NBER WORKING PAPER SERIES

THE MORTALITY AND MEDICAL COSTS OF AIR POLLUTION: EVIDENCE FROM CHANGES IN WIND DIRECTION

Tatyana Deryugina
Garth Heutel
Nolan H. Miller
David Molitor
Julian Reif

Working Paper 22796
http://www.nber.org/papers/w22796

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2016, Revised February 2018

We thank Allen C. Basala, Alex Hollingsworth, Pierre Léger, Christos Makridis, Mar Reguant, and seminar participants at Cornell University, IGPA, the University of Illinois Research Lunch, the University of Iowa, Georgia Tech, the Research Triangle Institute, the Midwestern Health Economics Conference, the AERE Summer Conference, the Southern Economics Association Conference, and the IZA Conference on Labor Market Effects of Environmental Policies for helpful comments. Eric Zou provided excellent research assistance. Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Number R01AG053350. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health nor of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Tatyana Deryugina, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction
Tatyana Deryugina, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif
NBER Working Paper No. 22796
November 2016, Revised February 2018
JEL No. I1,Q53

ABSTRACT

We estimate the causal effects of acute fine particulate matter (PM 2.5) exposure on mortality and health care use among the US elderly using Medicare administrative data and a novel instrument for air pollution: changes in the local wind direction. We then develop a new methodology that uses machine learning to estimate the number of life-years lost due to PM 2.5. We find that, while unhealthy individuals are disproportionately vulnerable to air pollution, the largest aggregate burden is borne by those with medium life expectancy, who are both vulnerable and comprise a large share of the elderly population.

Tatyana Deryugina
Department of Finance
University of Illinois at Urbana-Champaign
515 East Gregory Drive, MC-520
Champaign, IL 61820
and NBER
deryugin@illinois.edu

Garth Heutel
436 Andrew Young School
Department of Economics
Georgia State University
PO Box 3992
Atlanta, GA 30302-3992
and NBER
gheutel@gsu.edu

Nolan H. Miller
College of Business
University of Illinois
4033 BIF
515 East Gregory Drive
Champaign, IL 61820
and NBER
nmiller@illinois.edu

David Molitor
University of Illinois at Urbana-Champaign
340 Wohlers Hall
1206 S. Sixth Street
Champaign, IL 61820
and NBER
dmolitor@illinois.edu

Julian Reif
University of Illinois at Urbana-Champaign
College of Business
515 East Gregory Drive
Champaign, IL 61820
jreif@illinois.edu

An online appendix is available at http://www.nber.org/data-appendix/w22796

It is widely accepted that exposure to air pollution negatively affects human health, leading many countries to regulate air pollution levels. Accurately quantifying the health effects of marginal pollution reductions matters greatly for determining optimal environmental policy, especially for countries like the United States, where current pollution levels are relatively low and further reductions may be very costly. However, estimating the causal effect of pollution on health is complicated due to well-documented challenges, including separately identifying the effects of different pollutants, omitted variable bias, and measurement error. Quasi-experimental studies that use a plausibly exogenous source of pollution variation are typically confined to narrow geographic and temporal scales, raising questions of external validity. Such studies also lack power to detect changes in important but rare outcomes like adult mortality due to relatively small sample sizes, and thus they may overlook an important component of the social cost of pollution. Even if mortality effects are detected, estimating the mortality *cost* of pollution in terms of life-years lost is difficult because those who die prematurely from pollution may have shorter life expectancies than those who survive.

This paper presents the first large-scale, quasi-experimental investigation of the effects of acute (short-term) fine particulate matter exposure on mortality and medical costs among the elderly. We overcome the identification and statistical power challenges described above by exploiting daily variation in fine particulate matter (PM 2.5) concentrations caused by changes in daily wind direction to estimate the causal effect of pollution on three-day county-level elderly mortality rates, life-years lost, hospitalizations, and medical spending. The identifying assumption of our instrumental variables (IV) approach is that, after flexibly controlling for a large number of fixed effects and climatic variables, changes in a county's daily wind direction are unrelated to changes in the county's mortality or health care utilization except through air pollution.

A key innovation of our study relative to previous quasi-experimental designs exploiting wind variation is that our approach does not require understanding the detailed layout of an area (e.g., locations of roads, rivers, and population centers) or identifying the source of pollution to estimate its effects. This allows us to harness variation in pollution across a broad geographic scale and over a long time period, enabling us to estimate effects on rare health outcomes such as mortality and to explore effect heterogeneity across subpopulations. We do this by combining data on the universe of elderly Medicare beneficiaries, comprising approximately 97 percent of the US population aged 65 and older, with pollution and weather data from 1999 through 2011. The comprehensiveness of our dataset coupled with our novel methodology also allows us to separately identify the effects of different pollutants on mortality, which has proven to be extremely challenging.

