FINE PARTICULATE AIR POLLUTION AND ACCIDENT RISK: THREE ESSAYS

By

Matthew Lindsey Chambers

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

May 14, 2021

Nashville, Tennessee

Approved:

Matthew Zaragoza-Watkins, Ph.D.

W. Kip Viscusi, Ph.D.

Andrea Moro, Ph.D.

Yuya Sasaki, Ph.D.

To Stephanie, for her love and support. And to Kylie and Ellie, for hugs and smiles and being the most adorable distractions any father could have.

ACKNOWLEDGMENTS

I am deeply indebted to my advisor Matthew Zaragoza-Watkins, as well as to Kip Viscusi, Yuya Sasaki, and Andrea Moro for their invaluable suggestions and guidance. I am grateful to Arden Pope for his comments and feedback. I also thank the Vanderbilt Spatial Analysis Research Laboratory and Samantha Turley for crucial GIS support, and Bing Yang Tan for kindly providing data on thermal inversions, as well as helpful comments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

PM 2.5 and Construction Worksite Safety

This paper provides the first causal estimates of the impact of fine particulate matter air pollution (PM 2.5) on workplace accidents, which cost the U.S. economy an estimated \$273.3 billion in 2019. I construct a novel dataset composed of the universe of serious workplace accidents investigated by OSHA from 2003 to 2015 and exploit plausibly exogenous variation in PM 2.5 caused by thermal inversions to implement an instrumental variables research design. Focusing on the U.S. construction sector, which employed 7.5 million workers (4.6% of the U.S. workforce) and accounted for 5.7% of nonfatal and 19.9% of fatal U.S. workplace injuries in 2019, I find that decreasing PM 2.5 exposure by 1 $\mu g/m^3$ could lead to a 7% decrease in the risk of a serious construction worksite accident, representing an elasticity of 0.67.

1.1 Introduction

Fine particulate matter air pollution (PM 2.5) has been shown to impair cognition, health and productivity.[1] A growing awareness of these relationships has led policymakers in the United States and around the world to adopt regulations designed to limit PM 2.5 emissions. However, the variety of effects that PM 2.5 has on individuals and the economy are not fully understood, nor are the mechanisms which drive those effects. Identifying these effects and quantifying their costs is key to determining optimal pollution control policies.[2]

In this paper, I study a novel channel through which short-run exposure to PM 2.5 imposes costs on society: by increasing the probability of workplace accidents. According to the 2019 Census of Fatal Occupational Injuries (BLS 2020a) and Survey of Occupational Injuries and Illnesses (BLS 2020b), 5,333 fatal and 3,496,700 nonfatal workplace accidents, costing about \$273.3 billion,[3] occurred that year. In this paper, I specifically study the effect of PM 2.5 on construction worksite accidents, which accounted for 5.7% of nonfatal workplace injuries and illnesses and 19.9% of fatal workplace injuries in 2019 (BLS 2020a; BLS 2020b).

I utilize a unique dataset constructed by scraping OSHA's database of fatality and catastrophe investigations; data from the Quarterly Census of Employment and Wages (QCEW) and North American Regional

[1] An extensive literature documents the negative effects of PM 2.5 on physical health outcomes, including infant mortality and adult mortality due to respiratory disease, lung cancer, and cardiopulmonary disease (for example, see Chowdhury and Dey 2016; Deryugina et al. 2019; Dominici et al. 2006; Pope III et al. 2019; Woodruff et al. 2006). More recently, links between PM 2.5 and mental health and cognitive function have begun to surface. Even in the very short term, exposure to PM 2.5 has been found to affect cognitive ability (Ebenstein et al. 2016), and criminal behavior and anxiety (Lu et al. 2018). These cognitive/behavioral impacts, in turn, can lead to perhaps unexpected consequences, such as an increase in violent crime (Burkhardt et al. 2019) or in car accidents (Sager 2019).

[2] Aldy et al. (2020) document the influence of so-called "co-benefits," benefits other than the intended target of a regulation, on regulatory decision making. They focus specifically on the EPA, and develop a conceptual framework to demonstrate that any and all benefits (or costs) resulting from a proposed regulation should be accounted for equally as part of its benefits (or costs) rather than being relegated to "co-benefit" or "co-cost" status. In the case of proposed PM 2.5 regulations, this implies that all pathways by which PM 2.5 harms society (pathways by which abatement benefits society) should be identified and treated equally as benefits.

[3] In 2019 dollars. This calculation is based on a value of a statistical life (VSL) of \$10.6 million, from Viscusi (2018) (updated to 2019 dollars), and a value of a statistical injury (VSI) of \$62,000, from Viscusi and Aldy (2003) (updated to 2019 dollars).

Reanalysis (NARR); and a new, high resolution, daily PM 2.5 prediction dataset produced by Di et al. (2019). From these sources, I construct a daily panel covering all counties in the contiguous United States from January 1, 2003 to December 31, 2015.

Because PM 2.5 is largely generated by human activities, it is likely co-determined with other activities that impact worksite safety. To identify the causal effects of PM 2.5, I implement an instrumental variables research design, using thermal inversions (a meteorological phenomenon in which air above the earth's surface is warmer than air near the surface) as an instrument.[4] A thermal inversion reduces the upward circulation of air, which would otherwise tend to disperse air pollution. The identifying assumption of my research design is that, after controlling for surface temperature and precipitation, the difference in temperature between air at ground level and air 200 m above ground level has no effect on worksite accident risk except by its influence on air pollution.

I find that a significant portion of serious construction worksite accidents may be caused by short-run exposure to PM 2.5. Estimates from my primary specification, a linear panel fixed effect IV model, suggest that reducing construction workers' PM 2.5 exposure by 1 μg/m^3 could reduce the probability of a serious construction worksite accident by 7%, an elasticity of 0.67. By a back-of-the-envelope calculation, having 7% fewer workplace accidents (if a similar decrease were observed economy-wide) would save at least $19 billion annually. Across a variety of specifications, I find evidence for the hypothesis that PM 2.5 exposure increases the probability of worksite accidents, and that the size of this effect increases with the exposure level.

These results indicate that government interventions aimed at PM 2.5 reduction have as an important benefit a reduced risk of serious worksite accidents. This is a benefit that should be equally counted with all other benefits when evaluating proposed PM 2.5 regulations, especially since they may more effectively reduce workplace accident risk than regulations directly targeting workplace safety do.[5]

I also identify an important benefit to firms of decreased PM 2.5. Government regulations, including both safety regulations and environmental regulations, impose costs on firms. Debate over these regulations has been frequent and contentious. Because this paper examines a novel pathway by which pollution can influence firm profitability (via effects on worker safety), it draws attention to a set of possible benefits to firms from environmental regulation. Firms will benefit not only from their own adherence to pollution

[4] In using thermal inversions as an instrument, I follow Sager (2019). Deryugina et al. (2019) and Heyes and Zhu (2019) use wind direction as an instrument for PM 2.5. Ebenstein et al. (2016) and Archsmith et al. (2018) use advanced and inflexible scheduling of events as a source of exogenous variation in PM 2.5 exposure. Burkhardt et al. (2019) assume that PM 2.5 generation is orthogonal to criminal decision making.

[5] See Kniesner and Leeth (2013) for an overview of several studies finding little to no effect of OSHA regulations and enforcement on workplace accident rates. Some of the more optimistic studies find that OSHA inspections with penalties for violations have a stronger preventive effect on workplace accidents than those without, and that citations for failure to properly use personal protective gear have the strongest effect (A. Haviland et al. 2010; Mendeloff and Gray 2005). A. M. Haviland et al. (2012) suggest that inspections with penalties tend to reduce injuries for up to 2 years in midsize establishments (between 20 and 250 employees).

regulations, but will also receive positive externalities from other firms doing the same.

1.2 Background

1.2.1 Fine Particulate Matter (PM 2.5) and Health

Fine particulate matter (PM 2.5) air pollution consists of particles less than 2.5 micrometers (μm) in diameter; for comparison, the average human hair is 50-70 μm in diameter. These very small particles arise from a variety of sources. Particles emitted directly into the atmosphere are referred to as "primary" particles. "Secondary" particles form in the atmosphere from chemicals emitted into the atmosphere as gases, and may form some distance from where the gases were originally emitted. These secondary particles constitute a large proportion of PM 2.5, as opposed to coarser particulate matter, which contains more primary particles. Major contributors to the formation of PM 2.5 secondary particles are sulfur dioxide (SO2), which is emitted by the combustion of fossil fuels and the smelting of metal ores containing sulfur, and nitrogen oxides (NOx), which are emitted by high temperature combustion, such as that found in vehicle engines and power plants. These and other anthropogenic emissions account for a large portion of ambient PM 2.5.

Because of the small size of PM 2.5 particles, they are absorbed into the bloodstream when inhaled, rather than being filtered out by the lungs. They are then transmitted throughout the body. An extensive literature documents the negative physiological health effects of chronic exposure to PM 2.5, including increased likelihood of death due to heart disease, heart attack, and lung cancer (see Achilleos et al. 2017). Largely due to these negative physiological health effects, both coarse particulate matter (PM 10) and fine particulate matter (PM 2.5) air pollution are listed as criteria pollutants under the National Ambient Air Quality Standards (NAAQS) established by the EPA.[6]

Some of the particles making up PM 2.5 are small enough to cross from the bloodstream into the brain, where they can impact cognitive and mental health. The effects of PM 2.5 on cognition and the brain have been well documented in the last few years. Gatto et al. (2014) and Ailshire and Crimmins (2014) show that long term exposure to PM 2.5 is connected to decreased cognitive functioning in middle aged and older adults in the United States.[7] Weuve et al. (2012) find that chronic PM 2.5 exposure leads to an increased rate of decline in cognitive ability with age. Fonken et al. (2011), in a medical study using mice, find that mice chronically exposed to elevated levels of PM 2.5 displayed depressive behavioral symptoms, in addition to decreased cognitive ability, specifically impairments in spatial learning and memory. They describe the specific brain regions that were found to be affected, including the hippocampus, one of the centers of both spatial learning and memory formation. Block and Calderón-Garcidueñas (2009) and Costa et al. (2014)

[6]The current standard for PM 2.5 is 12 μg/m^3 (3-year average of annual mean concentration), significantly above the national mean PM 2.5 concentration (as of 2019) of 7.65 μg/m^3.

[7]The Gatto, et al. study is confined to Los Angeles.

suggest that PM 2.5 causes brain inflammation and oxidative stress, leading to various neuropathologies and central nervous system diseases.

The negative physical and cognitive/mental health effects of PM 2.5 are not only seen in those who experience chronic PM 2.5 exposure. Acute (short term) PM 2.5 exposure has also been shown to impact both physical and cognitive/mental health. Deryugina et al. (2019) use a large-scale database of Medicare recipients, and find that acute PM 2.5 exposure causes increased mortality, emergency hospitalization, and medical spending for up to two days following exposure. Ebenstein et al. (2016) find that acute exposure to elevated PM 2.5 levels causes an immediate decrease in cognitive ability. They use evidence from standardized testing of pre-college teens in Israel to show that elevated PM 2.5 at the testing location on test day causes lower scores, with subsequent persistent effects on educational attainment and earnings. Heyes and Zhu (2019) find that elevated levels of PM 2.5 lead to significantly increased sleeplessness, another possible pathway for physiological and cognitive harm.

Likely due to its impacts on cognition and behavior, short-term exposure to PM 2.5 has been shown by Burkhardt et al. (2019) to lead to an increase in violent crime. Interestingly, they find that lagged daily PM 2.5 has no impact on crime, suggesting that there are immediate, non-cumulative behavioral and cognitive effects of PM 2.5, although we have seen that chronic exposure has its own effects. Sager (2019) shows that short-term exposure to PM 2.5 causes an increase in vehicular accidents, and similarly to Burkhardt, et al., finds that lagged daily PM 2.5 has no impact.[8]

1.2.2 Worker Productivity and Safety

Fine particulate air pollution (PM 2.5) has been shown to directly impact the productivity of workers engaged in physical labor or mental labor, indoors or out. T. Chang et al. (2016) study pear packers engaged in manual labor inside a factory. They find that elevated levels of PM 2.5, which readily moves indoors, causes a decrease in worker productivity, while other pollutants, such as ozone, which do not readily move indoors, do not have an effect on indoor workers. They find that labor supply decisions appear to be unaffected by PM 2.5 levels, which they take as evidence that workers are not self-selecting their PM 2.5 exposure. Adhvaryu et al. (2014) find similar results among workers in a garment factory in India; additionally, they find substantial heterogeneity in treatment effects based on task and manager identity, suggesting that pollution impacts to productivity can be somewhat mitigated if managers choose to do so. Archsmith et al. (2018) use advanced scheduling of MLB baseball games as an instrument for umpire's PM 2.5 exposure and find that short-term PM 2.5 exposure leads to umpires, who work (mostly) outdoors, making more incorrect calls. T. Y. Chang

[8] This also suggests that the sleeplessness pathway proposed by Heyes and Zhu is not the only pathway by which PM 2.5 affects cognition and behavior.

et al. (2019) find that call center workers, who are indoors and performing almost entirely mental labor, are also impacted by PM 2.5 levels. It seems plausible that safe work practices could be affected along with productivity, whether via cognitive or other pathways.

OSHA, the U.S. government agency tasked with promulgating and enforcing safety rules and regulations to reduce the frequency of workplace accidents, regulates workplace exposure to a variety of toxic chemicals, and to dust in general, but does not regulate fine particulate matter. Similarly, literature on the connection between workplace safety and pollution has mainly focused on the health effects of pollution (especially toxic chemicals) generated by the workplace (Landrigan 1992) and the existence of compensating wages for workers in highly polluting industries (Cole et al. 2009), while ambient particulate air pollution as a potential cause of workplace accidents has not been studied. To my knowledge, this is the first paper which quantifies the effect of PM 2.5 on workplace accident risk.

1.3 Research Design

1.3.1 Data

To analyze the effect of acute PM 2.5 exposure on worksite accidents, I require data on workplace accidents that includes, at minimum, the date of each accident, the county it occurred in, and the employer's industrial classification. I make use of a unique dataset of OSHA Fatality and Catastrophe Investigation summaries, which provide the address where each incident occurres, the date it occurred, and the NAICS-6 industry classification of the employer. Addresses were linked to pollution and weather information using geolocation.[9]

Additionally, the victims in each incident are enumerated and categorized according to whether they were killed, hospitalized, or neither. While a fatality is (relatively) easy to define, catastrophes were defined by OSHA in 2005 as "the hospitalization of three or more employees resulting from a work-related incident or exposure" (OSHA 2005). However, more than half of the incidents in the Fatality and Catastrophe Investigation database do not fit this definition, for reasons that are unclear. For my main results, I use the set of all construction worksite accidents reported in the OSHA Fatality and Catastrophe Investigation database, but due to uncertainty surrounding the definition of "catastrophe," and the fact that this definition could have changed over time, I conduct a robustness check using a restricted dataset consisting of fatal accidents only.

Data on employment come from the Quarterly Census of Employment and Wages (QCEW). The QCEW is conducted quarterly, as the name suggests, and covers 95% of employees in the United States,[10] providing information on total employment by industry within each county. These data are drawn from each state's

[9]The address data from the investigation summaries is very clean. The Spatial Analysis Research Laboratory (SARL) at Vanderbilt University geocoded the addresses using ArcGIS and was able to assign a precise location to all but 172 addresses (about .015% of the total).

[10]There are minor exclusions for some government, agricultural, and railroad employees and self-employed persons.

unemployment insurance accounting program, which receives quarterly reports from each employer in the state, with employment and wage data for each of their establishments within the state. These statistics are reported by month within each quarter, so I build a dataset of monthly employment data by county and industry from the QCEW data.

For PM 2.5, I make use of a newly available dataset (Di et al. 2019) that uses a variety of data sources[11] and models of how pollutants are dispersed and transported through the atmosphere. The authors use an ensemble of machine learning algorithms to estimate daily PM 2.5 levels in 1 km by 1 km pixels across the contiguous United States for the years 2000 to 2015. Their final model predicts PM 2.5 levels well, with a 10-fold cross validated R^2 averaging 0.86. Directly monitored PM 2.5 has two major disadvantages for my analysis: first, not all counties have PM 2.5 monitors; second, many monitors do not collect (or at least do not report) data every day. Using the Di et al. (2019) PM 2.5 prediction dataset allows me to analyze the effect of PM 2.5 on worksite accidents across the entire contiguous U.S., and is more accurate than the interpolation methods I would have to use with monitor data to fill in missing days.

Temperature and precipitation data, used to identify thermal inversions and control for surface temperature and precipitation, come from the North American Regional Reanalysis (NARR) dataset (Mesinger et al. 2006). The NARR uses historical data from a variety of sources[12] to create a detailed picture of weather and climate conditions in 32×32 km square pixels, at three hour intervals each day. I make use of inversion data calculated by Tan (2020), who compares temperatures at the surface (2 m) and roughly 200 m above the surface. When the higher level of the atmosphere also has a higher temperature, a thermal inversion exists. For each 3-hourly time stamp, Tan calculates the proportion of the county in which an inversion exists. I average these measures over all the time stamps in a day to get a daily measure of the proportion of the county-day covered by an inversion.

