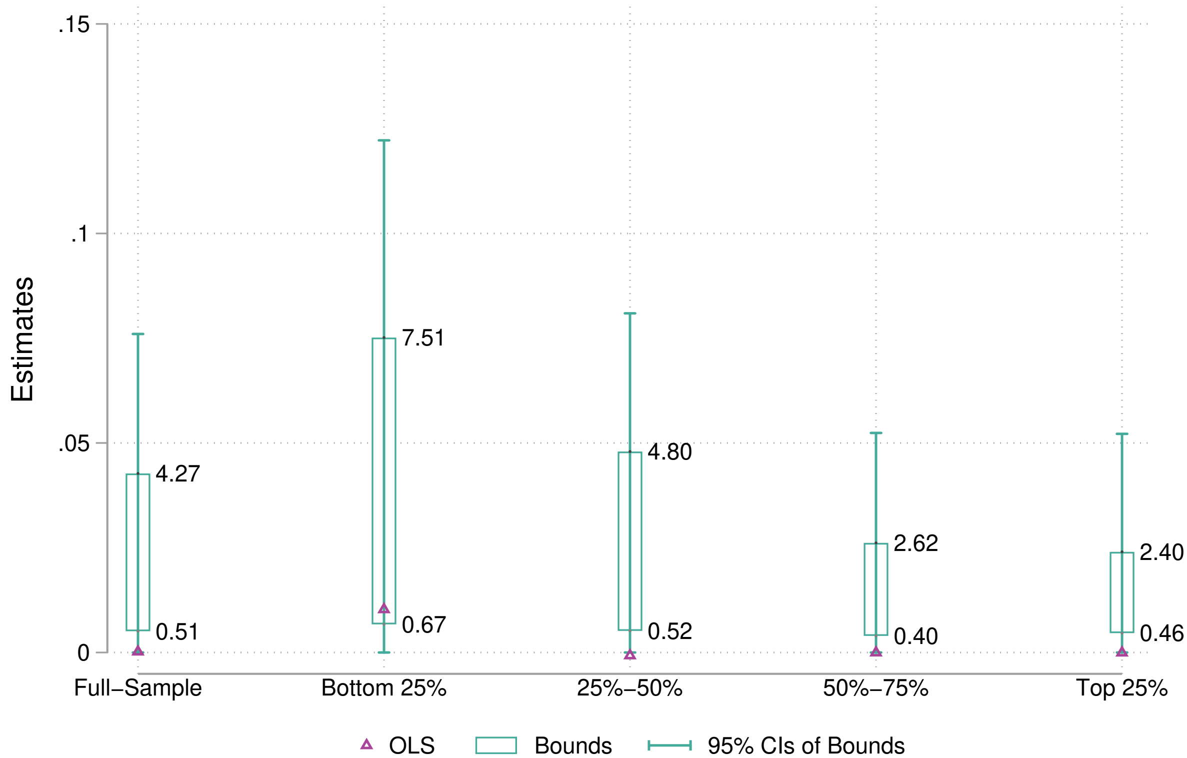


Figure 5: Point Estimates and 95% Confidence Intervals of Bounds by Population Density.

Notes: Figure 5 summarizes the estimated nonparametric bounds on the effect of PM2.5 pollution defined by the WHO standard ($T2$) by population density, employing the wind direction IV. The definition of the treatment variable follows Equation 2. Numbers aside the boxes presents the estimated bounds in percentage points.

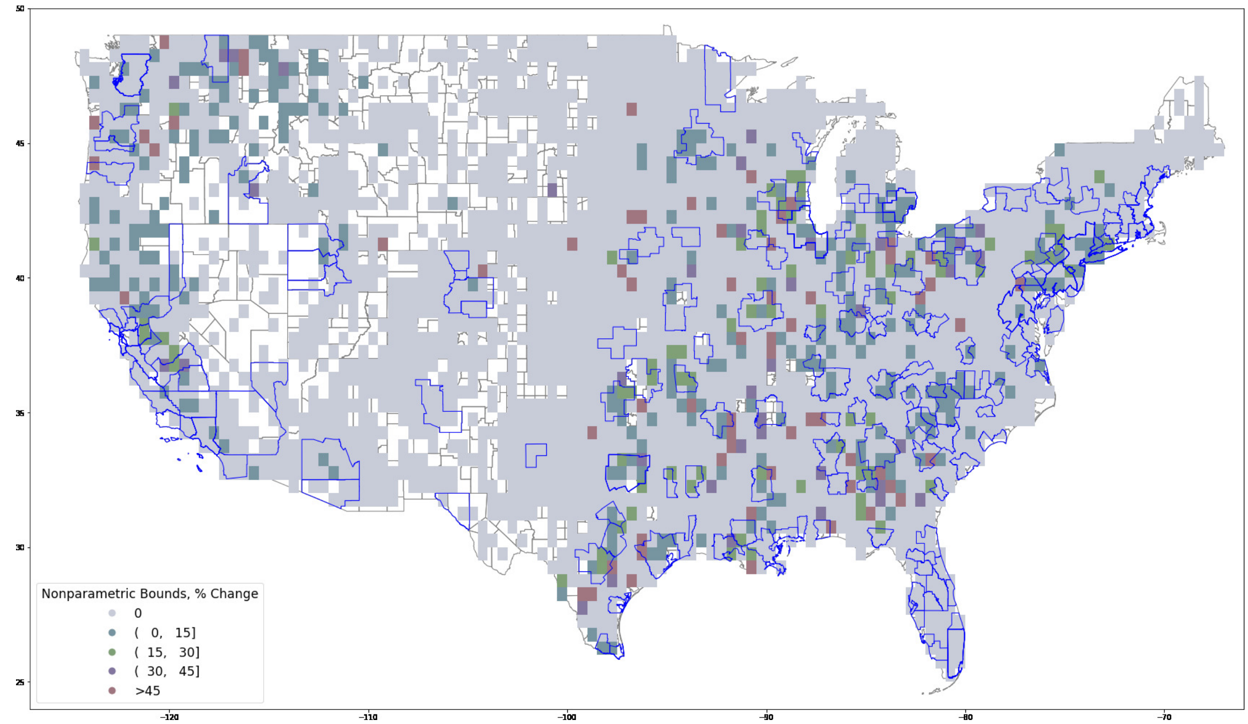


Figure 6: Nonparametric Bounds Under Assumptions SDC + LEI + MTR.

Notes: Figure 6 plots the estimated nonparametric bounds on the effect of PM2.5 pollution defined by the WHO standard ($T2$) by grid under the SDC + LEI + MTR assumptions, employing the wind direction IV. I transfer the estimated bounds to the percentage effect by dividing the point estimate by the sample average accident rate. Grids not in the sample are marked as white. Blue polygons flag the top 200 Core Based Statistical Areas (CBSAs) in terms of the total population in the Census 2010.

Table 1: Summary Statistics

Variable Name	Mean	Std.	Min	P25	P75	Max
Outcome Variable						
All Industries						
Accident Rate (per $10,000$ population)	0.102	2.61	0	0	0	338.26
Treatment Variable						
Monthly Average PM2.5 Level ($\mu g/m^3$)	7.280	2.87	1.31	5.37	8.68	85.59
PM2.5 Indicator						
$T1$: EPA standard 12 $\mu g/m^3$	0.212	0.41	0	0	0	1
$T2$: WHO standard 15 $\mu g/m^3$	0.197	0.40	0	0	0	1
Instrumental Variable						
Z_1: Rainfall	0.204	0.40	0	0	0	1
Z_2: Wind Direction	0.218	0.41	0	0	0	1
Covariate						
Monthly Average Precipitation (mm)	2.350	2.05	0.00	0.81	3.34	26.19
Monthly Average Wind Speed (m/s)	3.500	0.86	0.45	2.92	4.02	9.20
Monthly Average Temperature (C)	13.105	10.40	-18.62	5.41	21.66	36.81
Monthly Maximum Temperature (C)	19.725	8.97	-9.65	13.78	26.90	40.67
Monthly Minimum Temperature (C)	5.283	12.16	-31.42	-3.08	14.85	34.69
Population Count	112291	350462	1	5230	78585	9205318
Population Density (population$/km^2$)	51.886	192.72	0.00	2.36	33.47	6167.98
						$N = 132960$

Notes: Table 1 lists the sample mean, standard deviation, maximum, minimum and 25 percentile (P25) and 75 percentile (P75) of the outcome variable, treatment variable, instrumental variable and covariates. Outcome variable is defined as the monthly average rate per ten thousand population of workplace accidents that involve severe injuries. Treatment variables $T1$ and $T2$ indicate whether the location is exposure to more days of PM2.5 pollution in a month, as defined by Equation 2 with the cutoff set at the 75th percentile of the sample distribution of the number of PM2.5 polluted days in a month. Instrumental variables Z_1 and Z_2 denote the rainfall IV and the wind direction IV, respectively. The definition of Z_1 and Z_2 can be found in Equations 3 and 4. Covariates include monthly average temperature and average wind speed, maximum temperature, and minimum temperature. For the model of wind direction IV, I also control for monthly average precipitation.