We estimate that a 1 microgram per cubic meter ($\mu g/m^3$) (about 10 percent of the mean) increase in PM 2.5 exposure for one day causes 0.61 additional deaths per million elderly individuals over the three-day window consisting of the day of the increase and the following two days. The effect is largest for the oldest beneficiaries in absolute terms. However, the relative mortality risk changes non-monotonically with age, suggesting that age alone is a noisy predictor of vulnerability to air pollution. Our IV estimates are larger than both our corresponding ordinary least squares (OLS) estimates and the results reported in the prior literature, demonstrating the potential for substantial bias in observational studies of pollution exposure. Finally, the IV estimate is robust to simultaneously instrumenting for PM 2.5, carbon monoxide, and ozone, which is feasible because different wind directions transport varying amounts of each pollutant.

We also find that increases in PM 2.5 lead to more emergency room (ER) visits, more hospitalizations, and higher inpatient spending, driven entirely by admissions that originate in the ER. Each 1-$\mu g/m^3$ increase in PM 2.5 increases three-day ER visits by 2.3 per million beneficiaries and ER spending by over $15,000 per million. OLS estimates are again much smaller and sometimes significantly *negative*. As a placebo test, we find no effect of PM 2.5 on planned (non-ER) hospital admissions.

A central concern that arises when estimating mortality effects is whether those who die from pollution exposure would have died soon anyway, a phenomenon referred to as "mortality displacement" or "harvesting." If the mortality effect of pollution is concentrated among relatively old or sick individuals, then the mortality cost, as measured by the number of life-years lost, is likely to be much smaller than if the effect were more evenly distributed across the population. Some studies address this issue by using lags of the independent variables of interest to investigate whether mortality effects decrease as the length of time under consideration increases, as would be the case under harvesting (Schlenker and Walker 2016), or by averaging pollution fluctuations over longer time periods. We show that the effect of one-day PM 2.5 exposure on mortality grows if we expand the time window over which we measure mortality from 3 days to 5, 10, or 14 days, suggesting that our main results are not merely artefacts of short-term mortality displacement. However, this and other traditional approaches cannot account for displacement that occurs outside the time window spanned by the dependent variable.

We develop a novel, direct approach to estimate the number of life-years lost due to pollution exposure. To our knowledge, no other study has incorporated information beyond age and sex when accounting for life-years lost. By contrast, we employ machine learning techniques to incorporate over one thousand individual- and neighborhood-level variables from Medicare health histories and the American Community Survey into a survival model and use the results to predict remaining life expectancy for each individual in our sample. We then aggregate the estimates of decedents' counterfactual life expectancies up to the county level to provide daily measures of life-years lost per capita and use them to directly estimate the life-years lost due to pollution exposure.

Our life-years lost analysis reveals that accounting for decedents' age and gender reduces estimates of life-years lost by 31 percent compared to a naïve estimate that controls for neither. Accounting for the rich medical history data reduces the life-years lost estimate by an additional 55 percent relative to using only age and gender. Our preferred estimate is that a 1-μg/m^3 increase in PM 2.5 causes the loss of 2.7 life-years per million beneficiaries over three days. Due to their high mortality rates, those with a life expectancy of less than one year lose the largest number of life years per capita in both absolute (11.3 per million) and relative terms. Whereas the relative mortality effects of PM 2.5 are non-monotonic with respect to age, both our mortality and life-years lost estimates decrease steeply with counterfactual life expectancy. This suggests that our model of life expectancy identifies vulnerability to pollution shocks more effectively than age alone.

Although individuals with a life expectancy of less than one year bear the largest mortality and life-years lost burden in per capita terms, they are not the primary contributors to the *aggregate* social burden of pollution because they comprise less than one percent of all beneficiaries. The social cost of PM 2.5 is concentrated among the elderly with 5-10 years of remaining life expectancy, followed by those with 2-5 years remaining, because these groups represent a large fraction of the Medicare population and are affected non-trivially by acute particulate matter exposure.

Using a conventional value of $100,000 per statistical life year (Cutler, 2004), our estimates of life-years lost imply that the social mortality cost of a 1-μg/m^3 increase in PM 2.5 is $270,000 per million beneficiaries, which is an order of magnitude larger than our hospitalization cost estimate of $15,000 per million beneficiaries. To put these results into perspective, consider the national reduction in average PM 2.5 concentrations of 3.65 μg/m^3 that occurred during our study period, 1999-2011 (see Figure 1). Scaling our estimates linearly, by 2011 this reduction decreased the number of elderly deaths nationwide by 55,000 per year and the number of life-years lost by 150,000 per year. A standard value of $100,000 per statistical life-year implies a corresponding benefit of $15 billion per year, which represents a large fraction of the estimated annual costs of complying with air pollution regulations (EPA 2011).[1] By comparison, estimating life-years lost using an average life expectancy for the population increases this estimate by 220 percent, to $47 billion. Accounting for gender and age of the decedents mitigates this upward bias, but still causes the benefits to be overestimated by 120 percent.