From these data sources, I construct a daily panel of U.S. counties covering the 48 contiguous U.S. states and the years 2003 to 2015. For each county-day, I include a dummy for whether a construction worksite accident occurred, mean PM 2.5, mean inversion coverage, level of construction employment and total employment, mean temperature, and mean precipitation. Table 1.1 contains summary statistics.

1.3.2 Methods

As discussed earlier, PM 2.5 and the chemicals which react to produce PM 2.5 in the air are largely human generated, from sources such as power plants, manufacturing, and vehicular traffic. Consequently, PM 2.5 exposure may be co-determined with other activities that contribute to construction worksite risk, leading to

[11] These include satellite images, ground-based monitor data, meteorological data, land-use data, and reanalysis datasets.

[12] These include satellites, radiosondes (balloon mounted instruments), dropsondes (instruments dropped from aircraft), and surface instruments

Table 1.1: Summary Statistics by Season

	Spring	Summer	Fall	Winter	Total
Accident Occurred	0.00094	0.00121	0.00102	0.00089	0.00102
	(0.0307)	(0.0347)	(0.0319)	(0.0299)	(0.0319)
PM 2.5 (μg/m^3)	8.514	10.850	8.174	8.260	8.954
	(4.960)	(6.348)	(5.354)	(5.323)	(5.631)
Inversion Coverage	0.380	0.444	0.247	0.444	0.379
	(0.279)	(0.336)	(0.299)	(0.348)	(0.327)
Temperature (C)	12.75	24.25	14.35	1.65	13.35
	(8.158)	(4.699)	(8.058)	(8.396)	(10.95)
Precipitation (mm)	2.791	3.066	2.3513	2.135	2.589
	(6.889)	(6.684)	(6.822)	(5.917)	(6.602)
Construction	1954.9	2067.3	2041.2	1902.3	1991.7
Employment	(6785.8)	(7000.1)	(6969.0)	(6686.2)	(6862.4)
Observations	3,717,168	3,717,168	3,676,764	3,645,684	14,756,784

endogeneity. To account for this endogeneity, I need an exogenous source of variation in PM 2.5 exposure to use in an instrumental variables research design.

Various sources of exogenous variation in PM 2.5 have been used in other studies, including advanced and inflexible scheduling of events, wind direction, and thermal inversions.[13] Following Sager (2019), I use thermal inversions. Under normal weather conditions, air in the troposphere [14] decreases in temperature with altitude. Because warm air is lighter than cool air, air near the surface tends to rise and cool, dispersing pollutants, such as PM 2.5, generated near the surface. This circulation pattern can be disrupted if the surface layer of air is overlaid by a warmer layer of air; this is known as a thermal inversion. The lack of vertical air circulation during an inversion causes PM 2.5 to remain near the surface of the earth longer than usual. An inversion therefore causes an increase in surface PM 2.5 exposure that is orthogonal to human activity.

Thermal inversions are a simpler instrument for PM 2.5 than wind direction because the sign of their effect on surface PM 2.5 exposure is unambiguous, making it unnecessary to develop an air transport model, like Heyes and Zhu (2019), or utilize non-parametric methods, like Deryugina et al. (2019). Primarily for this

[13] Ebenstein et al. (2016) use the Israeli Bagrut college entrance exam, which is compulsory for all high school students, is scheduled years in advance, and which may not be rescheduled. This exam also takes place across several days, and scores are recorded by day, enabling the authors to use variation in PM 2.5 across different days for each individual student to identify the effect of PM 2.5 on test scores. Archsmith et al. (2018) use umpire schedules for Major League Baseball (MLB) games, which are set months in advance and which cannot readily be changed. These are excellent sources of exogenous variation in PM 2.5, but such ideal sources are not common. Furthermore, they are not well-suited to analyzing already-rare events such as workplace accidents.

Wind has a strong effect on PM 2.5 exposure in many locations. The chief difficulty in using (lateral) wind direction as an instrument for PM 2.5 exposure is that its effect is completely dependent on location, since a location may have pollution sources distributed arbitrarily around the points of the compass. To account for this, Heyes and Zhu (2019), who confine their study to 19 of China's most populous cities, construct a simple air transport model for each city in their sample. In contrast, Deryugina et al. (2019) use a non-parametric specification which allows the effect of wind direction on PM 2.5 to vary across groups of counties and makes no ex ante assumptions about the effect of wind in any given county.

[14] The atmospheric layer closest to the surface, up to about 3-6 miles.

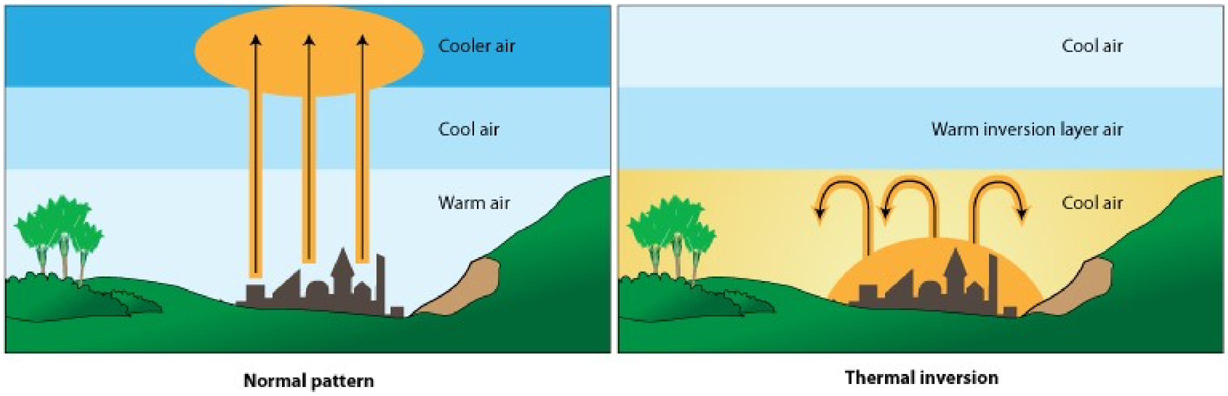

Figure 1.1: Thermal inversions and air circulation
Image credit: Science Learning Hub – Pokapū Akoranga Pūtaiao, University of Waikato, www.sciencelearn.org.nz

reason, I use thermal inversions as my instrument for PM 2.5 exposure.

A potential concern with using thermal inversions as an instrument is that inversions may be readily observable if they result in fog or trap enough pollution to create visible smog in the air. If this is the case, then people might change their behavior in response to observing an inversion. To check that this is not driving my results, I perform a robustness check by constructing a binary measure of nighttime inversions not followed by daytime inversions, and using this as an instrument for PM 2.5 exposure.[15]

1.3.3 Model Specification

For my main results, I use a panel fixed effect IV specification, with the second stage as follows:

$$Y_{it} = \mu + \alpha \hat{D}_{it} + \beta' \mathbf{X}_{it} + \varphi_i + \varepsilon_{it} \tag{1.1}$$

where Y_{it} is a binary variable indicating whether an accident occurred in county i on date t; $\hat{D}_{it}$ represents PM 2.5 exposure caused by thermal inversions in county i on day t; $\mathbf{X}_{it}$ is a vector of covariates, including the number of employees in the construction industry in the county that month, mean temperature and mean precipitation in that county that day, and day-of-week dummies; φ_i is a county fixed effect; and ε_{it} is an error term, clustered at the county level.[16]

The first stage of my model is:

$$D_{it} = \tau + \rho Z_{it} + \delta' \mathbf{X}_{it} + \theta_i + v_{it} \tag{1.2}$$

[15]This measure equals 1 if a county had high inversion coverage at night (defined as >0.75 average inversion coverage over the midnight, 3 a.m., and 6 a.m. observations) and low inversion coverage during the day (defined as <0.25 inversion coverage over the remaining observations: 9 a.m., noon, 3 p.m., 6 p.m., and 9 p.m.), and 0 otherwise.

[16]Table 1.2, column 1, also gives estimation results when standard errors are clustered by county and date.

where D_{it} is PM 2.5 exposure; Z_{it} is the proportion of the county-day experiencing an inversion; θ_i is a county level fixed effect; and ν_{it} is an error term, again clustered at the county level.

Because I am analyzing the effect of PM 2.5 on a probability (that of a serious worksite accident occurring), I also estimate an instrumental variables probit model with the second stage:

$$Y_{it} = \begin{cases} 1 & \mu + \alpha \hat{D}_{it} + \beta' \mathbf{X}_{it} + \varphi_i + \varepsilon_{it} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{1.3}$$

where all variables are as defined above, except that ε_{it} is assumed to be standard normally distributed. The first stage of this model is precisely as given in equation 1.2.

There could be heterogeneity in the responsiveness of worksite accidents to PM 2.5 across baseline exposure levels, but it is not necessarily clear what form that heterogeneity might take. I analyze such potential heterogeneity by assuming that the response function can be represented by a high degree polynomial, then allowing the data to dictate which terms of this polynomial are actually included in the model. I do this via the Least Absolute Shrinkage and Selection Operator (LASSO), a machine learning regularization technique designed to prevent overfitting of a model to data. LASSO does so by incorporating a regularization term which penalizes the absolute value of the sum of the coefficients in the model. For example, for a simple regression equation $y = \beta' \mathbf{x} + \varepsilon$, OLS coefficients β_{OLS} are given by $\beta_{OLS} = \text{argmin}_{\beta}\ \frac{1}{2N}(y - \beta' \mathbf{x})'(y - \beta' \mathbf{x})$. LASSO coefficients are given by $\beta_{LASSO} = \text{argmin}_{\beta}\ \frac{1}{2N}(y - \beta' \mathbf{x})'(y - \beta' \mathbf{x}) + \lambda \sum_{j=i}^{p} |b_j|$, for some λ. The addition of this regularization term pushes the coefficients β_j toward zero, and precisely to zero in some cases. Variables x_j with non-zero estimated coefficients are those which have been selected as part of the model.

In this case, I explore a polynomial of up to 5 degrees in PM 2.5 by using the second stage equation:

$$Y_{it} = \mu + \sum_{k=1}^{5} \alpha_k \hat{D}_{it}^{k} + \beta' \mathbf{X}_{it} + \varphi_i + \varepsilon_{it} \tag{1.4}$$

where LASSO adds the term $\lambda \sum_{k=1}^{5} |\alpha_k|$ to the objective function to be minimized. This pushes the coefficients α_k toward zero, and precisely to zero if possible, thus selecting only the most important terms $\hat{D}_{it}^{k}$ to be part of the final model. The first stage equations are:

$$D_{it}^{k} = \tau_k + \sum_{l=1}^{5} \rho_{kl} Z_{it}^{l} + \delta_k' \mathbf{X}_{it} + \theta_{ki} + \nu_{kit} \qquad i \in \{1,\ \ldots,\ 5\} \tag{1.5}$$

with terms as defined above. In both stages, clustering is at the county level.

I also use LASSO to explore polynomial non-linearities in the first stage, by using the second stage equation given in 1.1, with the first stage equation:

$$D_{it} = \tau + \sum_{l=1}^{5} \rho_l Z_{it}^l + \delta' \mathbf{X}_{it} + \theta_i + \nu_{it} \tag{1.6}$$

where LASSO adds $\lambda \sum_{l=1}^{5} |\rho_l|$ to the objective function, selecting only the most important powers of Z_{it} to contribute to the final model. Clustering is again at the county level.

1.4 Results

I first present results from my main specification, using thermal inversions as an instrument to estimate the causal effect of PM 2.5 exposure on the probability that a serious construction worksite accident occurs. I then present results from alternative model specifications and robustness checks.

1.4.1 Main Results

Table 1.2, column 1 gives results for my preferred linear panel fixed effects specification. Panel 1 shows first stage results, with an F-statistic of 1272.58, suggesting that thermal inversion are a strong instrument for PM 2.5 exposure.[17]

Panel 2 gives second stage estimates, which indicate that decreasing PM 2.5 exposure by 1 μg/m^3 would decrease the probability of a construction worksite accident occurring by about 0.007 percentage points; this represent roughly 7% of the baseline probability, for an elasticity of 0.67. Over the time frame of this study, this would have resulted in about 80 fewer fatal and catastrophic construction worksite accidents per year. If a similar percentage decrease were to hold across all industries and types of accidents, then BLS (2020a) and BLS (2020b) suggest that this decrease in PM 2.5 exposure would result in about 370 fewer fatal and 244,000 fewer non-fatal workplace injuries annually, saving about $19 billion per year.

These are large effects. For comparison, Sager (2019) estimates an elasticity of 0.06 for traffic accidents with respect to PM 2.5.[18] This is an important, newly quantified, benefit of PM 2.5 abatement that should be accounted for when doing cost-benefit analyses for PM 2.5 regulations and standards.

By way of illustration, the regulatory impact analysis for the 2012 revisions to the National Ambient Air Quality Standards (NAAQS) predicts that lowering the NAAQS standard for 3-year rolling PM 2.5 mean concentration to 11 μg/m^3 from 15 μg/m^3 would result in a population weighted decrease in PM 2.5 of 0.207 μg/m^3, with quantifiable health benefits of $13.9–$33.6 billion, updated to 2019 dollars (EPA 2012). If my results were to hold economy-wide, a 0.2 μg/m^{3}decrease in nationwide mean PM 2.5 would result in about 75 fewer fatal and 48,950 fewer non-fatal workplace injuries annually, saving about $3.8 billion per year. Taking

[17] The lowest first-stage F-statistic across all specifications is 120.80.

[18] Deryugina et al. (2019) estimate an elasticity of 0.019 for elderly three-day mortality rates with respect to PM 2.5; Ebenstein et al. (2016) estimate an elasticity of 0.034 for exam scores with respect to PM 2.5.

Table 1.2: Linear Panel Fixed Effects IV Estimates of the Effect of PM 2.5 on Accident Probability

	County & Weekday FEs (1)	County, Weekday, & Month FEs (2)	County, Weekday, & Season FEs (3)
Panel 1: First Stage			
Inversion Coverage	0.830***	0.254***	0.336***
	(35.67) [10.99]	(11.12)	(14.72)
Temperature	0.120***	0.146***	0.128***
	(64.83) [26.44]	(66.96)	(72.52)
Precipitation	−0.131***	−0.133***	−0.133***
	(−178.44) [−53.60]	(−172.73)	(−178.30)
Employment	0.0000881***	0.0000978***	0.0000924***
	(5.09) [3.93]	(3.99)	(4.00)
F-statistic	1272.58 [120.80]	123.61	216.73
Panel 2: Second Stage			
PM 2.5 (μg/m^3)	0.0000760**	0.000110	0.000105
	(2.56) [2.37]	(1.08)	(1.4)
Temperature	−0.00000416	0.0000153	−0.0000137
	(−1.06) [−0.99]	(−1.16)	(−1.36)
Precipitation	−0.00000461	0.0000134	0.000000345
	(−1.17) [−1.08]	(0.03)	(−0.03)
Employment	0.000000901***	0.000000897***	0.000000897***
	(5.09) [5.05]	(5.04)	(5.06)
County FEs	*X*	*X*	*X*
Weekday FEs	*X*	*X*	*X*
Calendar Month FEs		*X*	
Season FEs			*X*
Clustering	county [county & date]	county	county
Observations	14710686	14710686	14710686

t statistics corresponding to standard errors clustered by county are in parentheses. *t* statistics corresponding to standard errors clustered by county and date are in square brackets.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table reports IV estimates of equation 1.1 (with first stage given by equation 1.2) from the main text. Each observation represents a county-day. The second stage dependent variable in all cases is a dummy for the occurrence of a serious construction worksite accident investigated by OSHA. The proportion of the county-day covered by a thermal inversion is used as an instrument for PM 2.5 exposure, which is measured in μg/m^3. All regressions include county and day-of-week fixed effects, with columns 2 and 3 adding calendar month or season fixed effects, respectively. Standard errors are clustered at the county level in all regressions; column 1 additionally reports *t*- and *F*-statistics corresponding to two-way clustering of standard errors at the county and date levels, in square brackets. Clustering at the county level allows for arbitrary serial correlation of standard errors. Clustering at the date level allows for arbitrary contemporaneous geographic correlation of standard errors.

these benefits into account would increase the estimated benefits of PM 2.5 reduction by at least 11.3–27.5%, a significant increase.

More than this, these results suggest a pathway by which firms receive positive externalities from PM 2.5 abatement. Given that the effect of pollution regulations on firms themselves is often perceived as being at least weakly negative, it is important to understand the ways in which these regulations may actually help firms: in this case, by reducing the probability that they will incur the costs associated with an accident occurring at their establishment. These costs may include worker's compensation payments, OSHA fines, hiring and training replacement workers, etc. Firms that understand the positive externalities they can receive from PM 2.5 abatement should be less inclined to oppose new PM 2.5 regulations and standards.