Table 2: OLS Estimation

	(1)	(2)
	Outcome: Workplace Accident Rate (per 10,000 ppl)	
T1: EPA standard 12 $\mu g/m^3$	0.0011**	
	(0.0004)	
T2: WHO standard 15 $\mu g/m^3$		0.0003
		(0.0004)
N	132960	132960

Notes: Table 2 presents the OLS estimation results. Robust standard errors are presented in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. All specifications include time-varying weather controls, along with grid and month by year fixed effects. All models are weighted by grid-level population density in 2015. The treatment variable are binary dummies indicating whether the number of days with PM2.5 pollution at certain level in a month group is above a pre-specified threshold. In the main analysis, the threshold is set at the 75th percentile of the sample distribution of the number of days with PM2.5 pollution at certain level in that month group.

Table 3: 2SLS Estimation

	(1)	(2)	(3)	(4)
Panel A. First Stage	T1: EPA standard 12 $\mu g/m^3$		T2: WHO standard 15 $\mu g/m^3$	
IV Z1: Rainfall	-0.1070***		-0.0757***	
	(0.0110)		(0.0107)	
IV Z2: Wind Direction		-0.0825***		-0.0637***
		(0.0098)		(0.0101)
Montiel-Pflueger Effective F Statistics	98.02	53.27	51.04	29.46
Panel B. Second Stage	*Outcome: Workplace Accident Rate (per 10,000 ppl)*			
IV Z1: Rainfall	-0.0009		-0.0012	
	(0.0041)		(0.0059)	
IV Z2: Wind Direction		0.00003		0.00004
		(0.0051)		(0.0066)
N	132960	132960	132960	132960

Notes: Table 3 presents the first-stage and second-stage 2SLS estimation results. Heteroskedasticity and autocorrelation robust standard errors are presented in parentheses. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$. I report the Montiel-Pflueger effective first-stage F statistics. The 5% critical values for testing the null hypothesis that the 2SLS bias exceeds 10% (5%) of the OLS bias is 23.11 (37.42). All specifications include time-varying weather controls, along with grid and month by year fixed effects. All models are weighted by grid-level population density in 2015. The treatment variable are binary dummies indicating whether the number of days with PM2.5 pollution at certain level in a month group is above a pre-specified threshold. In the main analysis, the threshold is set at the 75th percentile of the sample distribution of the number of days with PM2.5 pollution at certain level in that month group.

Table 4: Parametric Bounds Estimation

	(1) Lower Bound	(2) Upper Bound	(3) Lower Bound	(4) Upper Bound
	Outcome: Workplace Accident Rate (per 10,000 ppl)			
	T1: EPA standard 12 $\mu g/m^3$		T2: WHO standard 15 $\mu g/m^3$	
Panel A. IV Z1: Rainfall				
Point Estimates	[-0.0009,	0.0009]	[-0.0012,	0.0002]
95% CI of Identified Set	[-0.0089,	0.0020]	[-0.0126,	0.0014]
90% CI of Identified Set	[-0.0076,	0.0018]	[-0.0108,	0.0012]
Panel B. IV Z2: Wind Direction				
Point Estimates	[0.00003,	0.0009]	[0.00004,	0.0002]
95% CI of Identified Set	[-0.0100,	0.0020]	[-0.0130,	0.0013]
90% CI of Identified Set	[-0.0084,	0.0012]	[-0.0109,	0.0012]
N	132960	132960	132960	132960

Notes: Table 4 presents the parametric bounds estimation results. Parametric bounds are estimated by the method of Nevo and Rosen (2012), employing two invalid IVs. The treatment variable are binary dummies indicating whether the number of days with PM2.5 pollution at certain level in a month group is above a pre-specified threshold. In the main analysis, the threshold is set at the 75th percentile of the sample distribution of the number of days with PM2.5 pollution at certain level in that month group. All specifications include time-varying weather controls, along with grid and month by year fixed effects. All models are weighted by grid-level population density in 2015.

Table 5: Nonparametric Bounds Estimation

	(1) *Lower Bound*	(2) *Upper Bound*	(3) *Lower Bound*	(4) *Upper Bound*
	Outcome: Workplace Accident Rate (per 10,000 ppl)			
	T1: EPA standard 12 $\mu g/m^3$		T2: WHO standard 15 $\mu g/m^3$	
Panel A. BOS + SDC + LEI				
IV Z1: Rainfall				
Point Estimates	[-0.1432,	0.1666]	[-0.1361,	0.1815]
95% CI of Identified Set	[-0.4197,	0.4050]	[-0.3999,	0.3985]
IV Z2: Wind Direction				
Point Estimates	[-0.1284,	0.1579]	[-0.1356,	0.1659]
95% CI of Identified Set	[-0.3809,	0.3731]	[-0.4048,	0.3801]
Panel B. BOS + SDC + LEI + MTR				
IV Z1: Rainfall				
Point Estimates	[0.0054	0.0393]	[0.0056,	0.0402]
95% CI of Identified Set	[0.0000,	0.0794]	[0.0000,	0.0792]
IV Z2: Wind Direction				
Point Estimates	[0.0068,	0.0392]	[0.0049,	0.0459]
95% CI of Identified Set	[0.0000,	0.0821]	[0.0000,	0.0816]
Mean Accident Rate		0.10		

Notes: Table 5 presents the nonparametric bounds estimation results. Nonparametric bounds are estimated by the method of Ban and Kédagni (2022), employing two different invalid IVs. Panel A shows bounds under the BOS, SDC and LEI assumptions. Panel B reports bounds under the BOS, SDC, LEI, and MTR assumptions. The treatment variable is a binary dummy indicating whether the number of days with PM2.5 pollution at certain level in a month group is above a pre-specified threshold. In the main analysis, the threshold is set at the 75th percentile of the sample distribution of the number of days with PM2.5 pollution at certain level in that month group.

Table 6: Evaluating the Impact on Workers' Compensation Costs

	(1) *Lower Bound*	(2) *Upper Bound*	(3) *Lower Bound*	(4) *Upper Bound*
	$T1$: EPA standard 12 $\mu g/m^3$		$T2$: WHO standard 15 $\mu g/m^3$	
$\hat{\beta}$	[0.007,	0.039]	[0.005,	0.046]
Δ	2		1	
C_{Δ}	[\$898,139,711,	\$5,177,511,272]	[\$647,188,909 ,	\$6,062,443,046]
Share of Total WC Paid in 2018	[1.43%,	8.23%]	[1.03%,	9.64%]
Aggregate WC Paid in 2018	\$62,900,000,000			

Notes: C_{Δ} implies the changes in the annual workers' compensation costs associated with an increase of air-polluted days by Δ days for average areas with total population at the sample mean (112,291) and average number of air-polluted days. The estimated workers' compensation paid per worker is set at \$42008. Δ denotes the pre-specified percentage change in air-polluted days a year. $\hat{\beta}$ denotes the estimated nonparametric bounds on the effect of PM2.5 pollution on the workplace accident rate employing the wind direction IIV in Table 5.

References

Anderson, Michael L. 2020. As the wind blows: the effects of long-term exposure to air pollution on mortality. *Journal of the European Economic Association* 18:1886–1927.

Andrews, Isaiah, James H. Stock, and Liyang Sun. 2019. Weak instruments in instrumental variables regression: theory and practice. *Annual Review of Economics* 11.

Auchincloss, Amy H., Ana V. Diez Roux, J. Timothy Dvonch, Patrick L. Brown, R. Graham Barr, Martha L. Daviglus, David C. Goff Jr, Joel D. Kaufman, and Marie S. O'Neill. 2008. Associations between recent exposure to ambient fine particulate matter and blood pressure in the multi-ethnic study of atherosclerosis (MESA). *Environmental health perspectives* 116:486–491.

Balietti, Anca, Souvik Datta, and Stefanija Veljanoska. 2022. Air pollution and child development in india. *Journal of Environmental Economics and Management* 113 (May): 102624.

Ban, Kyunghoon, and Désiré Kédagni. 2022. Nonparametric bounds on treatment effects with imperfect instruments. *The Econometrics Journal* 25:477–493.

Bedi, Arjun S., Marcos Y. Nakaguma, Brandon J. Restrepo, and Matthias Rieger. 2021. Particle pollution and cognition: evidence from sensitive cognitive tests in brazil. *Journal of the Association of Environmental and Resource Economists* 8:443–474.