Evidence supporting fine particulate matter regulation has come primarily from associational studies that have consistently found a relationship between PM 2.5 and increased morbidity and mortality, even after controlling for various confounding factors (e.g., Dockery et al. 1993, Pope et al. 1995, Laden et al. 2000, Samet et al. 2000, Pope and Dockery 2006, EPA 2009). The majority of these epidemiological,

[1] The EPA's calculation of the annual costs of meeting the 1990 Clean Air Act Amendment air quality standards (which include standards for all criteria pollutants, not just PM 2.5) increased from $19.9 billion in 2000 to $43.9 billion in 2010 (EPA 2011). Standards for PM 2.5 were first implemented in 1997, and then tightened in 2006.

associational studies focus on the effects of short-term (usually daily) exposure (Pope 2000), suggesting that even transient increases in particulate matter can have significant consequences. However, concerns about bias in associational estimates have caused both the scientific community and regulators to question how many deaths are avoided from reductions in particulate matter (OMB 2012; Dominici et al. 2014). While randomized controlled laboratory trials have shown that healthy volunteers exposed to ambient pollution for as little as one or two hours have worse cardiovascular performance than those exposed to very clean air, these studies face issues of external validity and are too small to draw conclusions about mortality effects (Brook et al. 2009, Langrish et al. 2013). Thus, significant uncertainty remains about the causal effects of acute PM 2.5 exposure on human health.

Our study addresses this uncertainty by providing the first quasi-experimental estimates of the causal effect of acute PM 2.5 exposure on adult mortality, hospitalizations, and medical costs. Our work contributes to the recent literature in economics that uses quasi-experimental approaches to estimate the effects of pollution on health. Much of this work has focused on the effect of pollutants other than fine particulate matter, such as TSP, PM 10, ozone, sulfur dioxide, or nitrogen oxides (Chay et al. 2003; Chay and Greenstone 2003; Currie and Neidell 2005; Currie et al. 2009; Moretti and Neidell 2011; Chen et al. 2013; Schlenker and Walker 2016; Deryugina et al. 2016; Knittel et al. (2016); Deschenes et al. 2016). Of these studies, only four consider non-infant mortality (Chay et al. 2003; Chen et al. 2013; Deryugina et al. 2016; Deschenes et al. 2016), but they do not estimate the effects of fine particulate matter. Ward (2015) focuses on PM 2.5, but only considers hospitalizations from respiratory causes in the province of Ontario. Finally, Anderson (2015) uses variation in wind direction across a highway in Los Angeles to proxy for changes in air pollution, but does not directly measure which pollutants are changing and focuses on chronic rather than acute pollution exposure.

Our study moves beyond these papers in three important ways. First, our approach allows us not only to estimate the causal impact of PM 2.5 on mortality, but also to separately identify the causal impact of other pollutants on mortality. We find that the PM 2.5-mortality relationship is more robust than that of other pollutants. Second, our study includes elderly mortality as an outcome, which few previous studies have considered. We find that the mortality costs of PM 2.5 among the elderly are an order of magnitude larger than the health care costs, demonstrating that ignoring mortality can cause researchers to overlook a primary social cost of pollution. Third, we estimate mortality costs more accurately than previous studies by developing and applying a novel method to estimate the life-years lost associated with pollution exposure. Our estimates suggest that traditional methods for estimating life-years lost are prone to significant upward bias. Moreover, our approach allows for an unprecedented investigation of the distributional effects of the mortality costs of air pollution by health status. For instance, we show that unhealthy individuals (as measured by life expectancy) bear a disproportionate share of the life-years lost burden of PM 2.5 despite their low life expectancies.

Our methodology for estimating life-years lost is general and can be applied to investigate mortality costs across a wide variety of contexts. For example, whether health insurance reduces mortality is an important question in health economics (Finkelstein and McKnight 2008; Card et al. 2009; Huh and Reif 2017). As in our study, estimating the social value of that mortality reduction depends on the number of life-years saved. Using our approach, this quantity can be estimated with administrative datasets such as Medicare or other surveys that contain information on demographics, health status, and mortality, such as the Health and Retirement Study or the Panel Study of Income Dynamics.

The rest of the paper is organized as follows. Section II provides a brief background on fine particulate matter, summarizes how air pollution is transported by the wind, and gives a preview of our estimation strategy. Section III describes our data. Section IV describes our econometric strategy in detail, including how we estimate the life-years lost. Section V presents results, and Section VI concludes.