I now report the results of alternative model specifications: two linear models with temporal fixed effects as well as county fixed effects, a simple instrumental variables probit model, a model using LASSO to explore polynomial non-linearities in the response function of accident probability to PM 2.5, and a model using LASSO to explore polynomial non-linearities in the first-stage responsiveness of PM 2.5 to thermal inversions. The results of robustness checks follow, in which I estimate the primary model using only fatal accidents and in which I estimate the primary model using as my instrument a dummy variable for nights with high inversion coverage followed by days with low inversion coverage.

1.4.2 Alternative Specifications

Table 1.2 gives results from including calendar month (column 2) or season (column 3) fixed effects in the model described in equations 1.1 and 1.2 (with standard errors clustered at the county level). Although these estimates are imprecise, they are in the same direction as my main results, and of similar magnitude. I believe the lack of precision arises because a substantial share of variation in thermal inversions, and hence PM 2.5 exposure caused by thermal inversions, occurs across seasonal boundaries and therefore does not contribute to the estimation when month or season fixed effects are included.

I also estimate an instrumental variables probit model, using the same set of variables as in the linear probability model described above. Results from this specification are given in table 1.3, with first stage results in column 1, and probit coefficients in column 2. Column 3 contains estimated average marginal effects at the means of all variables. These estimated effects are much larger than those from my primary specification, though less precise: they are significant only at the 10% level.

Table 1.4 gives results from the two stage LASSO process described in section 1.3.3, equations 1.4 and 1.5.[19] LASSO selects the fourth and fifth powers of PM 2.5, as shown. The coefficients assigned to these

[19] I use the package *lassopack* by Ahrens et al. (2020), implementing the square-root LASSO method of Belloni, Chernozhukov, et al. (2011) and Belloni, Chernozhukov, et al. (2014)

Table 1.3: IV Probit Estimates of the Effect of PM 2.5 on Accident Probability

	First Stage	Probit Coefficients	Avg. Marginal Effects
PM 2.5		0.0882^{***}	0.00143^{*}
		(6.29)	(1.70)
Inversion Coverage	0.830^{***}		
	(22.66)		
Temperature	0.141^{***}	-0.00887^{***}	-0.000144
	(57.47)	(-3.83)	(-1.46)
Precipitation	-0.105^{***}	0.000242	0.00000394
	(-81.49)	(0.14)	(0.14)
Employment	0.00000530	0.0000156^{***}	0.000000253^{**}
	(0.67)	(5.99)	(2.69)
Constant	-31.597^{***}	-1.622^{**}	
	(-44.74)	(-2.27)	
Observations	14710686	14710686	14710686

z statistics in parentheses. County-level fixed effects included.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table reports maximum likelihood estimates of the IV probit model described in equations 1.2 and 1.3 of the main text. Each observation represents a county-day. The second stage dependent variable is a dummy for the occurrence of a serious construction worksite accident investigated by OSHA. The proportion of the county-day covered by a thermal inversion is used as an instrument for PM 2.5 exposure, which is measured in $\mu g/m^3$. First-stage coefficients are in column 1. Second-stage probit coefficients are in column 2; since interpretation of probit coefficients is not straightforward, estimated average marginal effects are given in column 3. County level fixed effects are included, and clustering is at the county level.

two powers of PM 2.5 imply a significantly smaller effect of PM 2.5 near the (8.95 $\mu g/m^3$) mean level of PM 2.5 observed in my sample, but one which grows more rapidly with increasing PM 2.5, such that the two effects are equal at 26.7 $\mu g/m^3$of PM 2.5. This corresponds to the 98.8th percentile of PM 2.5 exposure in my sample. In other areas of the world, especially in India and China, such values are very common, even below average,[20] and PM 2.5 may be responsible for much larger increases in risk in those areas.

Results from using LASSO to analyze non-linearities in the first stage responsiveness of PM 2.5 to thermal inversions, as described in equation 1.6, are found in table 1.5, column 1.[21] As shown in panel 1, LASSO selects the first and fifth powers of inversion coverage and assign positive coefficients to each, suggesting a convex relationship between PM 2.5 and inversion coverage. Panel 2 gives second stage results, which are essentially the same as those in the linear model. In column 2, I present results from a model identical to my main specification, save for the addition of a quadratic term to equation 1.2, allowing for a convex relationship between PM 2.5 and inversion coverage, but one with a simpler form than the fifth degree polynomial selected by LASSO. The second stage results in panel 2 are essentially the same as in the LASSO model.

[20]See Y. Chen et al. (2020) and Tiwari et al. (2013)

[21]I use the package *pdslasso* by Ahrens et al. (2018), implementing Belloni, D. Chen, et al. (2012).

Table 1.4: LASSO and Polynomial Non-Linearity

	First Stage F-stat	LASSO Coefficients
PM 2.5	5629.64	
PM 2.5^2	1945.13	
PM 2.5^3	845.16	
PM 2.5^4	438.83	0.00000000352
PM 2.5^5	168.11	0.0000000000176
Temperature (K)		
Precipitation		-0.0000021
Employment		0.0000002
Observations	14710686	14710686

t statistics not given. County-level fixed effects included.
Notes: Table reports estimates of the two stage LASSO model described in equations 1.4 and 1.5 of the main text, using the square-root LASSO method of Belloni, Chernozhukov, et al. (2011) and Belloni, Chernozhukov, et al. (2014). Five first-stage equations (for the first through fifth powers of PM 2.5), with five powers of inversion coverage as instruments, were estimated using OLS; their *F*-statistics are reported in column 1. Then, predicted values of the five powers of PM 2.5 were used as inputs for LASSO to estimate equation 1.4. Estimated coefficients from LASSO are given in column 2. The package (*rlasso*) used for this estimation does not provide standard errors for LASSO estimation; in any case, such standard errors would be wrong given the manual nature of this two-stage estimation. Each observation represents a county-day, and the second stage dependent variable is a dummy for the occurrence of a serious construction worksite accident investigated by OSHA. County and day-of-week fixed effects are included.

1.4.3 Robustness Checks

I present the results of two robustness checks in table 1.6. First, I limit the sample of accidents from the OSHA dataset to those in which at least one worker was killed. Fatality is well defined, whereas around half of the incidents in the OSHA dataset do not conform to the definition of "catastrophe" in OSHA (2005), so that it is not completely clear what is being counted. By limiting the dataset to fatal accidents, it is possible to define more precisely what is being counted: fatal accidents that trigger an OSHA investigation. Results from this regression are given in column 1. The estimated effect of PM 2.5 is smaller than but comparable to that in my main specification, though not precisely estimated.

One concern regarding the use of thermal inversions as an instrument for PM 2.5 exposure is that inversions may lead to fog or smog which may be directly observable, thus potentially leading people to modify their behavior, or even reducing visibility, thus increasing worksite risk separately from PM 2.5. To address this, I create a dummy variable that equals one if and only if a night with high inversion coverage ($>$75%) is followed by a day with low inversion coverage ($<$25%), and then use this dummy variable as an instrument for PM 2.5 exposure. Results are given in column 2 of table 1.6. Panel 1 shows the first stage, suggesting that this dummy is a strong instrument for PM 2.5 exposure, though less so that the inversion coverage measure used in my main specification. Panel 2 gives the reduced form coefficients, which are very similar to those from my main specification, though again not precisely estimated.

Table 1.5: Polynomial Non-Linearity in the Responsiveness of PM 2.5 to Thermal Inversions

	LASSO-Estimated Model	Quadratic First Stage IV w/o LASSO
Panel 1: First Stage		
Inversion Coverage	0.591***	0.266***
	(21.87)	(5.19)
Inversion Coverage^2		0.637***
		(10.65)
Inversion Coverage^3		
Inversion Coverage^4		
Inversion Coverage^5	0.357***	
	(11.06)	
Temperature	0.120***	0.120***
	(64.98)	(65.20)
Precipitation	−0.131***	−0.131***
	(−178.66)	(−178.57)
Employment	0.000088***	0.000088***
	(4.00)	(4.00)
Panel 2: Second Stage		
PM 2.5	0.0000619**	0.0000655**
	(2.21)	(2.31)
Temperature	−0.00000246	−0.00000289
	(−0.66)	(−0.77)
Precipitation	−0.00000647*	−0.00000599
	(−1.69)	(−1.56)
Employment	0.000000902***	0.000000901***
	(5.09)	(5.09)
Observations	14710686	14710686

z (column 1) and *t* (column 2) statistics in parentheses. County-level fixed effects included.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table reports estimates of the LASSO model described in equations 1.1 and 1.6 of the main text. LASSO selects which of the first 5 powers of the inversion coverage instrument are best at predicting PM 2.5, and sets the coefficients of all other powers to 0, effectively removing them from the model. Results from this estimation are in column 1. Column 2 contains results from a standard panel IV fixed effects regression with inversion coverage and its square as instruments for PM 2.5, as a simpler functional form that allows for convex non-linearities roughly similar to those found by LASSO. Each observation represents a county-day, and the second stage dependent variable is a dummy for the occurrence of a serious construction worksite accident investigated by OSHA. The proportion of the county-day covered by a thermal inversion (and powers of this proportion) are used as instruments for PM 2.5 exposure, which is measured in $\mu g/m^3$. County level fixed effects are included.

Table 1.6: Robustness Checks

	Fatal Accidents as Dependent Variable (1)	Nighttime-Only Inversions as Instrument (2)
Panel 1: First Stage		
Inversion Instrument	0.830^{***}	0.194^{***}
(see note)	(35.67)	(14.33)
Temperature	0.120^{***}	0.120^{***}
	(64.83)	(64.13)
Precipitation	-0.131^{***}	-0.131^{***}
	(-178.44)	(-178.34)
Employment	0.0000881^{***}	0.0000865^{***}
	(4.00)	(4.98)
F-statistic	1272.58	205.40
Panel 2: Second Stage		
PM 2.5 (μg/m^3)	0.0000277	0.0000713
	(1.22)	(0.27)
Temperature	0.00000229	-0.00000359
	(0.82)	(-0.11)
Precipitation	-0.00000505	-0.00000523
	(-1.64)	(-0.15)
Employment	0.0000003^{***}	0.0000009^{***}
	(7.46)	(4.98)
Observations	14710686	14710686

t statistics in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table reports IV estimates of the robustness checks described in section 1.4.3. Each observation represents a county-day. The second stage dependent variable in column 1 is a dummy for the occureence of a fatal construction worksite accident investigated by OSHA, while that in column 2 is a dummy for the occurrence of a serious construction worksite accident investigated by OSHA (as in my main specification). The proportion of the county-day covered by a thermal inversion is used as an instrument for PM 2.5 exposure in column 1 (as in my main specification), while the instrument used in column 2 is a dummy for days where high nighttime inversion coverage (>75%) is followed by low (<25%) daytime inversion coverage. Both regressions include county and day-of-week fixed effects. Standard errors are clustered at the county level.

In summary, my results suggest that PM 2.5 exposure significantly increases the risk of construction worksite accidents, and that the rate of increase is also increasing in PM 2.5 levels. A variety of alternative specifications and robustness checks provide similar point estimates, less precisely estimated.

1.5 Conclusion

This paper makes important contributions to the literature on worker safety, on pollution and health (particularly cognitive health), and on pollution and worker productivity. I estimate that, at the U.S. average for PM 2.5 concentration, a reduction in PM 2.5 concentration of 1 $\mu g/m^3$could reduce the probability of a dangerous construction worksite accident by 0.007 percentage points, or roughly 7% of baseline risk, an elasticity of 0.67. This is an important benefit of regulations aimed at decreasing PM 2.5 that has not formerly been considered.

As this the first paper to seriously explore the connection between PM 2.5 and workplace safety, there is a great deal of room for more research in this area. Now that the effect of PM 2.5 on construction worksite accidents has been estimated, it is important to understand how government regulation and enforcement can help mitigate these effects. The OSHA dataset I collected also includes information about the unionization of establishments that have accidents, so the effects of unionization on the relationship between pollution and workplace safety can also be explored. The effect on construction worksite safety of other pollutants that have shown cognitive health effects can also be estimated, as well as the effect of both PM 2.5 and other pollutants on workplace safety in other industries.

CHAPTER 2

Fine Particulate Air Pollution and Manufacturing Worker Safety

This paper provides the first causal estimates of the impact of fine particulate matter air pollution (PM 2.5) on manufacturing workplace accidents. I construct a novel dataset composed of the universe of serious workplace accidents investigated by OSHA from 2003 to 2015 and exploit plausibly exogenous variation in PM 2.5 caused by changes in the height of the atmospheric planetary boundary layer height and the shape of the temperature-altitude curve in the lower atmosphere to implement an instrumental variables research design. Focusing on the U.S. manufacturing sector, which employed 12.7 million workers (8.5% of the U.S. workforce) and accounted for 12.1% of nonfatal and 6.5% of fatal U.S. workplace injuries in 2018, I find that decreasing PM 2.5 exposure by 1 μg/m^3 could lead to a 2.8% decrease in the risk of a serious construction worksite accident, representing an elasticity of 0.25.

2.1 Introduction

In recent years, an increased understanding of the harms, both physiological and cognitive,[1] imposed by fine particulate matter air pollution (PM 2.5) has led to increased national and international attention, and regulations intended to curb PM 2.5 emissions. These measures have met with some success, but billions of people around the world are still exposed to high levels of PM 2.5 on a daily basis. Additionally, the scope of the effects that PM 2.5 has on individuals and the economy, as well as the pathways by which these effects occur, are not fully understood. Identifying these effects and quantifying their costs is key to determining optimal pollution control policies.[2]

In this paper, I study an under-explored effect of short-run PM 2.5 exposure: decreased workplace safety, specifically in the U.S. manufacturing sector.[3] According to the 2018 Census of Fatal Occupational Injuries (BLS 2019a) and Survey of Occupational Injuries and Illnesses (BLS 2019b), 5,333 fatal and 3,496,700 non-fatal workplace accidents, costing about $273.3 billion,[4] occurred that year. This paper specifically examines the effect of PM 2.5 on manufacturing accidents, which accounted for 12.1% of nonfatal workplace injuries

[1] An extensive literature documents the negative effects of PM 2.5 on physical health outcomes, including infant mortality and adult mortality due to respiratory disease, lung cancer, and cardiopulmonary disease (for example, see Chowdhury and Dey 2016; Deryugina et al. 2019; Dominici et al. 2006; Pope III et al. 2019; Woodruff et al. 2006). More recently, links between PM 2.5 and mental health and cognitive function have begun to surface. Even in the very short term, exposure to PM 2.5 has been found to affect cognitive ability (Ebenstein et al. 2016), and criminal behavior and anxiety (Lu et al. 2018). These cognitive/behavioral impacts, in turn, can lead to perhaps unexpected consequences, such as an increase in violent crime (Burkhardt et al. 2019) or in car accidents (Sager 2019).

[2] Aldy et al. (2020) document the influence of so-called “co-benefits,” benefits other than the intended target of a regulation, on regulatory decision making. They focus specifically on the EPA, and develop a conceptual framework to demonstrate that any and all benefits (or costs) resulting from a proposed regulation should be accounted for equally as part of its benefits (or costs) rather than being relegated to “co-benefit” or “co-cost” status. In the case of proposed PM 2.5 regulations, this implies that all pathways by which PM 2.5 harms society (pathways by which abatement benefits society) should be identified and treated equally as benefits.

[3] In Chambers (2021b) I find that short-term PM 2.5 exposure leads to an increased risk of serious accidents in the U.S. construction sector. This paper’s results are important on their own, and also provide some evidence for the external validity of my earlier results.

[4] In 2019 dollars. This calculation is based on a value of a statistical life (VSL) of $10.6 million, from Viscusi (2018) (updated to 2019 dollars), and a value of a statistical injury (VSI) of $62,000, from Viscusi and Aldy (2003) (updated to 2019 dollars).

and illnesses and 6.5% of fatal workplace injuries in 2018 (BLS 2019a; BLS 2019b).

I utilize a unique dataset constructed by scraping OSHA's online database of fatality and catastrophe investigations; data from the Quarterly Census of Employment and Wages (QCEW) and North American Regional Reanalysis (NARR); and a new, high resolution, daily PM 2.5 prediction dataset produced by Di et al. (2019). From these sources, I construct a daily panel covering all counties in the contiguous United States from January 1, 2003 to December 31, 2015.

Because PM 2.5 is largely generated by human activities, it is likely co-determined with other activities that impact workplace safety. Especially in the manufacturing sector, it may even be the case that accidents influence the production of PM 2.5; for instance, a serious accident might lead to a plant shutting down briefly and producing less PM 2.5. To address this concern and identify the causal effects of PM 2.5, I need a plausibly exogenous source of variation in PM 2.5 to use in implementing an instrumental variables research design.

I use a suite of weather measurements capturing the capacity of the atmosphere to disperse pollutants generated near the surface upwards into the atmosphere as this source of exogenous variation in PM 2.5 exposure. The first of these measurements is the height of the *planetary boundary layer*, the portion of the atmosphere which is strongly influenced by the earth's surface due to vertical mixing of air. Second, I use a group of temperature-based measurements that capture the *convective stability*, or tendency to circulate vertically, of the air at different levels above the surface. When the planetary boundary layer is low, or convective stability is high, pollutants are not dispersed as far upwards as they would otherwise be, concentrating such pollutants near the surface where they are generated. The identifying assumption of my study is that, conditional on surface weather (temperature, precipitation, humidity, and wind speed), atmospheric characteristics above the surface should not affect workers at the surface except by their influence on air pollution.