Bharadwaj, Prashant, Matthew Gibson, Joshua Graff Zivin, and Christopher Neilson. 2017. Gray matters: fetal pollution exposure and human capital formation. *Journal of the Association of Environmental and Resource Economists* 4:505–542.

Center for International Earth Science Information Network - CIESIN - Columbia University. 2018. *Gridded population of the world, version 4 (GPWv4): population density, revision 11.*

Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2016. Particulate pollution and the productivity of pear packers. *American Economic Journal: Economic Policy* 8:141–69.

———. 2019. The effect of pollution on worker productivity: evidence from call center workers in china. *American Economic Journal: Applied Economics* 11:151–72.

Chernozhukov, Victor, Wooyoung Kim, Sokbae Lee, and Adam M. Rosen. 2015. Implementing intersection bounds in stata. *The Stata Journal* 15:21–44.

Chernozhukov, Victor, Sokbae Lee, and Adam M. Rosen. 2013. Intersection bounds: estimation and inference. *Econometrica* 81:667–737.

Connolly, Marie. 2008. Here comes the rain again: weather and the intertemporal substitution of leisure. *Journal of Labor Economics* 26 (January): 73–100.

Currie, Janet, Eric A. Hanushek, E. Megan Kahn, Matthew Neidell, and Steven G. Rivkin. 2009. Does pollution increase school absences? *The Review of Economics and Statistics* 91:682–694.

Depalo, Domenico, and Alessandro Palma. 2020. *Work accidents and air pollution.*

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. 2019. The mortality and medical costs of air pollution: evidence from changes in wind direction. *American Economic Review* 109:4178–4219.

EPA, US. 2009. *Integrated science assessment for particulate matter.* US.

———. 2012. *Revised air quality standards for particle pollution and updates to the air quality index (AQI).*

———. 2019. *Integrated science assessment for particulate matter.* US.

Glickman, Todd S., and Walter Zenk. 2000. *Glossary of meteorology.* AMS (American Meteorological Society).

Graff Zivin, Joshua, and Matthew Neidell. 2012. The impact of pollution on worker productivity. *American Economic Review* 102:3652–73.

———. 2014. Temperature and the allocation of time: implications for climate change. *Journal of Labor Economics* 32:1–26.

Guo, Ling-Chuan, Yonghui Zhang, Hualiang Lin, Weilin Zeng, Tao Liu, Jianpeng Xiao, Shannon Rutherford, Jing You, and Wenjun Ma. 2016. The washout effects of rainfall on atmospheric particulate pollution in two chinese cities. *Environmental Pollution* 215:195–202.

Guxens, Mònica, Malgorzata J. Lubczyńska, Ryan L. Muetzel, Albert Dalmau-Bueno, Vincent WV Jaddoe, Gerard Hoek, Aad van der Lugt, Frank C. Verhulst, Tonya White, and Bert Brunekreef. 2018. Air pollution exposure during fetal life, brain morphology, and cognitive function in school-age children. *Biological psychiatry* 84:295–303.

Hanna, Rema, and Paulina Oliva. 2015. The effect of pollution on labor supply: evidence from a natural experiment in mexico city. *Journal of Public Economics* 122:68–79.

Hansen, Anett C., and Harald K. Selte. 2000. Air pollution and sick-leaves. *Environmental and Resource Economics* 16:31–50.

Harada, Caroline N., Marissa C. Natelson Love, and Kristen L. Triebel. 2013. Normal cognitive aging. *Clinics in geriatric medicine* 29:737–752.

Hausman, Jerry A., Bart D. Ostro, and David A. Wise. 1984. *Air pollution and lost work.* National Bureau of Economic Research Cambridge, Mass., USA.

He, Jiaxiu, Haoming Liu, and Alberto Salvo. 2019. Severe air pollution and labor productivity: evidence from industrial towns in china. *American Economic Journal: Applied Economics* 11:173–201.

Heyes, Anthony, and Mingying Zhu. 2019. Air pollution as a cause of sleeplessness: social media evidence from a panel of chinese cities. *Journal of Environmental Economics and Management* 98:102247.

Holm, Sture. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics,* 65–70.

Imbens, Guido W., and Joshua D. Angrist. 1994. Identification and estimation of local average treatment effects. *Econometrica: Journal of the Econometric Society,* 467–475.

Insurance, Liberty Mutual. 2021. *Liberty mutual workplace safety index.*

La Nauze, Andrea, and Edson R. Severnini. 2021. *Air pollution and adult cognition: evidence from brain training.* National Bureau of Economic Research.

Manski, Charles F. 1990. Nonparametric bounds on treatment effects. *The American Economic Review* 80:319–323.

———. 1997. Monotone treatment response. *Econometrica: Journal of the Econometric Society,* 1311–1334.

Manski, Charles F., and John V. Pepper. 2000. Monotone instrumental variables: with an application to the returns to schooling. *Econometrica* 68:997–1010.

Manski, Charles F., and C. Sims. 1994. The selection problem. In *Advances in econometrics, sixth world congress,* 1:143–70.

Mesinger, Fedor, Geoff DiMego, Eugenia Kalnay, Kenneth Mitchell, Perry C. Shafran, Wesley Ebisuzaki, Dušan Jović, Jack Woollen, Eric Rogers, and Ernesto H. Berbery. 2006. North american regional reanalysis. *Bulletin of the American Meteorological Society* 87:343–360.

Montiel Olea, José Luis, and Carolin Pflueger. 2013. A robust test for weak instruments. *Journal of Business & Economic Statistics* 31 (July): 358–369.

Mourifié, Ismael, and Yuanyuan Wan. 2017. Testing local average treatment effect assumptions. *Review of Economics and Statistics* 99:305–313.

Nevo, Aviv, and Adam M. Rosen. 2012. Identification with imperfect instruments. *Review of Economics and Statistics* 94:659–671.

Ostro, Bart D. 1983. The effects of air pollution on work loss and morbidity. *Journal of Environmental Economics and Management* 10:371–382.

Park, R. Jisung, Nora Pankratz, and A. Patrick Behrer. 2021. *Temperature, workplace safety, and labor market inequality.* IZA Discussion Paper.

Peters, Annette, Emerson Liu, Richard L. Verrier, Joel Schwartz, Diane R. Gold, Murray Mittleman, Jeff Baliff, J. Annie Oh, George Allen, and Kevin Monahan. 2000. Air pollution and incidence of cardiac arrhythmia. *Epidemiology* 11:11–17.

Pönkä, Antti. 1990. Absenteeism and respiratory disease among children and adults in helsinki in relation to low-level air pollution and temperature. *Environmental Research* 52:34–46.

Pope, C. Arden. 2000. Epidemiological basis for particulate air pollution health standards. *Aerosol Science & Technology* 32:4–14.

Prado Bert, Paula de, Elisabet Mae Henderson Mercader, Jesus Pujol, Jordi Sunyer, and Marion Mortamais. 2018. The effects of air pollution on the brain: a review of studies interfacing environmental epidemiology and neuroimaging. *Current environmental health reports* 5:351–364.

Ransom, Michael R., and C. Arden Pope III. 1992. Elementary school absences and PM10 pollution in utah valley. *Environmental research* 58:204–219.

Rezaei, Mahdi, and Reinhard Klette. 2014. Look at the driver, look at the road: no distraction! no accident! In *Proceedings of the IEEE conference on computer vision and pattern recognition,* 129–136.

Rubin, Donald B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology* 66:688.

Shehab, M. A., and F. D. Pope. 2019. Effects of short-term exposure to particulate matter air pollution on cognitive performance. *Scientific reports* 9:1–10.

Swaen, G. M. H., LGPM Van Amelsvoort, U. Bültmann, and I. J. Kant. 2003. Fatigue as a risk factor for being injured in an occupational accident: results from the maastricht cohort study. *Occupational and environmental medicine* 60 (suppl 1): i88–i92.

Tai, Amos PK, Loretta J. Mickley, and Daniel J. Jacob. 2010. Correlations between fine particulate matter (PM2. 5) and meteorological variables in the united states: implications for the sensitivity of PM2. 5 to climate change. *Atmospheric environment* 44:3976–3984.

Talbot, Rachel, Helen Fagerlind, and Andrew Morris. 2013. Exploring inattention and distraction in the SafetyNet accident causation database. *Accident Analysis & Prevention* 60:445–455.