II. Background

Fine particulate matter, PM 2.5, refers to particles with diameters of 2.5 microns or less. Rather than having a single chemical composition, PM 2.5 is a mixture of various compounds including nitrates, sulfates, ammonium, and carbon (Kundu and Stone 2014). In addition to natural sources, PM 2.5 is created from atmospheric conversion of power plant and auto emissions.[2] For the past several decades, the Environmental Protection Agency (EPA) has been tightening its regulation of particulates, focusing increasingly on fine particulates. It has regulated particulate matter smaller than 100 micrometers in diameter (total suspended particulates or TSP) since 1971. Concerned with growing epidemiological evidence that smaller particulate matter was especially harmful, and that both acute and chronic (long-term) exposure mattered, in 1987 the EPA set a daily and an annual standard for particulate matter less than 10 micrometers in diameter (PM 10). Similar concerns prompted the EPA in 1997 to set limits on fine particulate matter (PM 2.5), defined as particles less than 2.5 micrometers in diameter. The daily limit for PM 2.5 was tightened in 2006, and the annual limit was tightened in 2012.[3]

The PM 2.5 present in a given location will consist of both locally-produced pollution and pollution produced elsewhere that is transported into the region by the wind.[4] The amount of transported pollution is

---

[2] While not themselves particulates, sulfur dioxide and nitrogen dioxide, two "criteria" pollutants regulated by the EPA under the Clean Air Acts, are precursors to sulfates and nitrates, which are components of PM 2.5.

[3] See https://www3.epa.gov/ttn/naaqs/standards/pm/s_pm_history.html.

[4] PM 2.5 is not unique in its ability to traverse considerable distances; other pollutants, including carbon monoxide, sulfur dioxide, nitrogen dioxide, and ozone precursors, can also be carried by the wind.

significant (Zhang et al. 2017). For example, the EPA estimates that most of the PM 2.5 in the Eastern United States was not produced locally, but instead transported from hundreds of miles away (EPA 2004). Pollution transport patterns depend on a host of factors, including the pollutant, the location of the pollution source, wind direction and speed, precipitation, the height of the planetary boundary layer, and the presence of other airborne molecules, which can react with the windborne pollutant. One way to exploit variation in pollution transport is to employ a sophisticated atmospheric science model (e.g., Muller and Mendelsohn 2007) to simulate daily pollution transport across the United States and use the resulting estimates as instruments. However, this is computationally infeasible at the daily level.

An instrumental variables approach, however, need not use all of the factors involved in pollution transport. Such an approach requires only that the instrument (a) be sufficiently correlated with the endogenous variable of interest and (b) not be correlated with any unobserved determinants of the outcome of interest. We instrument for changes in a county's daily average PM 2.5 concentrations using changes in the county's daily average wind direction, which we shall show is by itself an important determinant of pollution levels. We do not use *prevailing* wind directions because the predictability of prevailing winds may affect the placement of pollution monitors or cause individuals to sort into upwind or downwind locations, thereby biasing the estimates. Employing variation attributable to *changes* in wind direction alleviates these concerns, but also means that our method is most useful for examining acute, rather than chronic, exposure.

Wind may affect pollution measured by a particular monitor either by redistributing locally-produced pollution (e.g., from traffic or local power plants) or by transporting externally-produced pollution into the county. As we discuss in detail later, we construct our empirical specification to exploit primarily the wind-induced variation in pollution exposure that affects the whole county in a similar manner. This variation, which we argue is more likely to arise from transport of pollution produced outside the county, reduces the potential for measurement error in residents' pollution exposure due to within-county transport. The Online Appendix presents evidence that the relationship between PM 2.5 and wind direction is very stable across monitors in the same broad geographic area, suggesting that our variation is indeed driven by non-local pollution sources.

We now illustrate the type of variation used to estimate the causal effects of PM 2.5, relegating the details to Section IV. Figure 2 shows the relationship between the estimated daily wind direction at pollution monitors, in 10-degree bins, and PM 2.5 concentrations measured by these monitors in and around the San Francisco Bay Area, CA. Figure 3 shows the same relationship for pollution monitors in and around Greater Boston, MA. All estimates are relative to 260-270 degrees, where 270 degrees corresponds to a "Westerly" (blowing *from* the West) wind direction. The figures display results from a regression that controls for county, month-by-year, and state-by-month fixed effects, as well as a flexible set of controls for maximum and minimum temperatures, precipitation, wind speed, and the interactions between them.