I find that a significant portion of serious manufacturing workplace accidents may be caused by short-run exposure to PM 2.5. Estimates from my primary specification, a linear panel fixed effect IV model, suggest that reducing manufacturing workers' PM 2.5 exposure by 1 μg/m^3 could reduce the probability of a serious manufacturing workplace accident by 2.8%, (representing an elasticity of 0.25), and save $845 million per year.

These results indicate that the reduced risk of serious workplace accidents is an important benefit that should be accounted for when evaluating proposed PM 2.5 regulations, or other regulations that will impact PM 2.5. These results also suggest that such regulations may more effectively reduce workplace accident risk

[4] Meteorological conditions are a common class of instruments for PM 2.5 exposure. Sager (2019) uses thermal inversions, a phenomenon closely related to the measures of convective stability I use. Deryugina et al. (2019) and Heyes and Zhu (2019) use wind direction. Non-meteorological instruments include advanced and inflexible scheduling of events (see Archsmith et al. 2018; Ebenstein et al. 2016).

than regulations directly targeting workplace safety do.[5]

Additionally, these results suggest that PM 2.5 has a negative impact on firm profitability (through these worker safety effects). Consequently, firms will experience benefits from other firms' adherence to PM 2.5 regulations. Whether the benefits experienced by the individual firm offset their compliance costs will of course vary by industry, firm, and specific regulation, but these results suggest that firms may currently be overestimating their optimal level of opposition to such regulations, if they are not currently taking into account the positive externality they receive from other firms' compliance.

2.2 Background

2.2.1 Fine Particulate Matter (PM 2.5) and Health

Fine particulate matter (PM 2.5) air pollution consists of particles less than 2.5 micrometers (µm) in diameter; for comparison, the average human hair is 50-70 µm in diameter. These very small particles arise from a variety of sources. Particles emitted directly into the atmosphere are referred to as "primary" particles. "Secondary" particles form in the atmosphere from chemicals emitted into the atmosphere as gases, and may form some distance from where the gases were originally emitted. These secondary particles constitute a large proportion of PM 2.5, as opposed to coarser particulate matter, which contains more primary particles. Major contributors to the formation of PM 2.5 secondary particles are sulfur dioxide (SO2), which is emitted by the combustion of fossil fuels and the smelting of metal ores containing sulfur, and nitrogen oxides (NOx), which are emitted by high temperature combustion, such as that found in vehicle engines and power plants. These and other anthropogenic emissions account for a large portion of ambient PM 2.5.

Because of the small size of PM 2.5 particles, they are absorbed into the bloodstream when inhaled, rather than being filtered out by the lungs. They are then transmitted throughout the body. An extensive literature documents the negative physiological health effects of chronic exposure to PM 2.5, including increased likelihood of death due to heart disease, heart attack, and lung cancer (see Achilleos et al. 2017). Largely due to these negative physiological health effects, both coarse particulate matter (PM 10) and fine particulate matter (PM 2.5) air pollution are listed as criteria pollutants under the National Ambient Air Quality Standards (NAAQS) established by the EPA.[6]

Some of the particles making up PM 2.5 are small enough to cross from the bloodstream into the brain, where they can impact cognitive and mental health. The effects of PM 2.5 on cognition and the brain have

[5] See Kniesner and Leeth (2013) for an overview of several studies finding little to no effect of OSHA regulations and enforcement on workplace accident rates. Some of the more optimistic studies find that OSHA inspections with penalties for violations have a stronger preventive effect on workplace accidents than those without, and that citations for failure to properly use personal protective gear have the strongest effect (A. Haviland et al. 2010; Mendeloff and Gray 2005). A. M. Haviland et al. (2012) suggest that inspections with penalties tend to reduce injuries for up to 2 years in midsize establishments (between 20 and 250 employees).

[6] The current standard for PM 2.5 is 12 µg/m^3 (3-year average of annual mean concentration), significantly above the national mean PM 2.5 concentration (as of 2019) of 7.65 µg/m^3.

been well documented in the last few years. Gatto et al. (2014) and Ailshire and Crimmins (2014) show that long term exposure to PM 2.5 is connected to decreased cognitive functioning in middle aged and older adults in the United States.[7] Weuve et al. (2012) find that chronic PM 2.5 exposure leads to an increased rate of decline in cognitive ability with age. Fonken et al. (2011), in a medical study using mice, find that mice chronically exposed to elevated levels of PM 2.5 displayed depressive behavioral symptoms, in addition to decreased cognitive ability, specifically impairments in spatial learning and memory. They describe the specific brain regions that were found to be affected, including the hippocampus, one of the centers of both spatial learning and memory formation. Block and Calderón-Garcidueñas (2009) and Costa et al. (2014) suggest that PM 2.5 causes brain inflammation and oxidative stress, leading to various neuropathologies and central nervous system diseases.

The negative physical and cognitive/mental health effects of PM 2.5 are not only seen in those who experience chronic PM 2.5 exposure. Acute (short term) PM 2.5 exposure has also been shown to impact both physical and cognitive/mental health. Deryugina et al. (2019) use a large-scale database of Medicare recipients, and find that acute PM 2.5 exposure causes increased mortality, emergency hospitalization, and medical spending for up to two days following exposure. Ebenstein et al. (2016) find that acute exposure to elevated PM 2.5 levels causes an immediate decrease in cognitive ability. They use evidence from standardized testing of pre-college teens in Israel to show that elevated PM 2.5 at the testing location on test day causes lower scores, with subsequent persistent effects on educational attainment and earnings. Heyes and Zhu (2019) find that elevated levels of PM 2.5 lead to significantly increased sleeplessness, another possible pathway for physiological and cognitive harm.

Likely due to its impacts on cognition and behavior, short-term exposure to PM 2.5 has been shown by Burkhardt et al. (2019) to lead to an increase in violent crime. Interestingly, they find that lagged daily PM 2.5 has no impact on crime, suggesting that there are immediate, non-cumulative behavioral and cognitive effects of PM 2.5, although we have seen that chronic exposure has its own effects. Sager (2019) shows that short-term exposure to PM 2.5 causes an increase in vehicular accidents, and similarly to Burkhardt, et al., finds that lagged daily PM 2.5 has no impact.[8]

2.2.2 Worker Productivity and Safety

Fine particulate air pollution (PM 2.5) has been shown to directly impact the productivity of workers engaged in physical labor or mental labor, indoors or out. T. Chang et al. (2016) study pear packers engaged in manual

[7] The Gatto, et al. study is confined to Los Angeles.

[8] This also suggests that the sleeplessness pathway proposed by Heyes and Zhu is not the only pathway by which PM 2.5 affects cognition and behavior.

labor inside a factory. They find that elevated levels of PM 2.5, which readily moves indoors,[9] causes a decrease in worker productivity, while other pollutants, such as ozone, which do not readily move indoors, do not have an effect on indoor workers. They find that labor supply decisions appear to be unaffected by PM 2.5 levels, which they take as evidence that workers are not self-selecting their PM 2.5 exposure. Adhvaryu et al. (2014) find similar results among workers in a garment factory in India; additionally, they find substantial heterogeneity in treatment effects based on task and manager identity, suggesting that pollution impacts to productivity can be somewhat mitigated if managers choose to do so. Archsmith et al. (2018) use advanced scheduling of MLB baseball games as an instrument for umpire's PM 2.5 exposure and find that short-term PM 2.5 exposure leads to umpires, who work (mostly) outdoors, making more incorrect calls. T. Y. Chang et al. (2019) find that call center workers, who are indoors and performing almost entirely mental labor, are also impacted by PM 2.5 levels. It seems plausible that safe work practices could be affected along with productivity, whether via cognitive or other pathways.

OSHA, the U.S. government agency tasked with promulgating and enforcing safety rules and regulations to reduce the frequency of workplace accidents, regulates workplace exposure to a variety of toxic chemicals, and to dust in general, but does not regulate fine particulate matter. Similarly, literature on the connection between workplace safety and pollution has mainly focused on the health effects of pollution (especially toxic chemicals) generated by the workplace (Landrigan 1992) and the existence of compensating wages for workers in highly polluting industries (Cole et al. 2009), while ambient particulate air pollution as a potential cause of workplace accidents has not been studied (except in Chambers (2021b)).

2.3 Research Design

2.3.1 Data

To analyze the effect of acute PM 2.5 exposure on worksite accidents, I require data on workplace accidents that includes, at minimum, the date of each accident, the county it occurred in, and the employer's industrial classification. I make use of a unique dataset of OSHA Fatality and Catastrophe Investigation summaries, which provide the address where each incident occurres, the date it occurred, and the NAICS-6 industry classification of the employer. Addresses were linked to pollution and weather information using geolocation.[10]

Additionally, the victims in each incident are enumerated and categorized according to whether they were killed, hospitalized, or neither. While a fatality is (relatively) easy to define, catastrophes were defined by OSHA in 2005 as "the hospitalization of three or more employees resulting from a work-related incident or

[9] See Guo et al. (2010), Meier et al. (2015), and Zhao et al. (2015) for studies on the correlation between outdoor ambient PM 2.5 and indoor PM 2.5 exposure.

[10] The address data from the investigation summaries is very clean. The Spatial Analysis Research Laboratory (SARL) at Vanderbilt University geocoded the addresses using ArcGIS and was able to assign a precise location to all but 172 addresses (about .015% of the total).

exposure" (OSHA 2005). However, more than half of the incidents in the Fatality and Catastrophe Investigation database do not fit this definition, for reasons that are unclear. For my main results, I use the set of all accidents at manufacturing establishments reported in the OSHA Fatality and Catastrophe Investigation database, but due to uncertainty surrounding the definition of "catastrophe," and the fact that this definition could have changed over time, I conduct a robustness check using a restricted dataset consisting of fatal accidents only.

Data on employment come from the Quarterly Census of Employment and Wages (QCEW). The QCEW is conducted quarterly, as the name suggests, and covers 95% of employees in the United States,[11] providing information on total employment by industry within each county. These data are drawn from each state's unemployment insurance accounting program, which receives quarterly reports from each employer in the state, with employment and wage data for each of their establishments within the state. These statistics are reported by month within each quarter, so I build a dataset of monthly employment data by county and industry from the QCEW data.

For PM 2.5, I make use of a newly available dataset (Di et al. 2019) that uses a variety of data sources[12] and models of how pollutants are dispersed and transported through the atmosphere. The authors use an ensemble of machine learning algorithms to estimate daily PM 2.5 levels in 1 km by 1 km pixels across the contiguous United States for the years 2000 to 2015. Their final model predicts PM 2.5 levels well, with a 10-fold cross validated[13] R^2 averaging 0.86. Directly monitored PM 2.5 has two major disadvantages for my analysis: first, not all counties have PM 2.5 monitors; second, many existing monitors collect data only every three or every six days. Using the Di et al. (2019) PM 2.5 prediction dataset allows me to analyze the effect of PM 2.5 on workplace accidents across the entire contiguous U.S., and is more accurate than the interpolation methods I would have to use with monitor data to fill in missing days.

Weather data, including planetary boundary layer height and temperature at a variety of levels above the surface (used to calculate the environmental lapse rate) as well as surface temperature, precipitation, wind speed, and humidity (to control for surface weather), come from the North American Regional Reanalysis (NARR) dataset (Mesinger et al. 2006). The NARR uses historical data from a variety of sources[14] to create a detailed picture of weather and climate conditions in 32×32 km square pixels across all of North America.

From these data sources, I construct a daily panel of U.S. counties covering the 48 contiguous U.S. states and the years 2003 to 2015. For each county-day, I include a dummy for whether a manufacturing accident

[11] There are minor exclusions for some government, agricultural, and railroad employees and self-employed persons.

[12] These include satellite images, ground-based monitor data, meteorological data, land-use data, and reanalysis datasets.

[13] k-fold cross validation refers to a method of testing the predictive power of a model by randomly partitioning the data into k equal subsamples. Each of these subsamples, in turn, is used as the validation data for a model trained using the other 9 samples.

[14] These include satellites, radiosondes (balloon mounted instruments), dropsondes (instruments dropped from aircraft), and surface instruments

Table 2.1: Summary Statistics by Season

	Spring	Summer	Fall	Winter	Total
Accident Occurred	0.000691 (0.0263)	0.000694 (0.0263)	0.000627 (0.0250)	0.000636 (0.0252)	0.000662 (0.0257)
PM 2.5 (μg/m^3)	8.514 (4.960)	10.85 (6.348)	8.174 (5.354)	8.260 (5.323)	8.954 (5.631)
PBL Height	1007.1 (451.0)	1011.5 (427.0)	847.8 (393.6)	766.3 (397.7)	909.0 (431.1)
Lapse Rate (Layer 1)	0.00139 (0.00719)	0.00264 (0.00468)	−0.000500 (0.00764)	−0.000976 (0.00974)	0.000650 (0.00766)
Lapse Rate (Layer 2)	0.00547 (0.00342)	0.00650 (0.00212)	0.00492 (0.00361)	0.00253 (0.00614)	0.00487 (0.00433)
Lapse Rate (Layer 3)	0.00623 (0.00330)	0.00751 (0.00179)	0.00585 (0.00332)	0.00261 (0.00526)	0.00556 (0.00405)
Lapse Rate (Layer 3)	0.00592 (0.00331)	0.00730 (0.00173)	0.00550 (0.00325)	0.00246 (0.00469)	0.00531 (0.00383)
Cooling Degree Days	1.189 (2.346)	6.551 (4.075)	1.695 (2.946)	0.0597 (0.441)	2.387 (3.737)
Heating Degree Days	6.402 (6.770)	0.305 (1.216)	5.350 (6.173)	16.40 (8.267)	7.074 (8.496)
Precipitation (mm)	2.791 (6.889)	3.066 (6.684)	2.351 (6.822)	2.135 (5.917)	2.589 (6.602)
Surface Wind Speed	3.912 (1.890)	3.108 (1.454)	3.583 (1.828)	3.816 (1.952)	3.604 (1.818)
Relative Humidity	69.19 (15.55)	67.05 (16.56)	68.43 (15.72)	76.34 (14.54)	70.23 (16.02)
Manufacturing Employment	4168.3 (13302.1)	4187.3 (13292.6)	4165.1 (13217.7)	4161.6 (13264.0)	4170.6 (13269.3)
Observations	3,717,168	3,717,168	3,676,764	3,645,684	14,756,784

occurred, mean PM 2.5, planetary boundary layer height, environmental lapse rate in the bottom four layers of the atmosphere, level of manufacturing employment and total employment, temperature, precipitation, humidity, and surface wind speed. Table 2.1 contains summary statistics.

2.3.2 Methods

As discussed earlier, PM 2.5 and the chemicals which react to produce PM 2.5 in the air are largely human generated, from sources such as power plants, manufacturing, and vehicular traffic. Consequently, PM 2.5 exposure may be co-determined with other activities that contribute to construction worksite risk, leading to endogeneity. To account for this endogeneity, I need an exogenous source of variation in PM 2.5 exposure to use in an instrumental variables research design.

Various sources of exogenous variation in PM 2.5 have been used in other studies, including advanced and inflexible scheduling of events,[15] wind direction,[16] and thermal inversions.[17]

The set of instruments I use is similar to thermal inversions in that both are connected with the vertical dispersal, upwards into the atmosphere, of PM 2.5 generated near the earth's surface. First, I use the height of the *planetary boundary layer*. The planetary boundary layer is defined as the portion of the earth's surface in which wind and other atmospheric characteristics (temperature, pressure, and humidity) are strongly influenced by surface characteristics, due to vertical mixing of the air. The height of this layer, therefore, is a rough measure of the space in which pollutants generated near the surface may be dispersed due to this vertical mixing.

The second weather measurement, or group of measurements, that I use as instruments is a set of measurements of the rate of temperature change with height for each of four layers of air above the earth's surface. This rate of temperature change determines the vertical circulation of air via convection, based on the physical properties of air. Under theoretically ideal conditions, a parcel of air rising through the atmosphere will cool at a rate of 10°C/km, or 0.01°C/m; this rate is known as the *dry adiabatic lapse rate*. When the *environmental lapse rate*, or the rate of cooling with height actually observed in a region of the atmosphere, is less than the dry adiabatic lapse rate, then this region of the atmosphere is *convectively stable*, and vertical circulation due to convection is inhibited. The degree to which convection is inhibited depends on how low

[15] Ebenstein et al. (2016) use the Israeli Bagrut college entrance exam, which is compulsory for all high school students, is scheduled years in advance, and which may not be rescheduled. This exam also takes place across several days, and scores are recorded by day, enabling the authors to use variation in PM 2.5 across different days for each individual student to identify the effect of PM 2.5 on test scores. Archsmith et al. (2018) use umpire schedules for Major League Baseball (MLB) games, which are set months in advance and which cannot readily be changed. These are excellent sources of exogenous variation in PM 2.5, but such ideal sources are not common.