Thornton, M.M., R. Shrestha, Y. Wei, P.E. Thornton, S. Kao, and B.E. Wilson. 2020. *Daymet: daily surface weather data on a 1-km grid for north america, version 4.*

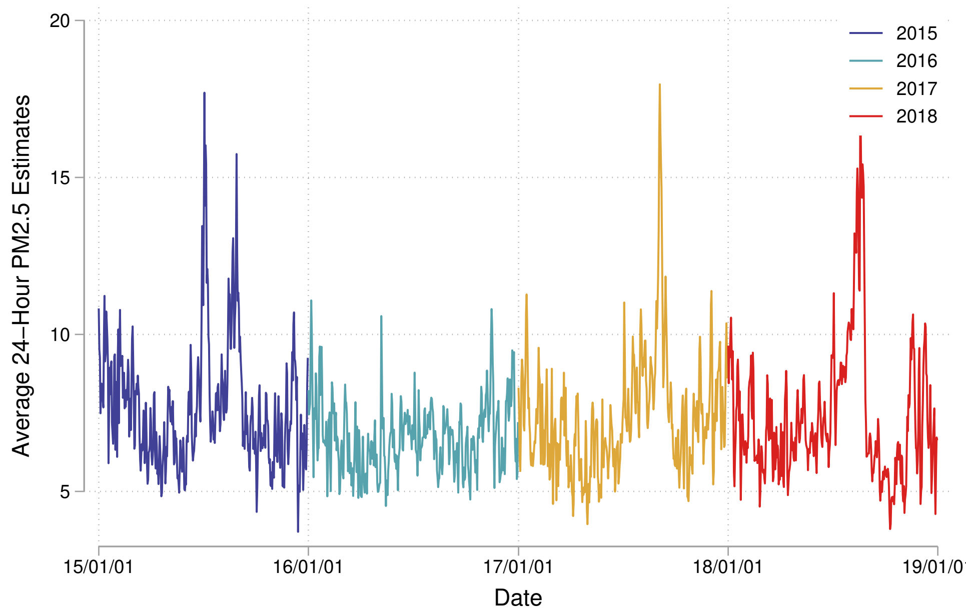
Tucker, Philip. 2003. The impact of rest breaks upon accident risk, fatigue and performance: a review. *Work & Stress* 17:123–137.

Weiss, Elaine, Griffin Murphy, and Leslie I. Boden. 2020. Workers' compensation: benefits, costs, and coverage (2018 data). *Washington, DC: National Academy of Social Insurance.*

Weuve, Jennifer, Robin C. Puett, Joel Schwartz, Jeff D. Yanosky, Francine Laden, and Francine Grodstein. 2012. Exposure to particulate air pollution and cognitive decline in older women. *Archives of internal medicine* 172:219–227.

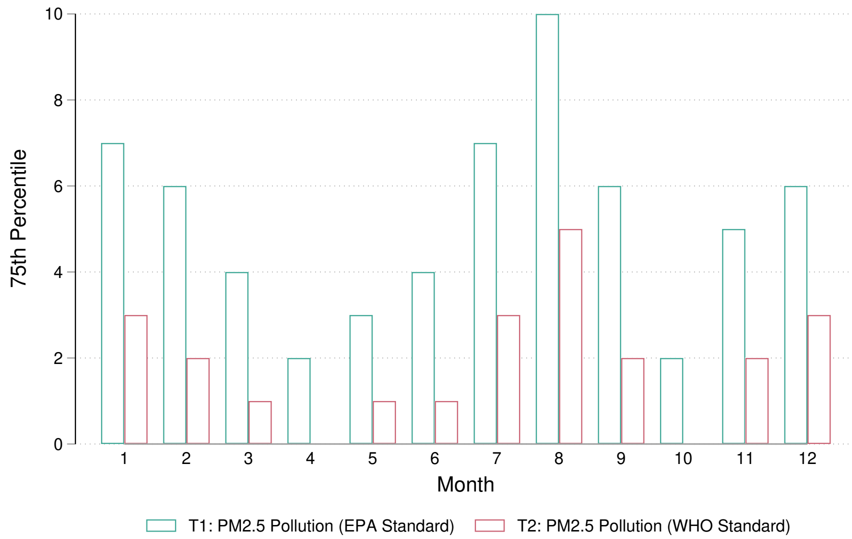
WHO. 2021. *WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide.* World Health Organization.

A Figures

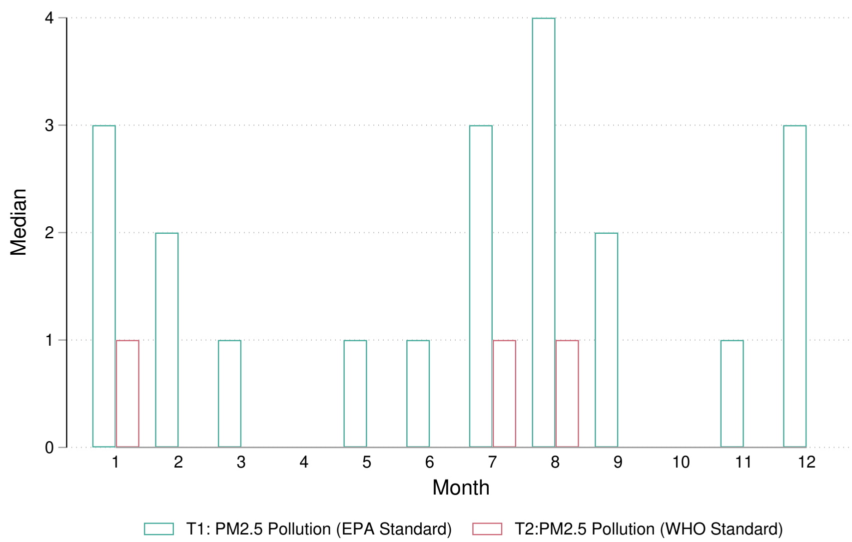


A.1: The Time Trend and Seasonality of PM2.5 in the U.S.

Notes: Figure A.1 shows the evolution of the 24-Hour average PM2.5 over all grids in the sample by date during 2015-2018. The level of average ambient PM2.5 is generally high in the summer and relatively lower in the spring and fall. We observe sizable annual and seasonal variations during the sample periods.