In both figures, the change in local wind direction is a very strong predictor of changes in local pollution levels, and the patterns are consistent with the geographic placement of the monitors. In and around the Bay Area, PM 2.5 levels are highest when the wind is blowing from the Southeast and lowest when the wind is blowing from the West and the North. In other words, more pollution is blown in from Southeast California than from the ocean and the Northern states like Oregon and Washington. In and around Boston, MA, pollution is highest when the wind is blowing from the Southwest, where New York City is located, and lowest when it is blowing from the East, North, Northeast and Northwest, where the ocean and sparsely populated areas dominate.

III. Data

A. Air pollution

We obtain air pollution data from the EPA's Air Quality System database, which provides hourly data at the pollution-monitor level for pollutants that are regulated by the Clean Air Act. Comprehensive data for PM 2.5 are available beginning in 1999. We focus on PM 2.5, but we also obtain data on four other criteria pollutants: ozone (O_3), carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen dioxide (NO_2).[5] As with PM 2.5, past literature has linked these air pollutants to infant mortality and other adverse health outcomes (Currie and Neidell 2005; Moretti and Neidell 2011; Ward 2015; Schlenker and Walker 2016). We aggregate monitor readings to the daily level by averaging across hourly observations and then construct county-level pollution measures by averaging all available pollution readings on a given day across all monitors located within the county.

Figure 1 displays aggregate trends in PM 2.5 over time. Average concentrations of PM 2.5 have been steadily falling from 13.0 micrograms per cubic meter ($\mu g/m^3$) in 1999 to 9.37 $\mu g/m^3$ in 2011. One unit of PM 2.5 thus represents about 10 percent of the average concentration during our time period. Figure 1 also shows that the number of PM 2.5 monitors has remained fairly constant since 2001. However, the set of monitored counties does change over time, and Grainger et al. (2016) find evidence that counties strategically place their pollution monitors in relatively clean areas. Because our instrumental variables approach exploits variation in pollution that is almost surely independent of monitor placement, our estimates should not be biased by changes in monitor composition. However, for completeness we test the

[5] Lead is also a criteria pollutant and can in principle be transported by the wind. However, there are only about 64,000 county-day level observations for lead between 1999 and 2011, and only 52,000 of these also contain observations of PM 2.5.

robustness of our results to imposing various continuity requirements on the sample of included pollution monitors and obtain very similar estimates (see discussion in section V.C).

B. Atmospheric conditions

Wind speed and wind direction data for the years 1999-2011 are from the North American Regional Reanalysis (NARR) daily reanalysis data published by the National Centers for Environmental Information (NCEI). Reanalysis combines multiple data sources in a systematic way to produce an internally and externally consistent dataset that is usually more detailed than any of its components. NARR incorporates raw data from land-based weather stations, aircraft, satellites, radiosondes (essentially, weather balloons), dropsondes (weather instruments dropped from aircraft), and other meteorological datasets.[6] Wind conditions are reported on a 32 by 32 kilometer grid, and consist of vector pairs, one for the East-West wind direction (u-component) and one for the North-South wind direction (v-component). We first interpolate between grid points in the original dataset to estimate the daily u- and v-components at each pollution monitor, using simple linear interpolation. We then use trigonometry to convert the average u- and v-components into wind direction and wind speed. The wind speed is calculated as $WS = \sqrt{u^2 + v^2}$, where u and v are the county-day-level vectors. To calculate the wind angle, we first calculate $\theta = \frac{180}{\pi} Arctan\left(\frac{|v|}{|u|}\right)$ and then translate θ into a 0-360 scale depending on the signs of u and v. Specifically, given θ, the wind angle, $WINDDIR$, is calculated as follows:

$$WINDDIR = \begin{cases} 270 - \theta & \text{if u} > 0 \text{ and v} > 0 \\ 270 + \theta & \text{if u} > 0 \text{ and v} < 0 \\ 90 + \theta & \text{if u} < 0 \text{ and v} > 0 \\ 90 - \theta & \text{if u} < 0 \text{ and v} < 0 \end{cases}$$

Expressed in this way, $WINDDIR$ indicates the wind direction the wind is blowing *from*, with 0 corresponding to wind blowing from the North and higher angles matching to compass directions in a clockwise fashion. We average the estimated monitor-day-level wind direction and speed to the county-day level.

Finally, we obtain daily temperature and precipitation data from Schlenker and Roberts (2009), who produce a daily weather grid using data from PRISM and weather stations.[7] These data include total daily precipitation, and daily maximum and minimum temperatures for each point on a 2.5 by 2.5 mile grid

---

[6] See https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/north-american-regional-reanalysis-narr for more information and for the NARR dataset itself.

[7] See http://www.prism.oregonstate.edu/ for the original PRISM dataset and http://www.wolfram-schlenker.com/dailyData/dataDescription.pdf for a more detailed description of the daily data.

covering the contiguous United States for the years 1999-2011. To aggregate the gridded data to the county level, we average the daily measures across all grid points in a particular county.