[16] Wind has a strong effect on PM 2.5 exposure in many locations. The chief difficulty in using (lateral) wind direction as an instrument for PM 2.5 exposure is that its effect is completely dependent on location, since a location may have pollution sources distributed arbitrarily around the points of the compass. To account for this, Heyes and Zhu (2019), who confine their study to 19 of China's most populous cities, construct a simple air transport model for each city in their sample. In contrast, Deryugina et al. (2019) use a non-parametric specification which allows the effect of wind direction on PM 2.5 to vary across groups of counties and makes no ex ante assumptions about the effect of wind in any given county.

[17] See Sager (2019).

Figure 2.1: Turbulence generated by wind passing over surface obstructions. The planetary boundary layer by definition lies above such turbulence.
Image credit: Pilot's Handbook of Aeronautical Knowledge, Federal Aviation Administration, page 12-10.

the environmental lapse rate is.

In the NARR (North American Regional Reanalysis) data, temperatures are given at different *pressure levels* of the atmosphere. A pressure level is defined as the point in the air column with a specified atmospheric pressure (e.g. 950 millibars). At any given point on the earth's surface, the height of each pressure level above the surface varies with atmospheric/weather conditions and is included as an additional variable in the NARR data. I calculate the environmental lapse rate between the surface and the lowest pressure level that is at least 100 meters above the surface, then between each consecutive pair of pressure levels for the next three pressure levels (going up). This gives me environmental lapse rates for 4 different layers of air at each point. I then take the average over each county to get the average lapse rate in each of these four layers of air for the county-day.

The environmental lapse rate instrument is related to the thermal inversion instrument used by Sager (2019) and in my earlier paper (Chambers 2021b). A thermal inversion exists when cool air near the surface is overlain by warmer air. The cool surface air therefore does not tend to rise, i.e. vertical circulation due to convection is again inhibited. But in this case, by definition, the environmental lapse rate will be negative (since it measures the rate of cooling with height), definitely less than 0.01°C/m. The environmental lapse rate instrument thus incorporates a more detailed picture of atmospheric convective stability than simple thermal inversions.

These measures of vertical pollution dispersion capacity are simpler instruments for PM 2.5 than wind direction because the sign of their effect on surface PM 2.5 exposure is unambiguous, making it unnecessary to develop an air transport model, like Heyes and Zhu (2019), or utilize non-parametric methods, like Deryugina et al. (2019). This is a primary reason I have chosen this suite of measurements as instruments for PM 2.5 exposure.

Another important factor in vertical dispersion of pollutants generated near the service is surface wind speed. Faster wind generates more turbulence as it passes over irregularities on the surface, and this turbulence contributes to vertical mixing of the air. However, surface wind speed could plausibly have a direct impact on worker safety, so I use it as a control variable and not as an instrument.

2.3.3 Model Specification

For my main results, I use a panel fixed effect IV specification, with the second stage as follows:

$$Y_{it} = \mu + \alpha \hat{D}_{it} + \beta' \mathbf{X}_{it} + \varphi_i + \varepsilon_{it} \tag{2.1}$$

where Y_{it} is a binary variable indicating whether an accident occurred in county i on date t; $\hat{D}_{it}$ represents PM 2.5 exposure caused by planetary boundary layer height and environmental lapse rate in four layers of air near the surface in county i on day t; $\mathbf{X}_{it}$ is a vector of covariates, including the number of employees in the manufacturing industry in the county that month, mean temperature, precipitation, relative humidity, and wind speed in that county that day, and day-of-week and calendar month dummies; φ_i is a county fixed effect; and ε_{it} is an error term, clustered at the county level.

The first stage of my model is:

$$D_{it} = \tau + \rho' \mathbf{Z}_{it} + \delta' \mathbf{X}_{it} + \theta_i + \nu_{it} \tag{2.2}$$

where D_{it} is PM 2.5 exposure; $\mathbf{Z}_{it}$ is a vector of instruments, including the height of the planetary boundary layer and environmental lapse rates for four layers of air above the earth's surface; θ_i is a county level fixed effect; and ν_{it} is an error term, again clustered at the county level.

Because I am analyzing the effect of PM 2.5 on a probability (that of a serious worksite accident occurring), I also estimate an instrumental variables probit model with the second stage:

$$Y_{it} = \begin{cases} 1 & \mu + \alpha \hat{D}_{it} + \beta' \mathbf{X}_{it} + \varphi_i + \varepsilon_{it} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

where all variables are as defined above, except that ε_{it} is assumed to be standard normally distributed. The first stage of this model is precisely as given in equation 2.2.

There could be heterogeneity in the responsiveness of workplace accidents to PM 2.5 across baseline exposure levels, but it is not necessarily clear what form that heterogeneity might take. I analyze such potential heterogeneity by assuming that the response function can be represented by a high degree polynomial, then allowing the data to dictate which terms of this polynomial are actually included in the model. I do this via the Least Absolute Shrinkage and Selection Operator (LASSO), a machine learning regularization technique designed to prevent overfitting of a model to data. LASSO does so by incorporating a regularization term which penalizes the absolute value of the sum of the coefficients in the model. For example, for a simple regression equation $y = \beta'\mathbf{x} + \varepsilon$, OLS coefficients β_{OLS} are given by $\beta_{OLS} = \text{argmin}_\beta \frac{1}{2N}(y - \beta'\mathbf{x})'(y - \beta'\mathbf{x})$. LASSO coefficients are given by $\beta_{LASSO} = \text{argmin}_\beta \frac{1}{2N}(y - \beta'\mathbf{x})'(y - \beta'\mathbf{x}) + \lambda \sum_{j=i}^{p} |b_j|$, for some λ. The addition of this regularization term pushes the coefficients β_j toward zero, and precisely to zero in some cases. Variables x_j with non-zero estimated coefficients are those which have been selected as part of the model.

In this case, I explore a polynomial of up to 5 degrees in PM 2.5 by using the second stage equation:

$$Y_{it} = \mu + \sum_{k=1}^{5} \alpha_k \hat{D}_{it}^k + \beta'\mathbf{X}_{it} + \varphi_i + \varepsilon_{it} \tag{2.4}$$

where LASSO adds the term $\lambda \sum_{k=1}^{5} |\alpha_k|$ to the objective function to be minimized. This pushes the coefficients α_k toward zero, and precisely to zero if possible, thus selecting only the most important terms $\hat{D}_{it}^k$ to be part of the final model. The first stage equations are:

$$D_{it}^k = \tau_k + \sum_{l=1}^{5} \rho_{kl}' \mathbf{Z}_{it}^l + \delta_k' \mathbf{X}_{it} + \theta_{ki} + \nu_{kit} \qquad i \in \{1, \ldots, 5\} \tag{2.5}$$

with terms as defined above. In both stages, clustering is at the county level.

2.4 Results

I first present results from my main specification, using planetary boundary layer height and environmental lapse rate as instruments to estimate the causal effect of PM 2.5 exposure on the probability that a serious manufacturing workplace accident occurs. I then present results from alternative model specifications and robustness checks.

2.4.1 Main Results

Column 2 of tables 2.2 and 2.3 gives results for my preferred linear panel fixed effects specification. Table 2.2 shows first stage results, with an F-statistic of 2415.98, suggesting that planetary boundary height layer and environmental lapse rate in the four pressure layers of air nearest the earth's surface are a strong set of instruments for PM 2.5 exposure.

Table 2.3 gives the second stage estimates, which indicate that decreasing PM 2.5 exposure by 1 μg/m^3 would decrease the probability of a fatal or catastrophic manufacturing accident occurring by about 0.0019 percentage points; this represent roughly 2.8% of the baseline probability, for an elasticity of 0.25. If a

Table 2.2: Linear Panel Fixed Effects IV Estimates of the Effect of PM 2.5 on Accident Probability

First Stage	County & Weekday FEs (1)	County, Weekday, & Month FEs (2)
Lapse Rate (Layer 1)	-83.325^{***}	-82.031^{***}
	(-41.64)	(-42.28)
Lapse Rate (Layer 2)	-28.306^{***}	-22.972^{***}
	(-7.24)	(-5.77)
Lapse Rate (Layer 3)	-111.095^{***}	-107.0972^{***}
	(-31.08)	(-30.15)
Lapse Rate (Layer 4)	-21.427^{***}	-33.178^{***}
	(-7.05)	(-10.77)
PBL Height	-0.000915^{***}	-0.00124^{***}
	(-18.77)	(-23.30)
Cooling Degree Days	0.530^{***}	0.592^{***}
	(109.26)	(92.65)
Heating Degree Days	-0.041^{***}	-0.088^{***}
	(-21.65)	(-58.97)
Precipitation	-0.118^{***}	-0.119^{***}
	(-167.72)	(-169.06)
Surface Wind Speed	-0.280^{***}	-0.265^{***}
	(-47.33)	(-45.61)
Relative Humidity	0.020^{***}	0.017^{***}
	(18.77)	(15.71)
Employment	0.000194^{***}	0.000194^{***}
	(3.28)	(3.27)
County FEs	*X*	*X*
Weekday FEs	*X*	*X*
Calendar Month FEs		*X*
F-statistic	1790.93	2415.98
Observations	14504337	14504337

t statistics in parentheses. Standard errors are clustered at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.3: Linear Panel Fixed Effects IV Estimates of the Effect of PM 2.5 on Accident Probability

Second Stage	County & Weekday FEs (1)	County, Weekday, & Month FEs (2)
PM 2.5 (μg/m^3)	0.0000117	0.0000188**
	(1.53)	(2.39)
Cooling Degree Days	−0.000000814	−0.00000345
	(0.22)	(−0.72)
Heating Degree Days	−0.000000930	0.00000285*
	(0.89)	(1.66)
Precipitation	−0.0000000436	0.000000879
	(0.03)	(0.58)
Surface Wind Speed	−0.00000259	0.00000509
	(0.57)	(1.05)
Relative Humidity	−0.000000215	−0.000000124
	(−0.25)	(−0.15)
Employment	0.000000793**	0.000000792**
	(2.16)	(2.16)
County FEs	*X*	*X*
Weekday FEs	*X*	*X*
Calendar Month FEs		*X*
Observations	14504337	14504337

t statistics in parentheses. Standard errors are clustered at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table reports IV estimates of equation 2.1 (with first stage given by equation 2.2) from the main text. Each observation represents a county-day. The second stage dependent variable in all cases is a dummy for the occurrence of a serious manufacturing accident investigated by OSHA. The proportion of the county-day covered by a thermal inversion is used as an instrument for PM 2.5 exposure, which is measured in μg/m^3. All regressions include county and day-of-week fixed effects, with columns 2 and 3 adding calendar month or season fixed effects, respectively. Standard errors are clustered at the county level in all regressions, allowing for arbitrary serial correlation of standard errors.

similar percentage decrease were to hold across all types of manufacturing accidents, then BLS (2019a) and BLS (2019b) suggest that this decrease in PM 2.5 exposure would result in about 9.6 fewer fatal and 12,000 fewer non-fatal manufacturing workplace injuries annually, saving more than $845 million per year.

These are large effects: smaller than those found in my earlier paper (Chambers 2021b), where I find that the probability of a fatal or catastrophic construction worksite accident has an elasticity of 0.67 with respect to PM 2.5, but still large compared to elasticities from some related studies, such as the elasticity of 0.06 that Sager (2019) finds for traffic accidents with respect to PM 2.5.[18] This paper adds to the evidence from Chambers (2021b) that workplace safety is an important benefit of PM 2.5 abatement.

To repeat the comparative illustration found in Chambers (2021b), the regulatory impact analysis for the 2012 revisions to the National Ambient Air Quality Standards (NAAQS) predicts that lowering the NAAQS standard for 3-year rolling PM 2.5 mean concentration to 11 $\mu g/m^3$ from 15 $\mu g/m^3$ would result in a population weighted decrease in PM 2.5 of 0.207 $\mu g/m^3$, with quantifiable health benefits of $13.9–$33.6 billion, updated to 2019 dollars (EPA 2012). If the results from this paper were to hold economy-wide, a 0.2 $\mu g/m^3$ decrease in nationwide mean PM 2.5 would result in 29.4 fewer fatal and 19,850 fewer non-fatal workplace injuries annually, saving about $1.5 billion per year: an increase of at least 4.5–10.8% over EPA's estimated benefits. These effects are smaller than those estimated for the construction industry in my earlier paper, but suggest that the negative effects of PM 2.5 on workplace safety are not limited to those working outdoors.

These results also add to those in Chambers (2021b) in suggesting that firms may receive direct, bottom-line benefits from PM 2.5 abatement, in the form of reductions to worker's compensation payments, OSHA fines, costs of hiring and training replacement workers, etc., due to reduced risk of workplace accidents. Whether these benefits outweigh firms' compliance costs is another question, one which depends on the specific firm and regulation in question, but the results from both these papers suggest that firms may be over-estimating the net costs associated with PM 2.5 regulations and therefore over-estimating the optimal level of opposition to such regulations.

I now report the results of alternative model specifications: a simple instrumental variables probit model, a model using LASSO to explore polynomial non-linearities in the response function of accident probability to PM 2.5, and a robustness check in which I estimate the primary model using only fatal accidents.

[18] On the other end of the scale, Deryugina et al. (2019) estimate an elasticity of 0.019 for elderly three-day mortality rates with respect to PM 2.5; Ebenstein et al. (2016) estimate an elasticity of 0.034 for exam scores with respect to PM 2.5.

2.4.2 Alternative Specifications

I also estimate an instrumental variables probit model, using the same set of variables as in the linear probability model described above. Results from this specification are given in table 2.4, with first stage results in column 1, and probit coefficients in column 2. Column 3 contains estimated average marginal effects at the means of all variables. In this case the estimated average marginal effects are statistically significant and nearly five times larger than the estimated effects in my primary, linear, specification.

Table 2.5 gives results from the two stage LASSO process described in section 2.3.3, equations 2.4 and 2.5.[19] LASSO does not select any of the powers of PM 2.5, so this model does not give us any information about nonlinearity of the relationship between PM 2.5 and manufacturing accident probability.

2.4.3 Robustness Checks

I present in table 2.6 the results of a robustness check in which I limit the sample of accidents from the OSHA dataset to those in which at least one worker was killed. Fatality is well defined, whereas around half of the incidents in the OSHA dataset do not conform to the definition of "catastrophe" in OSHA (2005), so that it is not completely clear what is being counted. By limiting the dataset to fatal accidents, it is possible to define more precisely what is being counted: fatal accidents that trigger an OSHA investigation. Results from this regression are given in the table. The estimated effect of PM 2.5 is smaller than but comparable to that in my main specification, though not precisely estimated.

2.5 Conclusion

This paper makes important contributions to the literature on worker safety, on pollution and health (particularly cognitive health), and on pollution and worker productivity. I estimate that, at the U.S. average for PM 2.5 concentration, a reduction in PM 2.5 concentration of 1 $\mu g/m^3$could reduce the probability of a fatal or catastrophic manufacturing accident by 0.0018 percentage points, or roughly 2.8% of baseline risk, implying an elasticity of 0.25. This is an important, newly quantified benefit of regulations aimed at decreasing PM 2.5.

As this paper and Chambers (2021b) are the first to estimate the effect of PM 2.5 on workplace safety (in construction and manufacturing, respectively), much research remains to be done. Estimating these effects in other industries is crucial for understanding what the total effect of changes in PM 2.5 is on worker safety. These results also suggest that other airborne pollutants, such as ozone, may be contributing to workplace accidents, and it would be good to investigate these effects. Finally, understanding the specific physiological

[19] I use the package *lassopack* by Ahrens et al. (2020), implementing the square-root LASSO method of Belloni, Chernozhukov, et al. (2011) and Belloni, Chernozhukov, et al. (2014)

Table 2.4: IV Probit Estimates of the Effect of PM 2.5 on Accident Probability

	First Stage	Probit Coefficients	Avg. Marginal Effects
PM 2.5		0.0339^{***}	0.0000888^{***}
		(4.70)	(2.80)
PBL Height	-0.00199^{***}		
	(−36.27)		
Lapse Rate (Layer 1)	-94.849^{***}		
	(−32.10)		
Lapse Rate (Layer 2)	−4.692		
	(−0.62)		
Lapse Rate (Layer 3)	4.810		
	(0.61)		
Lapse Rate (Layer 4)	-92.541^{***}		
	(−16.01)		
Cooling Degree Days	0.537^{***}	-0.0261^{***}	-0.0000683^{***}
	(92.07)	(−4.63)	(−2.86)
Heating Degree Days	-0.0660^{***}	-0.00369^{***}	-0.00000965^{**}
	(−19.21)	(−3.08)	(−2.37)
Precipitation	-0.121^{***}	0.00454^{***}	0.0000119^{***}
	(−131.69)	(4.21)	(2.92)
Surface Wind Speed	-0.179^{***}	0.00684	0.0000179
	(−21.16)	(1.30)	(1.18)
Relative Humidity	0.0628^{***}	-0.00631^{***}	-0.0000165^{***}
	(37.40)	(−3.43)	(−2.67)
Manufacturing Employment	0.0000183^{***}	0.00000652^{***}	0.0000000171^{**}
	(3.03)	(10.14)	(2.49)
Constant	6.828^{***}	-3.499^{***}	
	(47.76)	(−26.55)	
Observations	14504337	14504337	14504337

z statistics in parentheses. Clustering is at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table reports maximum likelihood estimates of the IV probit model described in equations 2.2 and 2.3. Each observation represents a county-day. The second stage dependent variable is a dummy for the occurrence of a serious manufacturing workplace accident investigated by OSHA. The planetary boundary layer height and environmental lapse rate in 4 layers of air near the earth's surface are used as instruments for PM 2.5 exposure, which is measured in μg/m^3. First-stage coefficients are in column 1. Second-stage probit coefficients are in column 2; since interpretation of probit coefficients is not straightforward, estimated average marginal effects are given in column 3. County level fixed effects are included, and clustering is at the county level.