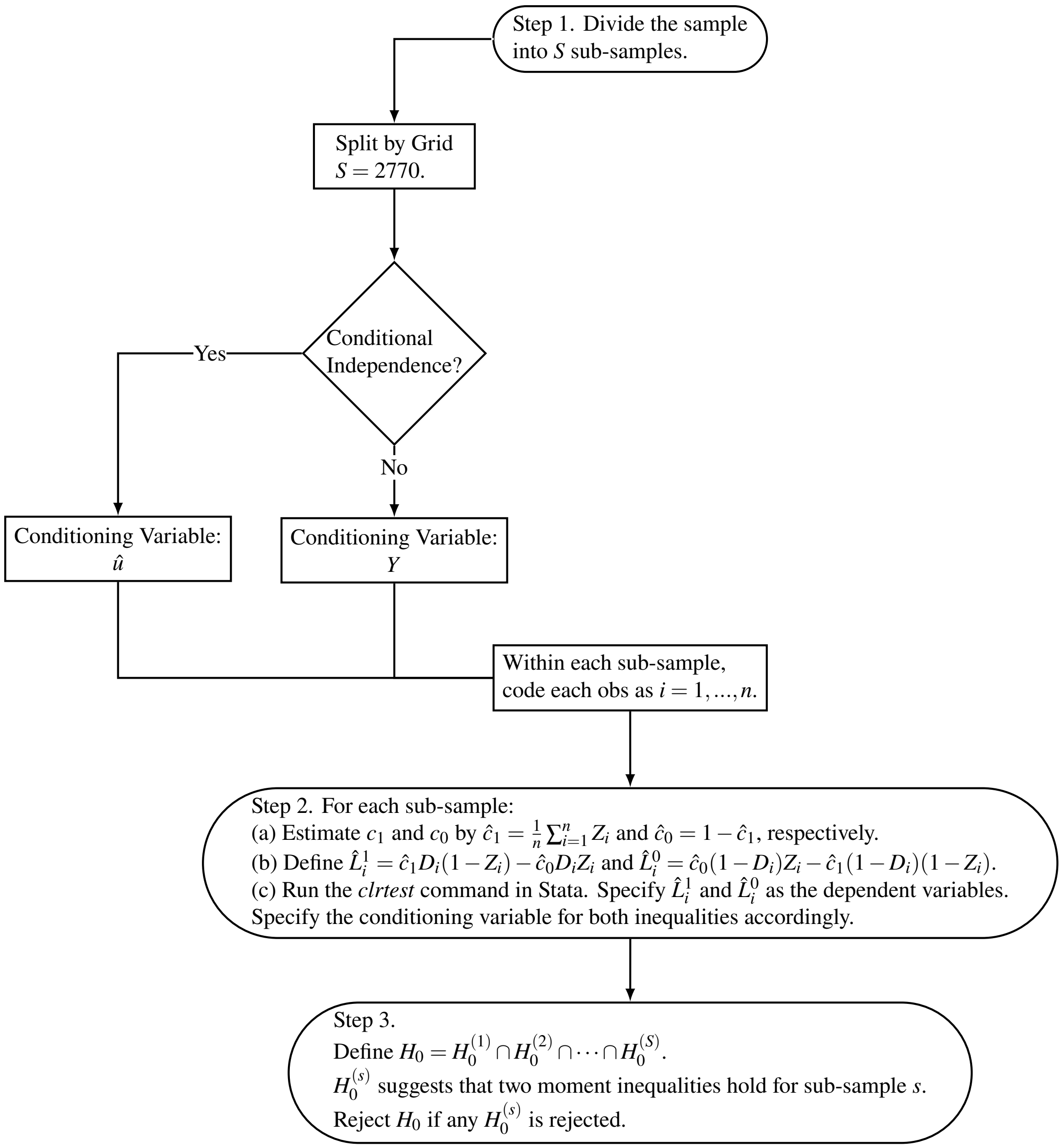
(a) 75th Percentile



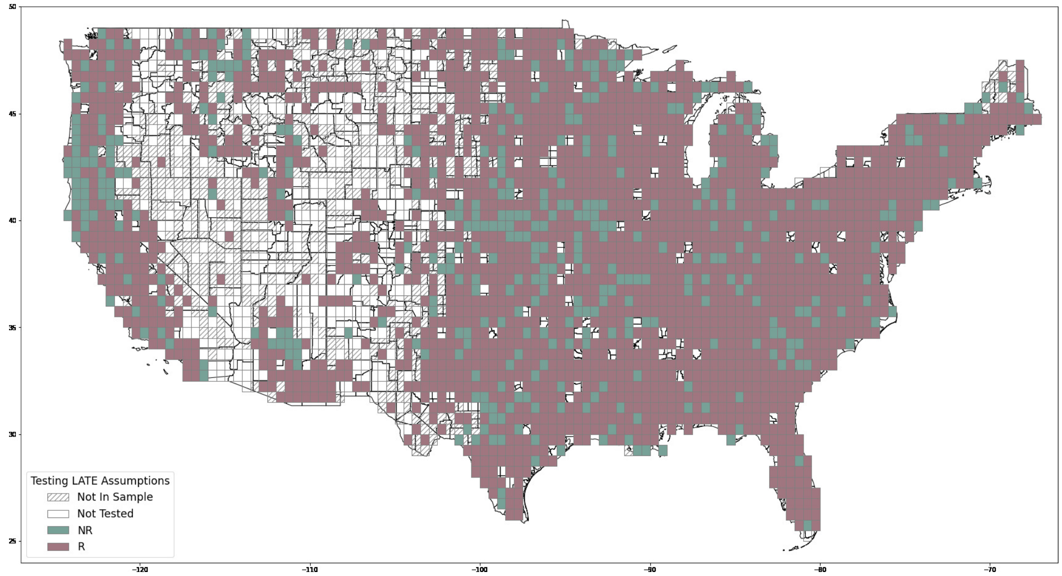
(b) Median

A.2: Treatment Variable Cutoffs

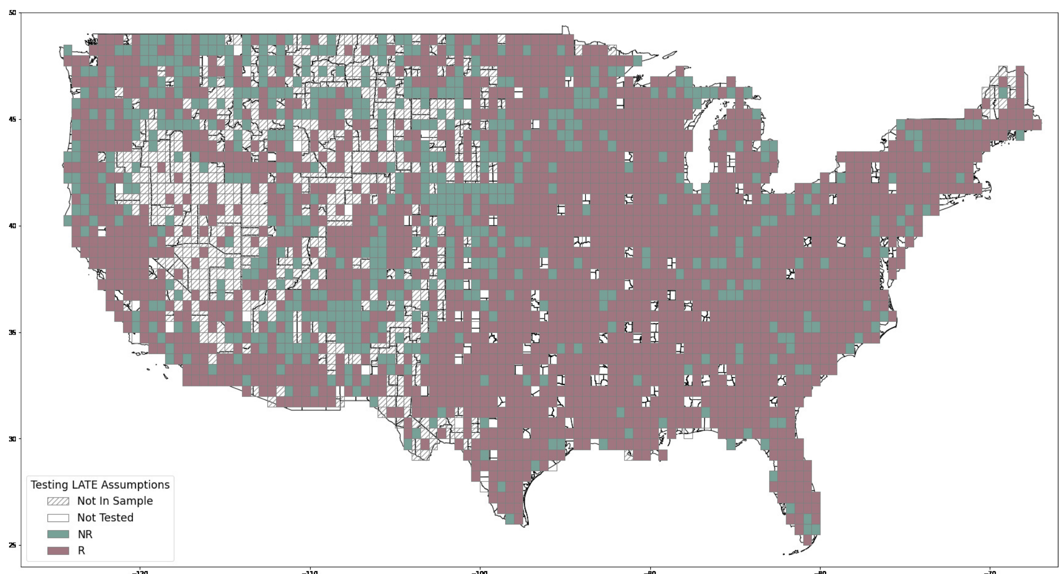
Notes: Figure A.2 plots the cutoffs of by month. Panel (a) shows the 75th percentile of the sample distribution of the number of PM2.5-polluted days, while Panel (b) plots the median of the sample distribution. For the PM2.5 pollution defined by the EPA standard ($T1$), the cutoffs vary from one to ten. The sample mean of the cutoffs is 5.16 for the 75th percentile threshold and 1.75 for the median threshold. For the PM2.5 pollution defined by the WHO standard ($T2$), the cutoffs vary from zero to four, with the mean of the cutoffs is 1.91 for the 75th percentile threshold and 0.25 for the median threshold.



A.3: IV Validity Test Flow Chart.



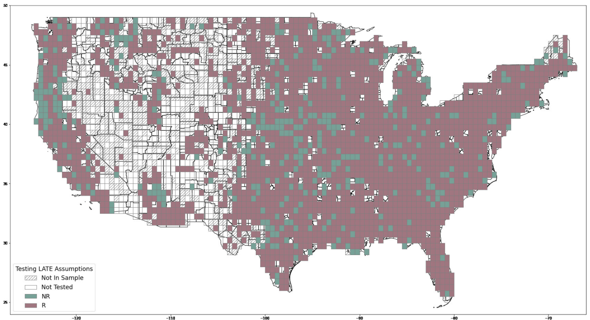
(a) $Z1$: Rainfall



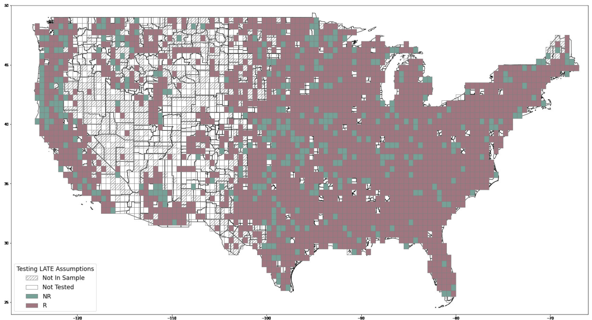
(b) $Z2$: Wind Direction

A.4: Results of the IV Validity Test, Alternative Treatment Variable

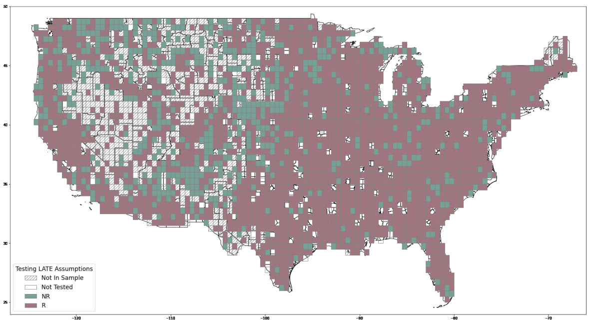
Notes: Figure A.4 shows the testing results for IV validity using sub-samples that are split by grid. The marginal significance level is set at 0.003%. The outcome variable is the adjusted accident rate. The treatment variable is the binary indicator of the number of days with PM2.5 pollution according to the EPA standard (PM2.5$>12\mu g/m^3$). The instrumental variable is either the rainfall indicator (Panel a) or the wind direction indicator (Panel b). The green grid implies that the null hypothesis is not rejected, while the red grid suggests that the null hypothesis is rejected. White grid indicates no test is conducted, either because the outcome distribution reduces to a masspoint of zero or because the distribution of the IV indicator is degenerated to zero. Grids that are not in the sample are hatched.