C. Mortality, morbidity, and health care costs

Our data on mortality, morbidity, and health care costs come from Medicare administrative data. We focus on beneficiaries aged 65-100, a sample that includes over 97 percent of U.S. elderly. Dates of death, age, sex, and county of residence are obtained for all beneficiaries from the 1999-2011 Medicare enrollment files. Health care utilization and costs are derived from the Medicare Provider Analysis and Review (MedPAR) File, which reports each inpatient stay in a hospital or skilled nursing facility for any beneficiary enrolled in Original (fee-for-service or FFS) Medicare. MedPAR observations are derived from the accumulation of facility (Medicare Part A) service claims corresponding to that stay, and include the date of admission, length of stay, and total cost of the stay.[8] The cost of these inpatient stays accounts for about 70 percent of all Medicare Part A spending and about 45 percent of all Medicare Parts A, B, and D spending on elderly FFS beneficiaries over 1999-2011. In addition to inpatient stays, we also measure outpatient emergency room visits that do not result in a hospital admission using Medicare outpatient claims. To construct county-level daily measures of hospital utilization and costs, we aggregate hospital visit records based on patients' county of residence and the admission date (for inpatient stays) or date of service (for outpatient emergency room visits). Prior air pollution studies generally focus on subcategories of hospitalization, e.g., emergency room visits for heart attacks, and rarely include medical spending data. Our study therefore provides the most comprehensive dataset of health care utilization and spending to date.

Individual-level indicators for the presence of 27 chronic conditions, such as heart disease, COPD, diabetes, and depression, which we use when estimating life-years lost, are obtained from the Chronic Conditions segment of the Master Beneficiary Summary File. Professional medical coders infer these conditions from detailed claims data, which are only available for beneficiaries enrolled in FFS Medicare. Because it may take some time for a relevant claim to appear in the data, information about chronic conditions will be most reliable for those who have been enrolled in FFS Medicare for multiple years.

Table 1 presents summary statistics for our main estimation sample, which consists of 1,600,846 observations at the county-day level.[9] The mean daily concentration of PM 2.5 in our estimation sample is

---

[8] Our measure of cost is the total allowed charges due to the provider, which includes payments made by Medicare, the beneficiary, or another payer.

[9] Our sample does not encompass the entire U.S. due to limitations in the EPA's pollution monitor coverage: PM 2.5 pollution measures are available for only 902 counties during our sample period (see Figure 4). But because pollution

10.86 micrograms per cubic meter, with a standard deviation of 7.34. There are on average 49,486 Medicare beneficiaries in each county, with close to half of these aged between 65 and 74. Because we focus on the elderly, the 3-day death rate in our sample is fairly high, ranging from 138 per million for those aged 65-69 to nearly 1,200 per million for those aged 85 and over.

We observe hospital spending only for beneficiaries who are enrolled in fee-for-service Medicare (FFS); these make up about 80 percent of the population in our sample.[10] For the life-years lost analysis, we focus on the subset of beneficiaries who have been continuously enrolled in FFS for at least two years (67 percent of the people in our sample) to ensure well-measured chronic conditions, as we described earlier. On average, there are 27,716 such individuals in each county, and their 3-day mortality rate is higher than the overall mortality rate in the Medicare population. There are at least two reasons for this. First, because of the continuous enrollment restriction, individuals in this population are at least 67 years old and thus older than average. Second, conventional wisdom and empirical evidence suggest that the fee-for-service population is generally sicker than the average Medicare beneficiary (McGuire et al. 2011).

Finally, the average 3-day hospital spending for the entire FFS population is about $34 per beneficiary, in nominal terms. About 40 percent of this inpatient spending originates from emergency room (ER) admissions. On average, there are 3,370 hospital admissions per million FFS beneficiaries over any given 3-day period, and 47 percent (1,579) of these admissions are through the ER. There are also many ER visits that do not result in admissions: the overall ER visit rate is 4,159 per million FFS beneficiaries.

IV. Empirical strategy

A. Effects of PM 2.5 on mortality and health care utilization

The key causal relationship we would like to estimate is the effect of short-run fluctuations in fine particulate matter on mortality, health, and health care spending, net of any potentially confounding factors. This relationship can be represented by the following regression equation:

$$Y_{cdmy} = \beta \text{PM2.5}_{cdmy} + X'_{cdmy}\boldsymbol{\gamma} + \alpha_c + \alpha_{sm} + \alpha_{my} + \epsilon_{cdmy}, \tag{1}$$

where the dependent variable is one of several possible outcomes in county c on day d in month m and year y. The parameter of interest is β, the coefficient on daily PM 2.5 levels. We first examine the effect of PM 2.5 on the death rate, measured in deaths per million Medicare beneficiaries. The other outcome

monitors tend to be placed in more populated counties, our main regression estimates still capture about 70 percent of the elderly Medicare population.