Table 2.5: LASSO and Polynomial Non-Linearity

	First Stage F-stat	LASSO Coefficients
PM 2.5	5638.21	
PM 2.5^2	2004.64	
PM 2.5^3	863.03	
PM 2.5^4	418.77	
PM 2.5^5	167.28	
Temperature (K)		
Precipitation		
Employment		
Observations	14504337	14504337

t statistics not given. County-level fixed effects included.
Notes: Table reports estimates of the two stage LASSO model described in equations 2.4 and 2.5 of the main text, using the square-root LASSO method of Belloni, Chernozhukov, et al. (2011) and Belloni, Chernozhukov, et al. (2014). Five first-stage equations (for the first through fifth powers of PM 2.5), with five powers of inversion coverage as instruments, were estimated using OLS; their *F*-statistics are reported in column 1. Then, predicted values of the five powers of PM 2.5 were used as inputs for LASSO to estimate equation 2.4. Estimated coefficients from LASSO are given in column 2. The package (*rlasso*) used for this estimation does not provide standard errors for LASSO estimation; in any case, such standard errors would be wrong given the manual nature of this two-stage estimation. Each observation represents a county-day, and the second stage dependent variable is a dummy for the occurrence of a serious construction worksite accident investigated by OSHA. County and day-of-week fixed effects are included.

and cognitive pathways by which PM 2.5 has these effects could help in developing measures to mitigate them.

Table 2.6: Effect of PM 2.5 on the Probability of Fatal Manufacturing Accidents

	Fatal Accidents as Dependent Variable (1)
Panel 1: First Stage	
Lapse Rate (Layer 1)	-82.031^{***}
	(-42.28)
Lapse Rate (Layer 2)	-22.972^{***}
	(-5.77)
Lapse Rate (Layer 3)	-107.097^{***}
	(-30.15)
Lapse Rate (Layer 4)	-33.178^{***}
	(-10.77)
PBL Height	-0.00124^{***}
	(-23.30)
Cooling Degree Days	0.592^{***}
	(-58.97)
Heating Degree Days	-0.0878^{***}
	(-58.97)
Precipitation	-0.1191^{***}
	(-169.06)
Surface Wind Speed	-0.265^{***}
	(-45.61)
Relative Humidity	0.0171^{***}
	(15.71)
Employment	0.000194^{***}
	(3.27)
F-statistic	2415.98
Panel 2: Second Stage	
PM 2.5 (μg/m^3)	0.00000450
	(1.27)
Heating Degree Days	0.00000271
	(1.00)
Cooling Degree Days	0.000000497
	(0.53)
Precipitation	0.000000199
	(0.25)
Surface Wind Speed	-0.00000156
	(-0.56)
Relative Humidity	0.000000564
	(1.62)
Employment	0.0000000477^{***}
	(6.25)
Observations	14710686

t statistics in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table reports IV estimates of the robustness check described in section 2.4.3. Each observation represents a county-day. The second stage dependent variable is a dummy for the occurrence of a fatal manufacturing accident investigated by OSHA. Both regressions include county, day-of-week, and month fixed effects. Standard errors are clustered at the county level.

CHAPTER 3

Fine Particulate Air Pollution and Traffic Safety

This paper provides causal estimates of the impact of fine particulate matter air pollution (PM 2.5) on vehicular traffic accidents in the United States, which resulted in 36,560 fatalities and 2.7 million injuries in 2018. I construct a panel dataset composed of the universe of fatal traffic accidents documented by NHTSA from 2003 to 2015 and exploit plausibly exogenous variation in PM 2.5 caused by changes in the height of the atmospheric planetary boundary layer height and the shape of the temperature-altitude curve in the lower atmosphere to implement an instrumental variables research design. I find that decreasing PM 2.5 exposure by 1 μg/m^3 could lead to a 2.4% decrease in the risk of fatal traffic accidents, representing an elasticity of 0.20.

3.1 Introduction

Fine particulate air pollution (PM 2.5) has long been known to cause a variety of physiological health harms,[1] but in recent years it has been shown to cause cognitive and behavioral harms as well.[2] Regulations intended to prevent these harms by reducing PM 2.5 have been somewhat successful, but high levels of PM 2.5 exposure are an ongoing health hazard in many parts of the world. However, the full range of effects that PM 2.5 has are still not full understood; neither are the pathways by which these effects occur. Designing optimal pollution control policies requires identifying these effects and quantifying their costs.[3]

Vehicular accidents impose huge costs on society. During 2018 in the U.S., 36,560 people were killed and 2.7 million were injured in traffic accidents, costing society an estimated $555 billion.[4] In this paper, I ask whether PM 2.5 exposure is a significant contributor to these accidents and resulting costs. Sager (2019) uses data from the United Kingdom, and finds that an increase of 1 μg/m^3 in PM 2.5 exposure leads to an increase of 0.4% in the number of vehicles involved in traffic accidents. He estimates that a one standard deviation reduction in London's PM 2.5 exposure for a single day could save £500,000. Given the much greater population of the U.S., and its greater cultural focus on driving, we might expect the U.S. effects to

[1] An extensive literature documents the negative effects of PM 2.5 on physical health outcomes, including infant mortality and adult mortality due to respiratory disease, lung cancer, and cardiopulmonary disease (for example, see Chowdhury and Dey 2016; Deryugina et al. 2019; Dominici et al. 2006; Pope III et al. 2019; Woodruff et al. 2006).

[2] More recently, links between PM 2.5 and mental health and cognitive function have begun to surface. Even in the very short term, exposure to PM 2.5 has been found to affect cognitive ability (Ebenstein et al. 2016), and criminal behavior and anxiety (Lu et al. 2018). These cognitive/behavioral impacts, in turn, can lead to perhaps unexpected consequences, such as an increase in violent crime (Burkhardt et al. 2019) or in car accidents (Sager 2019).

[3] Aldy et al. (2020) document the influence of so-called "co-benefits," benefits other than the intended target of a regulation, on regulatory decision making. They focus specifically on the EPA, and develop a conceptual framework to demonstrate that any and all benefits (or costs) resulting from a proposed regulation should be accounted for equally as part of its benefits (or costs) rather than being relegated to "co-benefit" or "co-cost" status. In the case of proposed PM 2.5 regulations, this implies that all pathways by which PM 2.5 harms society (pathways by which abatement benefits society) should be identified and treated equally as benefits.

[4] In 2019 dollars. This calculation is based on a value of a statistical life (VSL) of $10.6 million, from Viscusi (2018) (updated to 2019 dollars), and a value of a statistical injury (VSI) of $62,000, from Viscusi and Aldy (2003) (updated to 2019 dollars), and does not take into account property damage. It should therefore be considered a lower bound on the true cost to society.

be even more pronounced.

To estimate these effects, I construct a daily panel covering all counties in the contiguous United States from January 1, 2003 to December 31, 2015, using four data sources: the National Highway Transportation Safety Administration (NHTSA) Fatality Analysis Reporting System (FARS), the Quarterly Census of Employment and Wages (QCEW), the North American Regional Reanalysis (NARR), and new, high resolution, daily PM 2.5 prediction data produced by Di et al. (2019).

Because PM 2.5 is largely generated by human activities,[5] it is likely co-determined with other factors that impact traffic safety, such as number of cars on the road. To address this concern and identify the causal effects of PM 2.5, I need a plausibly exogenous source of variation in PM 2.5 to use in implementing an instrumental variables research design.

As this source of variation in PM 2.5 exposure, I use a suite of weather measurements that capture the capacity of the atmosphere to disperse pollutants generated near the surface upwards into the atmosphere. The first of these measurements is the height of the *planetary boundary layer*, the portion of the atmosphere which is strongly influenced by the earth's surface due to vertical mixing of air. Second, I use a group of temperature-based measurements that capture the *convective stability*, or tendency to circulate vertically, of the air at different levels above the surface. When the planetary boundary layer is low, or convective stability is high, pollutants are not dispersed as far upwards as they would otherwise be, concentrating such pollutants near the surface where they are generated. The identifying assumption of my study is that, conditional on surface weather (temperature, precipitation, humidity, and wind speed), atmospheric characteristics above the surface should not affect drivers at the surface except by their influence on air pollution.

I find that a significant portion of fatal traffic accidents may be caused by short-run exposure to PM 2.5. Estimates from my primary specification, a linear panel fixed effect IV model, suggest that reducing PM 2.5 exposure by 1 μg/m^3 could reduce the probability of a fatal traffic accident by 2.4%, (representing an elasticity of 0.20), and save \$13.3 billion per year.

These results indicate that the reduced risk of traffic accidents is an important benefit that should be accounted for when evaluating proposed PM 2.5 regulations, or other regulations that will impact PM 2.5.

[5] Especially via combustion processes, such as those which occur in car engines.

[5] Meteorological conditions are a common class of instruments for PM 2.5 exposure. Sager (2019) uses thermal inversions, a phenomenon closely related to the measures of convective stability I use. Deryugina et al. (2019) and Heyes and Zhu (2019) use wind direction. Non-meteorological instruments include advanced and inflexible scheduling of events (see Archsmith et al. 2018; Ebenstein et al. 2016).

3.2 Background

3.2.1 Fine Particulate Matter (PM 2.5) and Health

Fine particulate matter (PM 2.5) air pollution consists of particles less than 2.5 micrometers (μm) in diameter; for comparison, the average human hair is 50-70 μm in diameter. These very small particles arise from a variety of sources. Particles emitted directly into the atmosphere are referred to as "primary" particles. "Secondary" particles form in the atmosphere from chemicals emitted into the atmosphere as gases, and may form some distance from where the gases were originally emitted. These secondary particles constitute a large proportion of PM 2.5, as opposed to coarser particulate matter, which contains more primary particles. Major contributors to the formation of PM 2.5 secondary particles are sulfur dioxide (SO2), which is emitted by the combustion of fossil fuels and the smelting of metal ores containing sulfur, and nitrogen oxides (NOx), which are emitted by high temperature combustion, such as that found in vehicle engines and power plants. These and other anthropogenic emissions account for a large portion of ambient PM 2.5.

Because of the small size of PM 2.5 particles, they are absorbed into the bloodstream when inhaled, rather than being filtered out by the lungs. They are then transmitted throughout the body. An extensive literature documents the negative physiological health effects of chronic exposure to PM 2.5, including increased likelihood of death due to heart disease, heart attack, and lung cancer (see Achilleos et al. 2017). Largely due to these negative physiological health effects, both coarse particulate matter (PM 10) and fine particulate matter (PM 2.5) air pollution are listed as criteria pollutants under the National Ambient Air Quality Standards (NAAQS) established by the EPA.[6]

Some of the particles making up PM 2.5 are small enough to cross from the bloodstream into the brain, where they can impact cognitive and mental health. The effects of PM 2.5 on cognition and the brain have been well documented in the last few years. Gatto et al. (2014) and Ailshire and Crimmins (2014) show that long term exposure to PM 2.5 is connected to decreased cognitive functioning in middle aged and older adults in the United States.[7] Weuve et al. (2012) find that chronic PM 2.5 exposure leads to an increased rate of decline in cognitive ability with age. Fonken et al. (2011), in a medical study using mice, find that mice chronically exposed to elevated levels of PM 2.5 displayed depressive behavioral symptoms, in addition to decreased cognitive ability, specifically impairments in spatial learning and memory. They describe the specific brain regions that were found to be affected, including the hippocampus, one of the centers of both spatial learning and memory formation. Block and Calderón-Garcidueñas (2009) and Costa et al. (2014) suggest that PM 2.5 causes brain inflammation and oxidative stress, leading to various neuropathologies and central nervous system diseases.

---

[6] The current standard for PM 2.5 is 12 μg/m^3 (3-year average of annual mean concentration), significantly above the national mean PM 2.5 concentration (as of 2019) of 7.65 μg/m^3.

[7] The Gatto, et al. study is confined to Los Angeles.

The negative physical and cognitive/mental health effects of PM 2.5 are not only seen in those who experience chronic PM 2.5 exposure. Acute (short term) PM 2.5 exposure has also been shown to impact both physical and cognitive/mental health. Deryugina et al. (2019) use a large-scale database of Medicare recipients, and find that acute PM 2.5 exposure causes increased mortality, emergency hospitalization, and medical spending for up to two days following exposure. Ebenstein et al. (2016) find that acute exposure to elevated PM 2.5 levels causes an immediate decrease in cognitive ability. They use evidence from standardized testing of pre-college teens in Israel to show that elevated PM 2.5 at the testing location on test day causes lower scores, with subsequent persistent effects on educational attainment and earnings. Heyes and Zhu (2019) find that elevated levels of PM 2.5 lead to significantly increased sleeplessness, another possible pathway for physiological and cognitive harm.

Likely due to its impacts on cognition and behavior, short-term exposure to PM 2.5 has been shown by Burkhardt et al. (2019) to lead to an increase in violent crime. Interestingly, they find that lagged daily PM 2.5 has no impact on crime, suggesting that there are immediate, non-cumulative behavioral and cognitive effects of PM 2.5, although we have seen that chronic exposure has its own effects. Sager (2019) shows that short-term exposure to PM 2.5 causes an increase in vehicular accidents, and similarly to Burkhardt, et al., finds that lagged daily PM 2.5 has no impact.[8]

3.2.2 Pollution, Productivity, and Vehicular Safety

While the effect of pollution on traffic safety has not been extensively explored, very recent work by Sager (2019) and Wan et al. (2020) indicates that PM 2.5 contributes to an increased risk of traffic accidents. Additionally, PM 2.5 has been shown to directly impact worker productivity for both indoor and outdoor workers. Archsmith et al. (2018) use advanced scheduling of MLB baseball games as an instrument for umpire's PM 2.5 exposure and find that short-term PM 2.5 exposure leads to umpires, who work (mostly) outdoors, making more incorrect calls. T. Y. Chang et al. (2019) find that call center workers, who are indoors and performing almost entirely mental labor, are also impacted by PM 2.5 levels. It seems plausible that safe driving practices could be affected along with productivity. While the precise pathway by which productivity is affected has not been identified, cognition and behavior have both been shown to be impacted by PM 2.5, and could contribute to decreased productivity, unsafe driving behaviors, and impaired driving skill. Quantifying the effects of PM 2.5 on vehicular safety is particularly important because drivers are exposed to high levels of PM 2.5 while in traffic (see Knittel et al. 2016; Riediker et al. 2004).

[8]This also suggests that the sleeplessness pathway proposed by Heyes and Zhu is not the only pathway by which PM 2.5 affects cognition and behavior.

3.3 Research Design

3.3.1 Data

To analyze the effect of acute PM 2.5 exposure on traffic accidents, I require data on traffic accidents that includes, at minimum, the date of each accident and the county it occurred in. This data is readily available for fatal traffic accidents through the NHTSA Fatality Analysis reporting System (FARS). Since counties with more people in them might reasonably be expected to have more traffic accidents, I also need data on county populations, so I can control for them. For this purpose I use annual county population estimates from the Census Bureau.

For PM 2.5, I make use of a newly available dataset (Di et al. 2019) that uses a variety of data sources[9] and models of how pollutants are dispersed and transported through the atmosphere. The authors use an ensemble of machine learning algorithms to estimate daily PM 2.5 levels in 1 km by 1 km pixels across the contiguous United States for the years 2000 to 2015. Their final model predicts PM 2.5 levels well, with a 10-fold cross validated[10] R^2 averaging 0.86. Directly monitored PM 2.5 has two major disadvantages for my analysis: first, not all counties have PM 2.5 monitors; second, many existing monitors collect data only every three or every six days. Using the Di et al. (2019) PM 2.5 prediction dataset allows me to analyze the effect of PM 2.5 on traffic accidents across the entire contiguous U.S., and is more accurate than the interpolation methods I would have to use with monitor data to fill in missing days.

Weather data, including planetary boundary layer height and temperature at different levels above the surface (used to calculate the environmental lapse rate) as well as surface temperature, precipitation, wind speed, and humidity (to control for surface weather), come from the North American Regional Reanalysis (NARR) dataset (Mesinger et al. 2006). The NARR uses historical data from a variety of sources[11] to create a detailed picture of weather and climate conditions in 32×32 km square pixels across all of North America.