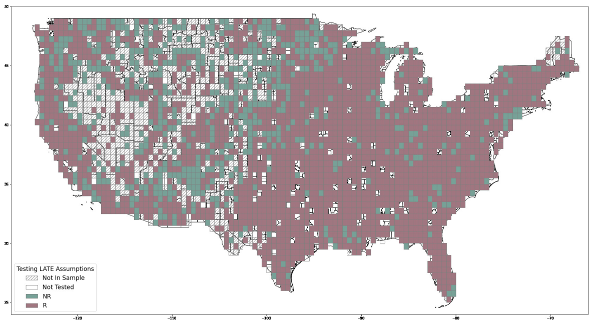
(a) $Z1$: Rainfall; EPA's standard



(b) $Z1$: Rainfall; WHO's standard



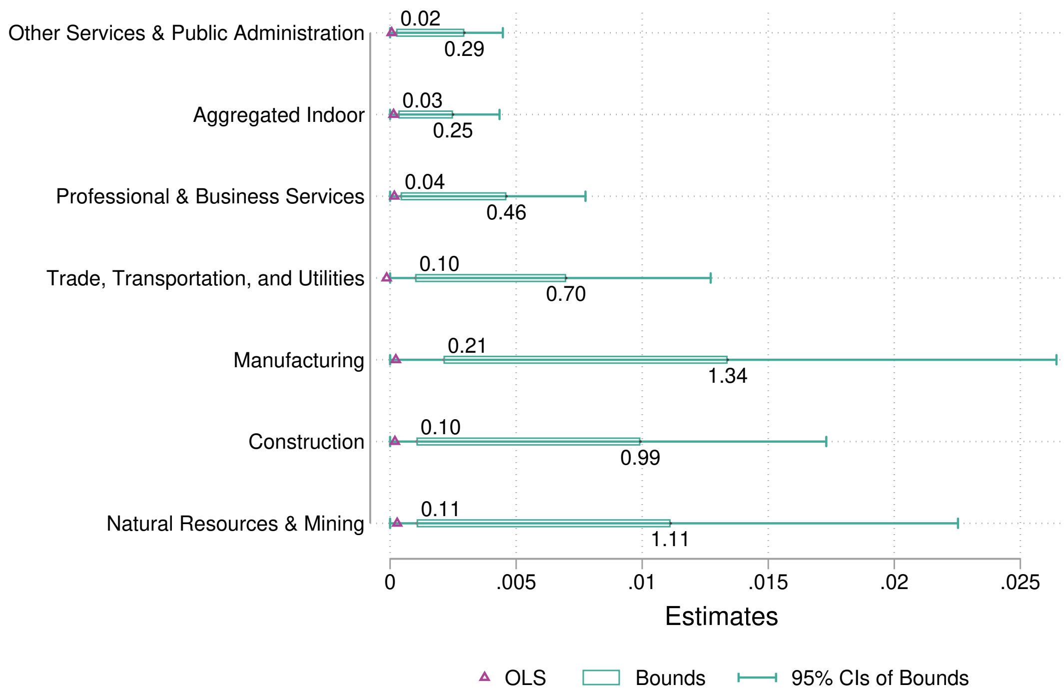
(c) $Z2$: Wind Direction; EPA's standard



(d) $Z2$: Wind Direction; WHO's standard

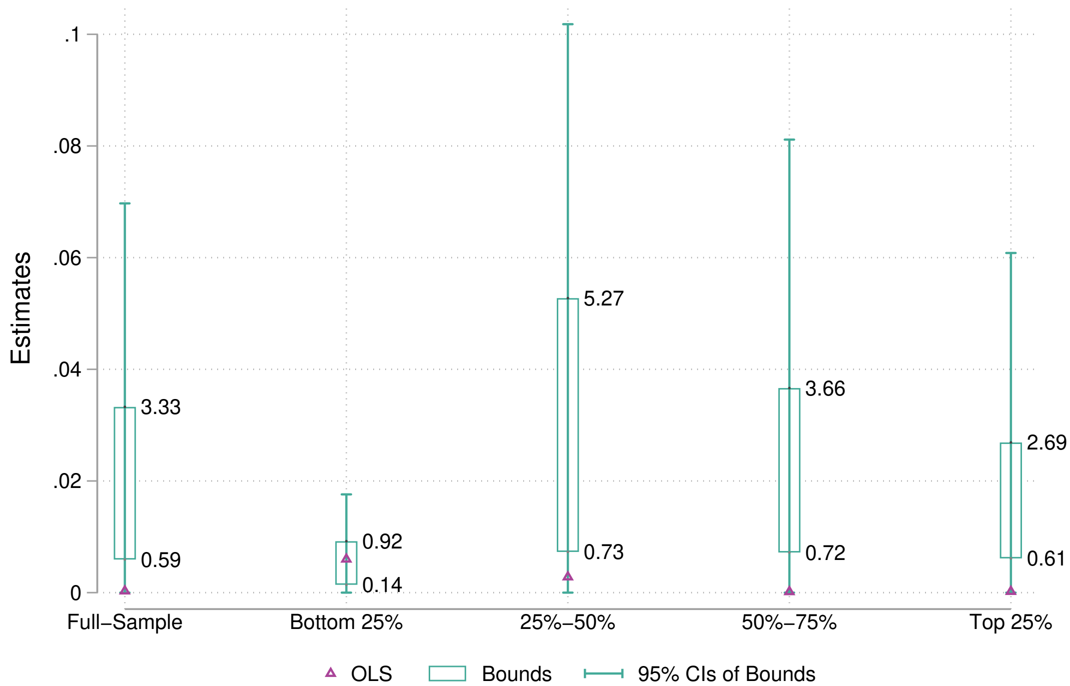
A.5: Results of the IV Validity Test, Alternative Marginal Significance Level

Notes: Figure A.5 shows the testing results for IV validity using sub-samples that are split by grid. The marginal significance level is set at 0.0018%. In other words, instead of controlling the FWER at around 10%, setting the marginal significance level at 0.0018% controls the FWER at around 5%. The outcome variable is the adjusted accident rate. The treatment variable is either the binary indicator of the number of days with PM2.5 pollution according to the WHO standard (PM2.5$>15\mu g/m^3$) or the binary indicator of days with moderate PM2.5 pollution based on the EPA standard (PM2.5$>12\mu g/m^3$). The instrumental variable is either the rainfall indicator (Panel a) or the wind direction indicator (Panel b). The green grid implies that the null hypothesis is not rejected, while the red grid suggests that the null hypothesis is rejected. White grid indicates no test is conducted, either because the outcome distribution reduces to a masspoint of zero or because the distribution of the IV indicator is degenerated to zero. Grids that are not in the sample are hatched.



A.6: Point Estimates and the 95% Confidence Intervals of Bounds by Industry.

Notes: Figure A.6 summarizes the estimated nonparametric bounds on the effect of PM2.5 pollution classified by the EPA criteria ($T1$) by industry, employing the wind direction IV. The definition of the treatment variable follows Equation 2. Numbers aside the boxes presents the estimated bounds in percentage points.



A.7: Point Estimates and 95% Confidence Intervals of Bounds by Population Density.

Notes: Figure A.7 summarizes the estimated nonparametric bounds on the effect of PM2.5 pollution defined by the EPA standard ($T1$) by population density, employing the wind direction IV. The definition of the treatment variable follows Equation 2. Numbers aside the boxes presents the estimated bounds in percentage points.

B Tables

B.1: Industry Category Crosswalk

NAICS Code	Industry Title	Industry Name in Paper
Goods-Producing Industries:		
11 21	Agriculture, Forestry, Fishing and Hunting Mining	Natural Resources & Mining
23	Construction	Construction
31-33	Manufacturing	Manufacturing
Service-Providing Industries:		
42 44-45 48-49 22	Wholesale Trade Retail Trade Transportation and Warehousing Utilities	Trade, Transportation, and Utilities
51	Information	Information
52 53	Finance and Insurance Real Estate Rental and Leasing	Financial Activities
55 56 54	Management of Companies and Enterprises Administrative and Support and Waste Management and Remediation Services Professional, Scientific, and Technical Services	Professional & Business Services
61 62	Educational Services Health Care and Social Assistance	Education & Health
71 72	Arts, Entertainment, and Recreation Accommodation and Food Services	Leisure & Hospitality
81	Other Services (except Public Administration)	Other Services
92	Public Administration	Public Administration

Note: Table B.1 lists the crosswalk mapping the 21 industries classified by 2-digit NAICS to the 11 industry sectors defined by BLS. See BLS's website for more details. The industry category used in this paper is the same as BLS's classification.