[10] The remaining 20 percent are enrolled in Medicare Advantage, a managed care alternative to traditional Medicare.

variables measure health care utilization. We observe all hospitalizations and the inpatient spending associated with them; we also observe all ER visits, but only see spending for ER visits that resulted in hospitalizations. Our main utilization outcomes of interest are total hospital spending and admissions per million beneficiaries. We calculate these measures for all hospital admissions and also for the subset of hospital admissions that originate through the emergency room (ER) admissions. We also estimate the effect of PM 2.5 on the total ER visit rate per million beneficiaries, which includes visits that did not result in a hospital admission. Finally, as a placebo test, we consider non-ER (planned) admissions, which should not be affected by short-run fluctuations in fine particulate matter.

The dependent variable Y_{cdmy} is a 3-day total, based on the day d and the following two days. For example, we estimate the effect of pollution on January 1st on the death rate calculated across January 1-3. This aggregation avoids including very short-run mortality displacement in our estimates. For example, a death that occurs 1-2 days earlier because of an increase in pollution exposure (e.g., on January 1st instead of January 2nd) would not change our 3-day measure. A 3-day measure also allows for lagged effects (e.g., exposure to pollution on January 1 may cause a death on January 2). To ensure that β is not capturing the effects of pollution and weather fluctuations over the following two days, which may be correlated with contemporaneous variation, we include two leads of weather conditions in X'_{cdmy}. Our OLS estimates also include two leads of PM 2.5 concentrations, while our IV estimates include two leads of the instruments.[11] To minimize concerns about autocorrelation, we also control for two lags of PM 2.5 concentrations (OLS) or two lags of the instruments (IV).[12] Our results are robust to the number of lags (Table 11).

The high granularity and comprehensive scope of our data allow us to estimate this regression with multiple sets of high-dimensional fixed effects. We control for weather, geography, time, and seasonality far more flexibly than previous studies have done. Specifically, we generate indicators for daily maximum temperatures falling into one of 17 bins, ranging from -15 degrees Celsius (5°F) or less to 30 degrees Celsius (86°F) or more, with 15 intermediate bins each spanning 3 degrees Celsius (5.4°F). We do the same for minimum temperatures. These variables are captured in X'_{cdmy} in the equation above. For daily precipitation and wind speed, we generate indicators for deciles of these variables. We then generate a set of indicators for all possible interactions of these temperature, precipitation, and wind speed variables and

[11] We do not instrument for lagged PM 2.5 with lagged wind direction to reduce the computational burden of the estimation. However, this choice does not affect our reported estimates.

[12] We do not include lags of the weather variables for computational reasons, as these add about 9,300 regressors per lagged day, as discussed below.

include it in all our regressions.[13] We are able to include all these weather controls without losing instrument power because there is a substantial amount of residual variation in wind direction, even after accounting for other climatic factors. We also estimate a series of alternative specifications to demonstrate that our estimates are robust to less flexible weather controls or omitting weather controls entirely (Table 10). Those results reinforce the assumption that our estimates are not driven by unobserved climatic factors that are correlated with both wind direction and mortality.

Our estimates also include county (α_c), state-by-month (α_{sm}), and month-by-year (α_{my}) fixed effects. The county fixed effects control for underlying differences in health and pollution that vary by geography. State-by-month fixed effects control for potential seasonal correlation between pollution, wind direction, and population health, allowing this correlation to vary by state. Finally, month-by-year fixed effects control flexibly for common time-varying shocks, such as those induced by any Medicare or environmental policy changes during our sample period. As with weather controls, we estimate alternative specifications with varying fixed effects to demonstrate the robustness of our results. We cluster all standard errors at the county level and weight all estimates by the relevant population in cases where the dependent variable is in per capita terms.[14] Our results are robust to different clustering choices, including clustering by pollution monitor group (results available upon request).