From these data sources, I construct a daily panel of U.S. counties covering the 48 contiguous U.S. states and the years 2003 to 2015. For each county-day, I include a dummy for whether a fatal traffic accident occurred, mean PM 2.5, planetary boundary layer height, environmental lapse rate in the bottom four layers of the atmosphere, county population, temperature, precipitation, humidity, and surface wind speed. Table 3.1 contains summary statistics.

[9] These include satellite images, ground-based monitor data, meteorological data, land-use data, and reanalysis datasets.

[10] k-fold cross validation refers to a method of testing the predictive power of a model by randomly partitioning the data into k equal subsamples. Each of these subsamples, in turn, is used as the validation data for a model trained using the other 9 samples, and a R^2 is calculated based on this validation.

[11] These include satellites, radiosondes (balloon mounted instruments), dropsondes (instruments dropped from aircraft), and surface instruments

Table 3.1: Summary Statistics by Season

	Spring	Summer	Fall	Winter	Total
Accident Occurred	0.0265	0.0298	0.0290	0.0247	0.0275
	(0.161)	(0.170)	(0.168)	(0.155)	(0.164)
PM 2.5 (μg/m^3)	8.514	10.85	8.174	8.260	8.954
	(4.960)	(6.348)	(5.354)	(5.323)	(5.631)
PBL Height	1007.1	1011.5	847.8	766.3	909.0
	(451.0)	(427.0)	(393.6)	(397.7)	(431.1)
Lapse Rate (Layer 1)	0.00139	0.00264	−0.000500	−0.000976	0.000650
	(0.00719)	(0.00468)	(0.00764)	(0.00974)	(0.00766)
Lapse Rate (Layer 2)	0.00547	0.00650	0.00492	0.00253	0.00487
	(0.00342)	(0.00212)	(0.00361)	(0.00614)	(0.00433)
Lapse Rate (Layer 3)	0.00623	0.00751	0.00585	0.00261	0.00556
	(0.00330)	(0.00179)	(0.00332)	(0.00526)	(0.00405)
Lapse Rate (Layer 3)	0.00592	0.00730	0.00550	0.00246	0.00531
	(0.00331)	(0.00173)	(0.00325)	(0.00469)	(0.00383)
Cooling Degree Days	1.189	6.551	1.695	0.0597	2.387
	(2.346)	(4.075)	(2.946)	(0.441)	(3.737)
Heating Degree Days	6.402	0.305	5.350	16.40	7.074
	(6.770)	(1.216)	(6.173)	(8.267)	(8.496)
Precipitation (mm)	2.791	3.066	2.351	2.135	2.589
	(6.889)	(6.684)	(6.822)	(5.917)	(6.602)
Surface Wind Speed	3.912	3.108	3.583	3.816	3.604
	(1.890)	(1.454)	(1.828)	(1.952)	(1.818)
Relative Humidity	69.19	67.05	68.43	76.34	70.23
	(15.55)	(16.56)	(15.72)	(14.54)	(16.02)
Population	97824.4	97824.4	97824.4	97822.4	97823.9
	(313552.7)	(313552.7)	(313552.7)	(313546.3)	(313551.1)
Observations	3,717,168	3,717,168	3,676,764	3,645,684	14,756,784

3.3.2 Methods

As discussed earlier, PM 2.5 and the chemicals which react to produce PM 2.5 in the air are largely human generated, from sources such as power plants, manufacturing, and vehicular traffic. Consequently, PM 2.5 exposure is co-determined with other activities (e.g. traffic itself) that contribute to traffic risk, leading to endogeneity. To account for this endogeneity, I use an instrumental variables research design with a suite of weather measurements as instruments for PM 2.5 exposure.

Various sources of exogenous variation in PM 2.5 have been used in other studies, including advanced and inflexible scheduling of events,[12] wind direction,[13] and thermal inversions.[14]

The set of instruments I use is similar to thermal inversions in that both are connected with the vertical dispersal, upwards into the atmosphere, of PM 2.5 generated near the earth's surface. First, I use the height of the *planetary boundary layer*. The planetary boundary layer is defined as the portion of the earth's surface in which wind and other atmospheric characteristics (temperature, pressure, and humidity) are strongly influenced by surface characteristics, due to vertical mixing of the air. The height of this layer, therefore, is a rough measure of the space in which pollutants generated near the surface may be dispersed due to this vertical mixing.

The second weather measurement, or group of measurements, that I use as instruments is a set of measurements of the rate of temperature change with height for each of four layers of air above the earth's surface. This rate of temperature change determines the vertical circulation of air via convection, based on the physical properties of air. Under theoretically ideal conditions, a parcel of air rising through the atmosphere will cool at a rate of 10°C/km, or 0.01°C/m; this rate is known as the *dry adiabatic lapse rate*. When the *environmental lapse rate*, or the rate of cooling with height actually observed in a region of the atmosphere, is less than the dry adiabatic lapse rate, then this region of the atmosphere is *convectively stable*, and vertical circulation due to convection is inhibited. The degree to which convection is inhibited depends on how low the environmental lapse rate is.

In the NARR (North American Regional Reanalysis) data, temperatures are given at different *pressure levels* of the atmosphere. A pressure level is defined as the point in the air column with a specified atmospheric

[12] Ebenstein et al. (2016) use the Israeli Bagrut college entrance exam, which is compulsory for all high school students, is scheduled years in advance, and which may not be rescheduled. This exam also takes place across several days, and scores are recorded by day, enabling the authors to use variation in PM 2.5 across different days for each individual student to identify the effect of PM 2.5 on test scores. Archsmith et al. (2018) use umpire schedules for Major League Baseball (MLB) games, which are set months in advance and which cannot readily be changed. These are excellent sources of exogenous variation in PM 2.5, but such ideal sources are not common.

[13] Wind has a strong effect on PM 2.5 exposure in many locations. The chief difficulty in using (lateral) wind direction as an instrument for PM 2.5 exposure is that its effect is completely dependent on location, since a given location may have pollution sources distributed arbitrarily around the points of the compass. To account for this, Heyes and Zhu (2019), who confine their study to 19 of China's most populous cities, construct a simple air transport model for each city in their sample. In contrast, Deryugina et al. (2019) use a non-parametric specification which allows the effect of wind direction on PM 2.5 to vary across groups of counties and makes no ex ante assumptions about the effect of wind in any given county.

[14] See Sager (2019).

Figure 3.1: Turbulence generated by wind passing over surface obstructions. The planetary boundary layer by definition lies above such turbulence.
Image credit: Pilot's Handbook of Aeronautical Knowledge, Federal Aviation Administration, page 12-10.

pressure (e.g. 950 millibars). At any given point on the earth's surface, the height of each pressure level above the surface varies with atmospheric/weather conditions and is included as an additional variable in the NARR data. I calculate the environmental lapse rate between the surface and the lowest pressure level that is at least 100 meters above the surface, then between each consecutive pair of pressure levels for the next three pressure levels (going up). This gives me environmental lapse rates for 4 different layers of air at each point. I then take the average over each county to get the average lapse rate in each of these four layers of air for the county-day.

The environmental lapse rate instrument is related to the thermal inversion instrument used by Sager (2019) and in my earlier paper (Chambers 2021b). A thermal inversion exists when cool air near the surface is overlain by warmer air. The cool surface air therefore does not tend to rise, i.e. vertical circulation due to convection is again inhibited. But in this case, by definition, the environmental lapse rate will be negative (since it measures the rate of cooling with height), definitely less than 0.01°C/m. The environmental lapse rate instrument thus incorporates a more detailed picture of atmospheric convective stability than simple thermal inversions.

These measures of vertical pollution dispersion capacity are simpler instruments for PM 2.5 than wind direction because the sign of their effect on surface PM 2.5 exposure is unambiguous, making it unnecessary to develop an air transport model, like Heyes and Zhu (2019), or utilize non-parametric methods, like Deryugina et al. (2019). This is a primary reason I have chosen this suite of measurements as instruments for PM 2.5 exposure.

Another important factor in vertical dispersion of pollutants generated near the service is surface wind speed. Faster wind generates more turbulence as it passes over irregularities on the surface, and this turbulence contributes to vertical mixing of the air. However, surface wind speed could plausibly have a direct impact on traffic safety, so I use it as a control variable and not as an instrument.

3.3.3 Model Specification

For my main results, I use a panel fixed effect IV specification, with the second stage as follows:

$$Y_{it} = \mu + \alpha \hat{D}_{it} + \beta' \mathbf{X}_{it} + \varphi_i + \varepsilon_{it} \tag{3.1}$$

where Y_{it} is a binary variable indicating whether a fatal traffic accident occurred in county i on date t; $\hat{D}_{it}$ represents PM 2.5 exposure caused by planetary boundary layer height and environmental lapse rate in four layers of air near the surface in county i on day t; $\mathbf{X}_{it}$ is a vector of covariates, including the population of the county that year, mean temperature, precipitation, relative humidity, and wind speed in that county that day, and day-of-week and calendar month dummies; φ_i is a county fixed effect; and ε_{it} is an error term, clustered at the county level.

The first stage of my model is:

$$D_{it} = \tau + \rho' \mathbf{Z}_{it} + \delta' \mathbf{X}_{it} + \theta_i + \nu_{it} \tag{3.2}$$

where D_{it} is PM 2.5 exposure; $\mathbf{Z}_{it}$ is a vector of instruments, namely the height of the planetary boundary layer and environmental lapse rates for four layers of air above the earth's surface; θ_i is a county level fixed effect; and ν_{it} is an error term, again clustered at the county level.

Because I am analyzing the effect of PM 2.5 on a probability (that of a serious worksite accident occurring), I also estimate an instrumental variables probit model with the second stage:

$$Y_{it} = \begin{cases} 1 & \mu + \alpha \hat{D}_{it} + \beta' \mathbf{X}_{it} + \varphi_i + \varepsilon_{it} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

where all variables are as defined above, except that ε_{it} is assumed to be standard normally distributed. The first stage of this model is precisely as given in equation 3.2.

There could be heterogeneity in the responsiveness of traffic accidents to PM 2.5 across baseline exposure levels, but it is not necessarily clear what form that heterogeneity might take. I analyze such potential heterogeneity by assuming that the response function can be represented by a high degree polynomial, then allowing the data to dictate which terms of this polynomial are actually included in the model. I do this via the Least Absolute Shrinkage and Selection Operator (LASSO), a machine learning regularization technique designed to prevent overfitting of a model to data. LASSO does so by incorporating a regularization term which penalizes the absolute value of the sum of the coefficients in the model. For example, for a simple regression equation $y = \beta'\mathbf{x} + \varepsilon$, OLS coefficients β_{OLS} are given by $\beta_{OLS} = \text{argmin}_\beta \frac{1}{2N}(y - \beta'\mathbf{x})'(y - \beta'\mathbf{x})$. LASSO coefficients are given by $\beta_{LASSO} = \text{argmin}_\beta \frac{1}{2N}(y - \beta'\mathbf{x})'(y - \beta'\mathbf{x}) + \lambda \sum_{j=i}^{p} |b_j|$, for some λ. The addition of this regularization term pushes the coefficients β_j toward zero, and precisely to zero in some cases. Variables x_j with non-zero estimated coefficients are those which have been selected as part of the model.

In this case, I explore a polynomial of up to 5 degrees in PM 2.5 by using the second stage equation:

$$Y_{it} = \mu + \sum_{k=1}^{5} \alpha_k \hat{D}_{it}^k + \beta'\mathbf{X}_{it} + \varphi_i + \varepsilon_{it} \tag{3.4}$$

where LASSO adds the term $\lambda \sum_{k=1}^{5} |\alpha_k|$ to the objective function to be minimized. This pushes the coefficients α_k toward zero, and precisely to zero if possible, thus selecting only the most important terms $\hat{D}_{it}^k$ to be part of the final model. The first stage equations are:

$$D_{it}^k = \tau_k + \sum_{l=1}^{5} \rho_{kl}'\mathbf{Z}_{it}^l + \delta_k'\mathbf{X}_{it} + \theta_{ki} + \nu_{kit} \qquad i \in \{1, \ldots, 5\} \tag{3.5}$$

with terms as defined above. In both stages, clustering is at the county level.

3.4 Results

I first present results from my main specification, using planetary boundary layer height and environmental lapse rate as instruments to estimate the causal effect of PM 2.5 exposure on the probability that a fatal traffic accident occurs. I then present results from alternative model specifications and robustness checks.

3.4.1 Main Results

Column 2 of tables 3.2 and 3.3 gives results for my preferred linear panel fixed effects specification. Table 3.2 shows first stage results, with an F-statistic of 2403.14, suggesting that planetary boundary height layer and environmental lapse rate in the four pressure layers of air nearest the earth's surface are a strong set of instruments for PM 2.5 exposure.

Table 3.3 gives the second stage estimates, which indicate that decreasing PM 2.5 exposure by 1 µg/m^3 would decrease the probability of a fatal or catastrophic manufacturing accident occurring by about 0.0624 percentage points; this represents roughly 2.4% of the baseline probability, for an elasticity of 0.20. If a similar percentage decrease were to hold across all types of traffic accidents, then this decrease in PM 2.5 exposure would result in about 877.4 fewer traffic fatalities and 64,800 fewer non-fatal traffic accident injuries annually, saving more than $13.3 billion per year.

These effects are larger than those found by Sager (2019), who estimates an elasticity of 0.06 for traffic

Table 3.2: Linear Panel Fixed Effects IV Estimates of the Effect of PM 2.5 on Accident Probability

First Stage	County & Weekday FEs (1)	County, Weekday, & Month FEs (2)
Lapse Rate (Layer 1)	-81.039^{***}	-79.965^{***}
	(-40.58)	(-41.33)
Lapse Rate (Layer 2)	-27.111^{***}	-21.940^{***}
	(-6.94)	(-5.52)
Lapse Rate (Layer 3)	-110.621^{***}	-106.722^{***}
	(-31.02)	(-30.13)
Lapse Rate (Layer 4)	-24.137^{***}	-35.994^{***}
	(-8.00)	(-11.76)
PBL Height	-0.000898^{***}	-0.00123^{***}
	(-18.89)	(-23.53)
Cooling Degree Days	0.527^{***}	0.588^{***}
	(110.42)	(93.26)
Heating Degree Days	-0.0423^{***}	-0.0881^{***}
	(-23.00)	(-60.30)
Precipitation	-0.118^{***}	-0.119^{***}
	(-169.87)	(-171.08)
Surface Wind Speed	-0.278^{***}	-0.263^{***}
	(-47.81)	(-45.96)
Relative Humidity	0.0196^{***}	0.0167^{***}
	(18.63)	(15.65)
Population	-0.0000179^{***}	-0.0000179^{***}
	(-5.69)	(-5.69)
County FEs	*X*	*X*
Weekday FEs	*X*	*X*
Calendar Month FEs		*X*
F-statistic	1772.41	2403.14
Observations	14754148	14754148

t statistics in parentheses. Standard errors are clustered at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.3: Linear Panel Fixed Effects IV Estimates of the Effect of PM 2.5 on Accident Probability

Second Stage	County & Weekday FEs (1)	County, Weekday, & Month FEs (2)
PM 2.5 (μg/m^3)	0.000624***	0.000471***
	(14.58)	(11.28)
Cooling Degree Days	−0.000276***	−0.000275***
	(−10.21)	(−8.35)
Heating Degree Days	−0.000280***	−0.000244***
	(−29.89)	(−17.80)
Precipitation	0.0000406***	0.000025***
	(4.43)	(2.71)
Surface Wind Speed	0.000105***	0.000072**
	(3.51)	(2.35)
Relative Humidity	−0.000057***	−0.0000548***
	(−12.73)	(−12.30)
Employment	−0.000000138***	0.000000141**
	(−5.06)	(−5.11)
County FEs	*X*	*X*
Weekday FEs	*X*	*X*
Calendar Month FEs		*X*
Observations	14754148	14754148

t statistics in parentheses. Standard errors are clustered at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table reports IV estimates of equation 3.1 (with first stage given by equation 3.2) from the main text. Each observation represents a county-day. The second stage dependent variable in all cases is a dummy for the occurrence of a serious manufacturing accident investigated by OSHA. The proportion of the county-day covered by a thermal inversion is used as an instrument for PM 2.5 exposure, which is measured in μg/m^3. All regressions include county and day-of-week fixed effects, with columns 2 and 3 adding calendar month or season fixed effects, respectively. Standard errors are clustered at the county level in all regressions, allowing for arbitrary serial correlation of standard errors.

accidents with respect to PM 2.5.[15] This paper adds to the evidence from Sager (2019) and Wan et al. (2020) that traffic safety is an important benefit of PM 2.5 abatement.

To repeat the comparative illustration found in Chambers (2021a) and Chambers (2021b), the regulatory impact analysis for the 2012 revisions to the National Ambient Air Quality Standards (NAAQS) predicts that lowering the NAAQS standard for 3-year rolling PM 2.5 mean concentration to 11 μg/m^3 from 15 μg/m^3 would result in a population weighted decrease in PM 2.5 of 0.207 μg/m^3, with quantifiable health benefits of $13.9–$33.6 billion, updated to 2019 dollars (EPA 2012). The results from this paper suggest that a 0.2 μg/m^3 decrease in nationwide mean PM 2.5 could result in 175 fewer traffic fatalities fatal and 12,960 fewer non-fatal traffic accident injuries annually, saving about $2.6 billion per year: an increase of at least 7.7–18.7% over EPA's estimated benefits. Together with Chambers (2021b) and Chambers (2021a), these results begin to document an entire class of safety benefits resulting from PM 2.5 abatement that have not formerly been considered.