B.2: Summary Statistics Appendix

Variable Name	N	Mean	Std.	Min	Max
Accident Rate per 10,000 population					
All industries	132960	0.102	2.610	0	338.26
By Industry					
Natural Resources & Mining		0.042	1.951	0	270.61
Construction		0.015	0.734	0	148.37
Manufacturing		0.016	0.217	0	37.32
Trade/Transportation/Utilities		0.013	0.570	0	135.30
Information		0.000	0.020	0	4.06
Financial Activities		0.001	0.186	0	67.65
Professional and Business Services		0.004	0.253	0	76.84
Education & Health		0.002	0.094	0	25.23
Leisure & Hospitality		0.004	0.567	0	148.47
Other Service		0.001	0.201	0	67.65
Public Administration		0.004	0.838	0	301.88
Top-Coded Accident Rate					
All industries	132960	0.070	0.605	0	18.90
By Industry					
Natural Resources & Mining		0.020	0.321	0	9.63
Construction		0.008	0.101	0	3.17
Manufacturing		0.014	0.096	0	2.26
Trade/Transportation/Utilities		0.008	0.086	0	2.62
Information		0.000	0.004	0	0.15
Financial Activities		0.000	0.003	0	0.11
Professional and Business Services		0.002	0.029	0	0.92
Education & Health		0.001	0.013	0	0.40
Leisure & Hospitality		0.001	0.020	0	0.68
Other Service		0.000	0.008	0	0.25
Public Administration		0.001	0.017	0	0.62
Alternative Treatment Variable					
PM2.5 Indicator					
$T1$: EPA standard 12 $\mu g/m^3$		0.448	0.497	0	1
$T2$: WHO standard 15 $\mu g/m^3$		0.356	0.479	0	1
Alternative Instrumental Variable					
Z_1: Rainfall		0.394	0.489	0	1
Z_2: Wind Direction		0.453	0.498	0	1

Note: In Table B.2, I present the mean, standard deviation, maximum, and minimum of the accident outcome variables by industry. To eliminate outliers and ensure the BOS assumption, I define a top-coded (at 0.05th percentile) accident rate. The bottom panel also presents the summary statistics for treatment variables and instrumental variables defined by an alternative cutoff – the median of the sample distribution of a given variable.

B.3: Robustness Check: Parametric Bounds Estimation

	(1) *Lower Bound*	(2) *Upper Bound*	(3) *Lower Bound*	(4) *Upper Bound*
	Outcome: Workplace Accident Rate (per 10,000 ppl)			
	T1: EPA standard 12 $\mu g/m^3$		T2: WHO standard 15 $\mu g/m^3$	
Panel A. IIV Z1: Rainfall				
Point Estimates	[-0.0002,	0.0006]	[-0.0003,	0.0001]
95% CI of Identified Set	[-0.0199,	0.0020]	[-0.0402,	0.0013]
Panel B. IIV Z2: Wind Direction				
Point Estimates	[-0.0038,	0.0005]	[-0.0040,	0.0001]
95% CI of Identified Set	[-0.0198,	0.0019]	[-0.0209,	0.0013]
N	132960	132960	132960	132960

Note: Bounds on the average treatment effect of PM2.5 pollution on the workplace accident rate. Parametric bounds are estimated by the method of Nevo and Rosen (2012), employing two invalid IV (IIV). The treatment variable is a binary dummy indicating whether the number of air-polluted days in a month is above a certain threshold. As a robustness check, the cutoff are set at the 50th percentile of the sample distribution of the number of air-polluted days. All specifications include time-varying weather controls, along with grid and month by year fixed effects. All models are weighted by grid-level population density in 2015.

B.4: Summary Statistics: Intertemporal Labor Supply and Rainfalls

	N	Mean	Std.	Min	Max
Intertemporal Labor Supply:					
Working Time (Narrow)	175893	155.27	239.73	0	1380
Working Time (Broad)	175893	156.87	240.24	0	1380
1(Absence)	175893	0.03	0.16	0	1
Weather:					
$1(Prcp > 7.6)$	175893	0.31	0.46	0	1
Minimum Temperature (C)	140059	7.56	10.05	-38	33
Maximum Temperature (C)	140059	20.31	10.67	-22	49
Accumulated Precipitation (mm/day)	140059	3.05	7.91	0	185
Day Length (s/day)	140059	43077.26	6685.32	28921	57432
$1(PrcpLastWeek > 30)$	175893	0.40	0.49	0	1
# Days Max Temp ¿35 Last Week	140059	0.32	1.23	0	7
Weekly Avg Precipitation Last Week	140059	21.11	27.69	0	899
Weekly Avg Day Length Last Week	140059	43051.70	6688.40	29053	57468
Covariates:					
Diary day a holiday	175893	0.02	0.13	0	1
Male	175893	0.44	0.50	0	1
Married	175893	0.50	0.50	0	1
Has Child ¡18	175893	0.45	0.50	0	1
Age	175893	47.84	17.90	15	85
% Age ¿65	175893	0.19	0.39	0	1
Employed Part-Time	175893	0.13	0.34	0	1
Paid Hourly	175893	0.30	0.46	0	1
% Reside in Urban Area	175893	0.83	0.38	0	1
% Hispanic	175893	0.14	0.35	0	1
% Black	175893	0.14	0.35	0	1
% Asian	175893	0.04	0.19	0	1
% ¡ High School	173876	0.14	0.34	0	1
% High School Graduate	173876	0.26	0.44	0	1
% Some College	173876	0.48	0.50	0	1

[1] Variables “Working Time (Narrow)” and “Working Time (Broad)” are measured in minutes.

B.5: The Effect of Rainfalls on Workers’ Intertemporal Labor Supply

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Working Time (Narrow)			Working Time (Broad)			$\mathbb{1}(Absence)$	
	Employed	Outdoor	Outdoor & Male	Employed	Outdoor	Outdoor & Male	Employed	Outdoor
$\mathbb{1}(Prcp_{ct} > 7.6)$	0.0628	-19.70***	-16.17**	0.00808	-20.12***	-16.47**		
	(2.229)	(7.057)	(7.980)	(2.230)	(7.057)	(7.982)		
$\mathbb{1}(Prcp_{cw} > 30)$							0.00333**	0.00913*
							(0.00161)	(0.00493)
N	86690	9929	7936	86690	9929	7936	86690	9929

[1] Standard errors in parentheses. * p¡0.10 ** p¡0.05 *** p¡0.01. Variables “Working Time (Narrow)” and “Working Time (Broad)” are measured in minutes.

C More on IV Validity Test

This section describes the implementation of the instrument validity test with more details. Recall the notation, the outcome variable Y denotes the rate of workplace accidents that involve severe workplace injuries per million population, while the treatment variables $T1$ and $T2$ are binary indicators indicating whether the number of PM2.5 pollution days in a month passes a pre-specified threshold. The testing involves two instrumental variables: the rainfall IV (denoted as $Z1$) and the wind direction IV (denoted as $Z2$) as defined in Section 4.2. Since the direction of the conditional moment inequalities is correlated with the direction of the monotonicity condition, inequalities (5)-(6) and (8)-(9) hold under the assumption of $T_0 \leq T_1$. For the two instrumental variables $Z1$ and $Z2$, the direction of monotonicity under testing is $T_1 \leq T_0$. Therefore, to directly conduct the testing for the two proposed IVs, it is needed to shift the direction of the conditional moment inequalities. Because the *clrtest* command tests the null hypothesis that a collection of lower intersection bounds is no greater than zero, adjustment on either the test statistics or the instrumental variable is required for the intersection bound test to perform correctly. Thus, I define the instrumental variable for the test — $Z1'$ and $Z2'$ and test the joint conditions LI and LM for $Z1$ and $Z2$ using $Z1'$ and $Z2'$, respectively.

For each treatment and instrumental variable, I implement the testing procedures separately. To be more specific, I implement the testing procedures according to the following steps:

1. Divide the sample into S sub-samples. Specifically, I divide the sample into $S = 2770$ sub-samples and each sub-sample s consists of data spanning the 48 month-year for each grid. In other words, the panel data is reduced to a set of time-series data for each grid. Within each sub-sample, code each observation as $i = 1, ..., n$, and the total number of observations is n.

2. For each sub-sample, conduct the test as follows:

   (a) Estimate c_1 and c_0 by $\hat{c}_1 = \frac{1}{n}\sum_{i=1}^{n} Z_i$ and $\hat{c}_0 = 1 - \hat{c}_1$, respectively.

   (b) Define $\hat{L}_i^1 = \hat{c}_1 T_i(1 - Z_i) - \hat{c}_0 T_i Z_i$ and $\hat{L}_i^0 = \hat{c}_0(1 - T_i)Z_i - \hat{c}_1(1 - T_i)(1 - Z_i)$.

   (c) Run the *clrtest* command (Chernozhukov et al. 2015) using the two moment inequalities. Specify $\hat{L}_i^1$ and $\hat{L}_i^0$ as the dependent variables for each conditional inequality, respectively. For the conditional independence assumption, specify $\hat{u}_i$ as the conditioning variable for both inequalities. For the unconditional independence assumption, specify Y_i as the conditioning variable.

3. Define the null hypothesis as $H_0 = H_0^{(1)} \cap H_0^{(2)} \cap \cdots \cap H_0^{(S)}$, where $H_0^{(s)}$ suggests that the moment inequalities hold for sub-sample s. Rejection of any $H_0^{(s)}$ implies the rejection of H_0.