I now report the results of alternative model specifications: a simple instrumental variables probit model and a model using LASSO to explore polynomial non-linearities in the response function of accident probability to PM 2.5.

3.4.2 Alternative Specifications

I also estimate an instrumental variables probit model, using the same set of variables as in the linear probability model described above. Results from this specification are given in table 3.4, with first stage results in column 1, and probit coefficients in column 2. Column 3 contains estimated average marginal effects at the means of all variables, which are essentially identical to the results from my primary specification.

Table 3.5 gives results from the two stage LASSO process described in section 3.3.3, equations 3.4 and 3.5.[16] LASSO selects the first and fifth powers of PM 2.5, as shown. The coefficients assigned to these two powers of PM 2.5 imply a significantly smaller effect of PM 2.5 near the (8.95 μg/m^3) mean level of PM 2.5 observed in my sample, but one which grows more rapidly with increasing PM 2.5, such that the two effects are equal at 34.1 μg/m^{3}of PM 2.5. This corresponds to the 99.7th percentile of PM 2.5 exposure in my sample. In other areas of the world, especially in India and China, such values are very common, even below average,[17] and PM 2.5 may be responsible for much larger increases in risk in those areas.

[15] On the other end of the scale, Deryugina et al. (2019) estimate an elasticity of 0.019 for elderly three-day mortality rates with respect to PM 2.5; Ebenstein et al. (2016) estimate an elasticity of 0.034 for exam scores with respect to PM 2.5.

[16] I use the package *lassopack* by Ahrens et al. (2020), implementing the square-root LASSO method of Belloni, Chernozhukov, et al. (2011) and Belloni, Chernozhukov, et al. (2014)

[17] See Y. Chen et al. (2020) and Tiwari et al. (2013)

Table 3.4: IV Probit Estimates of the Effect of PM 2.5 on Accident Probability

	First Stage	Probit Coefficients	Avg. Marginal Effects
PM 2.5		0.00866***	0.000495***
		(3.10)	(3.05)
PBL Height	−0.00200***		
	(−36.98)		
Lapse Rate (Layer 1)	−94.828***		
	(−32.15)		
Lapse Rate (Layer 2)	3.360		
	(0.43)		
Lapse Rate (Layer 3)	3.869		
	(0.49)		
Lapse Rate (Layer 4)	−96.263***		
	(−16.74)		
Cooling Degree Days	0.534***	−0.00239*	−0.000137*
	(93.69)	(−1.71)	(−1.70)
Heating Degree Days	−0.0695***	−0.0101***	−0.000577***
	(−20.29)	(−27.89)	(−22.25)
Precipitation	−0.122***	0.0000661	0.00000378
	(−134.75)	(0.15)	(0.15)
Surface Wind Speed	−0.175***	−0.00328*	0.000110*
	(−20.96)	(−1.69)	(−1.70)
Relative Humidity	0.0641***	0.000743	0.0000425
	(40.17)	(1.42)	(1.42)
Population	0.000000215*	0.000000751***	0.0000000430***
	(1.66)	(5.99)	(6.20)
Constant	6.775***	−2.0562***	
	(48.70)	(−47.68)	
Observations	14754148	14754148	14754148

z statistics in parentheses. Clustering is at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table reports maximum likelihood estimates of the IV probit model described in equations 3.2 and 3.3. Each observation represents a county-day. The second stage dependent variable is a dummy for the occurrence of a serious manufacturing workplace accident investigated by OSHA. The planetary boundary layer height and environmental lapse rate in 4 layers of air near the earth's surface are used as instruments for PM 2.5 exposure, which is measured in μg/m^3. First-stage coefficients are in column 1. Second-stage probit coefficients are in column 2; since interpretation of probit coefficients is not straightforward, estimated average marginal effects are given in column 3. County level fixed effects are included, and clustering is at the county level.

Table 3.5: LASSO and Polynomial Non-Linearity

	First Stage F-stat	LASSO Coefficients
PM 2.5	5723.12	0.0000157
PM 2.5^2	2022.87	
PM 2.5^3	860.93	
PM 2.5^4	407.31	
PM 2.5^5	151.90	0.000000000450
Temperature (K)		
Precipitation		-0.0000021
Employment		0.0000002
Observations	14754148	14754148

t statistics not given. County-level fixed effects included.
Notes: Table reports estimates of the two stage LASSO model described in equations 3.4 and 3.5 of the main text, using the square-root LASSO method of Belloni, Chernozhukov, et al. (2011) and Belloni, Chernozhukov, et al. (2014). Five first-stage equations (for the first through fifth powers of PM 2.5), with five powers of inversion coverage as instruments, were estimated using OLS; their F-statistics are reported in column 1. Then, predicted values of the five powers of PM 2.5 were used as inputs for LASSO to estimate equation 3.4. Estimated coefficients from LASSO are given in column 2. The package (*rlasso*) used for this estimation does not provide standard errors for LASSO estimation; in any case, such standard errors would be wrong given the manual nature of this two-stage estimation. Each observation represents a county-day, and the second stage dependent variable is a dummy for the occurrence of a serious construction worksite accident investigated by OSHA. County and day-of-week fixed effects are included.

3.5 Conclusion

This paper makes important contributions to the literature on traffic safety and on pollution and health (particularly cognitive health). I estimate that, at the U.S. average for PM 2.5 concentration, a reduction in PM 2.5 concentration of 1 μg/m^{3}could reduce the probability of a fatal traffic accident by 0.0624 percentage points, or roughly 2.4% of baseline risk, implying an elasticity of 0.20. This is an important, newly quantified benefit of regulations aimed at decreasing PM 2.5.

Additional research is needed to identify the effects on traffic safety of other airborne pollutants, such as ozone. Additionally, understanding the specific physiological and cognitive pathways by which PM 2.5 has these effects will help in developing measures to mitigate them.

References

Achilleos, S., Kioumourtzoglou, M.-A., Wu, C.-D., Schwartz, J. D., Koutrakis, P., & Papatheodorou, S. I. (2017). Acute effects of fine particulate matter constituents on mortality: A systematic review and meta-regression analysis. *Environment international*, *109*, 89–100.

Adhvaryu, A., Kala, N., & Nyshadham, A. (2014). Management and shocks to worker productivity: Evidence from air pollution exposure in an indian garment factory. *Unpublished working paper, University of Michigan*.

Ahrens, A., Hansen, C. B., & Schaffer, M. E. (2018). Pdslasso and ivlasso: Programs for post-selection and post-regularization ols or iv estimation and inference. http://ideas.repec.org/c/boc/bocode/s458459.html

Ahrens, A., Hansen, C. B., & Schaffer, M. E. (2020). Lassopack: Model selection and prediction with regularized regression in stata. arxiv:1901.05397

Ailshire, J. A., & Crimmins, E. M. (2014). Fine particulate matter air pollution and cognitive function among older us adults. *American journal of epidemiology*, *180*(4), 359–366.

Aldy, J. E., Kotchen, M., Evans, M. F., Fowlie, M., Levinson, A., & Palmer, K. (2020). *Co-benefits and regulatory impact analysis: Theory and evidence from federal air quality regulations* (tech. rep.). National Bureau of Economic Research.

Archsmith, J., Heyes, A., & Saberian, S. (2018). Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation. *Journal of the Association of Environmental and Resource Economists*, *5*(4), 827–863.

Belloni, A., Chen, D., Chernozhukov, V., & Hansen, C. (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, *80*(6), 2369–2429.

Belloni, A., Chernozhukov, V., & Wang, L. (2011). Square-root lasso: Pivotal recovery of sparse signals via conic programming. *Biometrika*, *98*(4), 791–806.

Belloni, A., Chernozhukov, V., Wang, L. et al. (2014). Pivotal estimation via square-root lasso in nonparametric regression. *The Annals of Statistics*, *42*(2), 757–788.

Block, M. L., & Calderón-Garcidueñas, L. (2009). Air pollution: Mechanisms of neuroinflammation and cns disease. *Trends in neurosciences*, *32*(9), 506–516.

Bureau of Labor Statistics. (2019a). *Census of fatal occupational injuries* (tech. rep.). Bureau of Labor Statistics.

Bureau of Labor Statistics. (2019b). *Survey of occupational injuries and illnesses* (tech. rep.). Bureau of Labor Statistics.

Bureau of Labor Statistics. (2020a). *Census of fatal occupational injuries* (tech. rep.). Bureau of Labor Statistics.

Bureau of Labor Statistics. (2020b). *Survey of occupational injuries and illnesses* (tech. rep.). Bureau of Labor Statistics.

Burkhardt, J., Bayham, J., Wilson, A., Carter, E., Berman, J. D., O'Dell, K., Ford, B., Fischer, E. V., & Pierce, J. R. (2019). The effect of pollution on crime: Evidence from data on particulate matter and ozone. *Journal of Environmental Economics and Management*, *98*, 102267.

Chambers, M. (2021a). Fine particulate air pollution and manufacturing worker safety. *Unpublished working paper, Vanderbilt University*.

Chambers, M. (2021b). Pm 2.5 and construction worksite safety. *Unpublished working paper, Vanderbilt University*.

Chang, T., Graff Zivin, J., Gross, T., & Neidell, M. (2016). Particulate pollution and the productivity of pear packers. *American Economic Journal: Economic Policy*, *8*(3), 141–69.

Chang, T. Y., Graff Zivin, J., Gross, T., & Neidell, M. (2019). The effect of pollution on worker productivity: Evidence from call center workers in china. *American Economic Journal: Applied Economics*, *11*(1), 151–72.

Chen, Y., Wild, O., Conibear, L., Ran, L., He, J., Wang, L., & Wang, Y. (2020). Local characteristics of and exposure to fine particulate matter (pm2. 5) in four indian megacities. *Atmospheric Environment: X*, *5*, 100052.

Chowdhury, S., & Dey, S. (2016). Cause-specific premature death from ambient pm2. 5 exposure in india: Estimate adjusted for baseline mortality. *Environment international*, *91*, 283–290.

Cole, M. A., Elliott, R. J., & Lindley, J. K. (2009). Dirty money: Is there a wage premium for working in a pollution intensive industry? *Journal of risk and uncertainty*, *39*(2), 161–180.

Costa, L. G., Cole, T. B., Coburn, J., Chang, Y.-C., Dao, K., & Roque, P. (2014). Neurotoxicants are in the air: Convergence of human, animal, and in vitro studies on the effects of air pollution on the brain. *BioMed research international*, *2014*.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, *109*(12), 4178–4219.

Di, Q., Amini, H., Shi, L., Kloog, I., Silvern, R., Kelly, J., Sabath, M. B., Choirat, C., Koutrakis, P., Lyapustin, A. et al. (2019). An ensemble-based model of pm2. 5 concentration across the contiguous united states with high spatiotemporal resolution. *Environment international*, *130*, 104909.

Dominici, F., Peng, R. D., Bell, M. L., Pham, L., McDermott, A., Zeger, S. L., & Samet, J. M. (2006). Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *Jama*, *295*(10), 1127–1134.

Ebenstein, A., Lavy, V., & Roth, S. (2016). The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics*, *8*(4), 36–65.

EPA. (2012). Regulatory impact analysis for the final revisions to the national ambient air quality standards for particulate matter. (EPA-452/R-12-005).

Fonken, L. K., Xu, X., Weil, Z. M., Chen, G., Sun, Q., Rajagopalan, S., & Nelson, R. J. (2011). Air pollution impairs cognition, provokes depressive-like behaviors and alters hippocampal cytokine expression and morphology. *Molecular psychiatry*, *16*(10), 987–995.

Gatto, N. M., Henderson, V. W., Hodis, H. N., John, J. A. S., Lurmann, F., Chen, J.-C., & Mack, W. J. (2014). Components of air pollution and cognitive function in middle-aged and older adults in los angeles. *Neurotoxicology*, *40*, 1–7.

Guo, H., Morawska, L., He, C., Zhang, Y. L., Ayoko, G., & Cao, M. (2010). Characterization of particle number concentrations and pm 2.5 in a school: Influence of outdoor air pollution on indoor air. *Environmental Science and Pollution Research*, *17*(6), 1268–1278.

Haviland, A., Burns, R., Gray, W., Ruder, T., & Mendeloff, J. (2010). What kinds of injuries do osha inspections prevent? *Journal of safety research*, *41*(4), 339–345.

Haviland, A. M., Burns, R. M., Gray, W. B., Ruder, T., & Mendeloff, J. (2012). A new estimate of the impact of osha inspections on manufacturing injury rates, 1998–2005. *American journal of industrial medicine*, *55*(11), 964–975.

Heyes, A., & Zhu, M. (2019). Air pollution as a cause of sleeplessness: Social media evidence from a panel of chinese cities. *Journal of Environmental Economics and Management*, *98*, 102247.

Kniesner, T. J., & Leeth, J. D. (2013). Regulating occupational and product risks. In M. Machina & W. K. Viscusi (Eds.), *Handbook of the economics of risk and uncertainty* (pp. 601–650). Newnes.

Knittel, C. R., Miller, D. L., & Sanders, N. J. (2016). Caution, drivers! children present: Traffic, pollution, and infant health. *Review of Economics and Statistics*, *98*(2), 350–366.

Landrigan, P. J. (1992). Commentary: Environmental disease–a preventable epidemic. *American journal of public health*, *82*(7), 941–943.

Lu, J. G., Lee, J. J., Gino, F., & Galinsky, A. D. (2018). Polluted morality: Air pollution predicts criminal activity and unethical behavior. *Psychological Science*, *29*(3), 340–355.

Meier, R., Eeftens, M., Phuleria, H. C., Ineichen, A., Corradi, E., Davey, M., Fierz, M., Ducret-Stich, R. E., Aguilera, I., Schindler, C. et al. (2015). Differences in indoor versus outdoor concentrations of ultrafine particles, pm 2.5, pm absorbance and no 2 in swiss homes. *Journal of exposure science & environmental epidemiology*, *25*(5), 499–505.

Mendeloff, J., & Gray, W. B. (2005). Inside the black box: How do osha inspections lead to reductions in workplace injuries? *Law & Policy*, *27*(2), 219–237.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H. et al. (2006). North american regional reanalysis. *Bulletin of the American Meteorological Society*, *87*(3), 343–360.

OSHA. (2005). Fatality/catastrophe investigation procedures. (CPL 02-00-137).

Pope III, C. A., Lefler, J. S., Ezzati, M., Higbee, J. D., Marshall, J. D., Kim, S.-Y., Bechle, M., Gilliat, K. S., Vernon, S. E., Robinson, A. L. et al. (2019). Mortality risk and fine particulate air pollution in a large, representative cohort of us adults. *Environmental health perspectives*, *127*(7), 077007.

Riediker, M., Cascio, W. E., Griggs, T. R., Herbst, M. C., Bromberg, P. A., Neas, L., Williams, R. W., & Devlin, R. B. (2004). Particulate matter exposure in cars is associated with cardiovascular effects in healthy young men. *American journal of respiratory and critical care medicine*, *169*(8), 934–940.

Sager, L. (2019). Estimating the effect of air pollution on road safety using atmospheric temperature inversions. *Journal of Environmental Economics and Management*, *98*, 102250.

Tan, B. Y. (2020). Making up for lost time: The effect of historical pollution on covid-19 morbidity and mortality. *Unpublished working paper, Vanderbilt University*.

Tiwari, S., Srivastava, A., Bisht, D., Parmita, P., Srivastava, M. K., & Attri, S. (2013). Diurnal and seasonal variations of black carbon and pm2. 5 over new delhi, india: Influence of meteorology. *Atmospheric Research*, *125*, 50–62.

Viscusi, W. K. (2018). *Pricing lives: Guideposts for a safer society*. Princeton University Press.

Viscusi, W. K., & Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *Journal of risk and uncertainty*, *27*(1), 5–76.

Wan, Y., Li, Y., Liu, C., & Li, Z. (2020). Is traffic accident related to air pollution? a case report from an island of taihu lake, china. *Atmospheric Pollution Research*, *11*(5), 1028–1033.

Weuve, J., Puett, R. C., Schwartz, J., Yanosky, J. D., Laden, F., & Grodstein, F. (2012). Exposure to particulate air pollution and cognitive decline in older women. *Archives of internal medicine*, *172*(3), 219–227.

Woodruff, T. J., Parker, J. D., & Schoendorf, K. C. (2006). Fine particulate matter (pm2. 5) air pollution and selected causes of postneonatal infant mortality in california. *Environmental Health Perspectives*, *114*(5), 786–790.

Zhao, L., Chen, C., Wang, P., Chen, Z., Cao, S., Wang, Q., Xie, G., Wan, Y., Wang, Y., & Lu, B. (2015). Influence of atmospheric fine particulate matter (pm2. 5) pollution on indoor environment during winter in beijing. *Building and Environment*, *87*, 283–291.