A flow chart illustrating the testing procedures is shown in Figure A.3. Notice that for the test of the conditional independence assumption, instead of conditioning on Y, I estimate and condition on the adjusted accident rate $\hat{u}$ which can be viewed as the dependent variable for which the unit-specific and time-specific confounders are removed. $\hat{u}$ is the residuals of a fixed-effects regression using the grid-month panel data:

$$Y_{im} = \gamma_i + \tau_{my} + X_{im} + u_{im}$$

where Y_{im} denotes the accident rate and X_{im} denotes a vector of weather controls including monthly average temperature and wind speed, etc. γ_i and τ_{my} represent the grid and month by year fixed effects, respectively. The intersection bound test is based on local linear estimation. The choice of bandwidth and kernel functions is the same as in Mourifié and Wan (2017) and is recommended by Chernozhukov, Lee, and Rosen (2013) and Chernozhukov et al. (2015). Because the test is conducted for a large number of sub-samples, it is necessary to control the family-wise error rate (FWER) at targeted levels. I follow Mourifié and Wan (2017) and Holm (1979) and adapt a multiple testing procedure. For each test on the sub-sample, I subsequently check whether $H_0^{(s)}$ is rejected at the significance level 10%, 5%, 1%, 0.5%, and 0.003% to find the marginal level of rejection. Any rejection of $H_0^{(s)}$ at the significance level 0.003% implies that H_0 is rejected with the FWER controlled by no greater than $0.003\% \times 2770 = 8.3\%$.[46] I show in Figures A.4 and 3 the testing results for sub-samples that are divided by grid and treatment variables $T1$ and $T2$, respectively. The marginal significance level is set up at 0.003%. Green grid denotes that the null hypothesis is not rejected, while red grid denotes that H_0 is rejected. No test is conducted in white grids, either because the outcome distribution reduces to a masspoint of zero or because the distribution of the IV indicator is degenerated to zero. We can infer from the existence of red grids that the null hypothesis is rejected with the FWER controlled by no more than 8.3%.

[46]Any number below $10\%/2770 = 0.0036\%$ should work as a marginal significance level. I arbitrarily pick the marginal significance level 0.003% to control the FWER at around 10%.

D The Effect of Rainfalls on Workers' Intertemporal Labor Supply

D.1 Data and Empirical Strategy

Data used for this preliminary analysis combine two data sources: the American Time Use Survey (ATUS) data and the weather data. The American Time Use Survey collects information about how individuals spend their time during a diary day. The sample is randomly selected from a subset of households that have completed their eighth month of interviews for the Current Population Survey (CPS). The ATUS data provides information about workers' intertemporal labor supply (hours worked and work absence status) and various demographical, educational, and employment characteristics at the individual level. More importantly, it contains information about the diary day's date and the respondent's geographic location, which allows researchers to link the ATUS data with the weather data.

The weather data come from Daymet (Thornton et al. 2020), which provides weather parameters of daily surfaces of minimum and maximum temperature, precipitation, vapor pressure, radiation, snow water equivalent, and day length on a 1 km $\times$ 1 km gridded surface. I map these gridded daily weather parameters to counties using the longitude and latitude coordinates of the centroid of each county.

I merge the 2005-2019 ATUS data with the Daymet weather data using the county FIPS or the CBSA/MSA code.[47]

I am interested in whether and to what extent employed workers change their work schedule in regards to rainfalls. I consider two working time outcomes: (a) a narrow definition that measures time spend on work and work-related activities; (b) a broad definition that measures time spend on not only work and work-related activities but also on income-generating activities.[48] I call the definition (a) "Working Time (narrow)" and the definition (b) "Working Time (broad)". The treatment variable is defined as a day with heavy rain if the daily accumulated precipitation is greater than 7.6 mm (same as the definition of the precipitation IV).

Let's use i to denote a respondent, t to denote a diary day, and c to denote the geographic unit the respondent resides. I estimate the following model:[49]

$$Y_i = \beta_0 + \beta_1 \mathbb{1}(Prcp_{ct} > 7.6) + \gamma X_i + \sigma Z_{ct} + f(month, year, dow, c) + \varepsilon_i \qquad (12)$$

where Y_i measures the hours-worked (in minutes) for individual i on a diary day and $\mathbb{1}(Prcp_{ct} > 7.6)$ is the heavy rain indicator for location c on day t. I control for other time-varying weather attributes (Z_{ct} that are potentially correlated with working hours, such as daylight and daily maximum and minimum temperature. X_i denotes a vector of individual-level covariates as listed in Table B.4. $f(month, year, dow, c)$ includes day of week dummy variables to account for differences in schedules

[47]I am able to identify the geographic location for $140,059$ individuals (80% of the whole sample). For these observations, about 56% of them are identified by the county FIPS code, 40% of them are identified by the CBSA/MSA code, and 4% of them are identified by the NECTA code.

[48]For work and work-related activities, I include time spend on activities in Categories 0501 and 0502. In the broad definition, I further add time spend on other income-generating activities, i.e., the Category 0503. See American Time Use Survey Activity Lexicon 2003-2019 for more details.

[49]This is a similar model specification as in Connolly (2008) and Graff Zivin and Neidell (2014).

throughout the week and year and month dummy variables to control for seasonal and annual time trends in the intertemporal labor supply. It also includes location fixed effects that capture all time-invariant observable and unobservable attributes that affect labor supply decisions. The parameter of interests is β_1, which captures the changes in the hours-worked in days with heavy rainfalls.

The ATUS has a question about whether the respondent were absent from work in the past week. Thus, I construct a binary indicator which equals one if the respondent were absent from work during the past seven days. I further calculate the accumulate precipitation in the past week. To examine the effect of rainfalls on work absenteeism, I estimate the following model:

$$\mathbb{1}(Absence) = \alpha_0 + \alpha_1 \mathbb{1}(Prcp_{cw} > 30) + \gamma X_i + \sigma W_{cw} + f(month, year, c) + \varepsilon_i \tag{13}$$

w denotes the week before the diary day t. Since the absent from work indicator is measured in the past week, I change the treatment variable from a binary indicator of daily precipitation to a binary indicator of accumulated precipitation in the past week and the threshold is changed from 7.6 mm to 30 mm.[50] Daily weather attributes Z_{ct} are changed to weekly weather attributes Z_{cw}, including weekly mean daylight, the number of days when daily maximum temperature is greater than 35 Celsius. I also remove the day of week fixed effects from Equation (13). Similar to β_1, α_1 captures the changes of the probability of work absenteeism with respect to the heavy rainfalls.

Lastly, all models are conditional on employment, i.e., applying to employed individuals. All estimates are weighted by the ATUS final weight.

D.2 Main Findings

Table B.4 shows the summary statistics. I summarize estimation results in Table B.5. Columns (1) - (3) present results for the narrowly defined hours-worked outcome, Columns (4) - (6) report results for the broadly defined hours-worked outcome, and the last two Columns show results for the absent from work indicator.

First, I estimate Equation (12) using the employed individual sample. Results are shown in Columns (1) and (4). I do not find statistically significant effect of heavy rainfalls on hours-worked. Restricting to respondents who work in industries that usually requires workers to perform tasks in outdoor or semi-outdoor environments, I find workers work about 20 minutes fewer on days with heavy rainfalls, see Columns (2) and (5).[51] Columns (3) and (6) further restrict the sample to male workers and suggest that heavy rain is associated with a reduction in hours-worked of about 16 minutes.

Then, I consider the effect on work absenteeism. Column (8) suggests that having heavy rainfalls in the past week is associated with a statistically significant increase in the probability of work absence of about 0.3 percentage points for employed respondents. For outdoor workers,

[50] Recall that daily accumulated precipitation at 7.6 mm is equivalent to having a heavy rain for about an hour. Thus, having weekly accumulated precipitation of 30 mm is equivalent to having heavy rainfalls for an hour for three days a week.

[51] These industries are Agriculture, forestry, fishing, and hunting, Mining, Construction, and Transportation and utilities. I classify respondents' working industry using the major industry code for the main job provided by the ATUS.

the effect is larger at about 0.9 percentage points and is statistically significant at 10% significance level.

Using the ATUS data in 2003 and 2004, Connolly (2008) do not find evidence that female workers adjust their intertemporal labor supply on rainy days, while on average male workers work about 30 minutes more on rainy days. Results in Table B.5 are not directly comparable to the findings in Connolly (2008) as i) although both of us use the ATUS data, we consider different sample periods (2005-2019 in this paper and 2003-2004 in Connolly (2008)); ii) we consider different treatments: Connolly (2008) define the treatment as having rainfalls of at least 0.1 inches over 24 hours, while I consider the treatment as heavy rainfalls with daily precipitation over 7.6 mm.