OLS estimates of equation (1) are prone to bias because exposure to PM 2.5 is not randomly assigned and is likely to be measured with error. We address this by employing an instrumental variables (IV) strategy, using daily wind direction in the county as an instrument for pollution. Because the effect of wind direction on PM 2.5 levels varies by geography, as illustrated by Figures 1 and 2, we allow the effect of the wind instruments in our first stage to also vary according to geography. The specification for our first stage is:

$$\mathrm{PM2.5}_{cdmy} = \sum_{g=1}^{100}\sum_{b=0}^{2} \beta_b^g 1[G_c = g] \times WINDDIR_{cdmy}^{90b} + X'_{cdmy}\boldsymbol{\sigma} + \alpha_c + \alpha_{sm} + \alpha_{my} + \epsilon_c. \quad (2)$$

The excluded instruments are the variables $1[G_c = g] \times WINDDIR_{cdmy}^{90b}$. Each variable in the set $WINDDIR_{cdmy}^{90b}$ is equal to 1 if the daily average wind direction in county c falls in the 90-degree interval

[13] Thus, we have up to 28,899 ($=17 \times 17 \times 10 \times 10 - 1$) weather indicators included in our regression for each of the three days we control for. In practice, not all possible combinations are realized in the data, so the actual number of included weather controls is about 9,300 per day (i.e., about 27,900 weather indicators per regression). Although wind speed affects local concentrations of pollution, we do not use it as an instrument because wind speed also affects the rate of water evaporation from skin and can thus have direct mortality impacts under some conditions.

[14] For example, if the dependent variable is the elderly mortality rate, then we weight by the number of Medicare beneficiaries in the county at the beginning of the reference day; if the dependent variable is the mortality rate for those 85 and older, then we weight by the number of beneficiaries who are 85 and older.

$[90b, 90b + 90)$ and 0 otherwise. The omitted category is the interval [270,360). The variable $1[G_c = g]$ is an indicator for county c being classified into monitor group g. The coefficient on the interaction between these two variables, β_b^g, is thus allowed to vary across 100 different geographic regions, as explained below. The other control variables (the included instruments) X'_{cdmy} and the fixed effects are defined as in equation (1).

Like other studies in this literature, we face the challenge that a county's pollution monitor readings may not adequately measure the average pollution exposure for county residents due to the sparse placement of monitors within counties. We structured equation (2) to mitigate this problem by forcing the equation to estimate a common effect of county wind direction on pollution for all monitors within each of the 100 geographic areas, all of which span multiple counties. We use the *k*-means cluster algorithm to classify all the pollution monitors in our dataset into 100 spatial groups based on their location.[15] Cluster analysis is a standard tool used to assign observations (in our case, pollution monitors) into a pre-specified number of groups based on their characteristics (in our case, longitude and latitude). The resulting groups are displayed in Figure 4. Intuitively, monitors that are close to each other are more likely to be assigned to the same group than monitors that are far apart. On average, each geographic area (group) contains 21 monitors with PM 2.5 readings and 9 counties. As we discuss later, our results are robust to using more or fewer monitor groups. A detailed analysis of the variation driving the first stage, including a full set of estimates from a related specification, is presented in the Online Appendix.

Restricting the coefficient on wind direction to be the same for all monitors in a group is appealing because it reduces the influence of pollution variation emitted by local sources from our estimates, thereby minimizing measurement error. Locally-produced pollution is likely to have different effects on monitors within a monitor group depending on the relative location of the source and monitor. In contrast, non-local sources that are systematically located to one side or another of the entire monitor group are more likely to have a similar effect on all (or most) monitors in the group. Consequently the non-local sources are more likely to drive the pollution variation captured by equation (2).[16] This is beneficial because pollution from local sources is unlikely to reach all individuals residing within the area encompassed by a monitor group, thereby generating measurement error.

---

[15] If a county has monitors belonging to more than one group, we assign the larger integer group number to the county, which is effectively random assignment.

[16] We cannot test for this directly. However, in order for our approach to pick up emissions from local sources within a county, those sources would have to be located consistently in the same direction away from all the monitors in their cluster group. Because the monitors are fairly dispersed geographically (see Figures 2, 3, and 4), we think this is unlikely to be the case. Furthermore, we have also estimated alternative specifications that include 50 cluster groups instead of 100, so that each group spans an even larger portion of the country. Those estimates are reported in our robustness section and are similar to estimates from our primary specification.

To understand how local pollution transport can introduce measurement error and bias the estimated effect of air pollution on health, consider a power plant with a short smokestack located in the center of a county. Suppose an air pollution monitor is located just to the east of the plant. When the wind blows from the west, the monitor will record high levels of pollution, and when it blows from the east, it will record low levels of pollution. Yet, in either case, pollution exposure increases for only one half of the county; in the other half of the county, pollution exposure actually decreases. On net, there might not be any observable health effects at the county level, leading a researcher who uses such variation to conclude that short-term fluctuations in pollution have no effect on health. More generally, pollution transport that affects measured pollution concentrations at a particular monitor in a manner that is not representative of ambient pollution levels for the average individual residing in the area will generate measurement error and subsequent bias. By contrast, we provide evidence in Section V.C and the Online Appendix that the pollution variation we employ is largely driven by non-local transport and thus should have a uniform effect on the entire county. The Appendix also provides evidence that our first stage variation is not driven by a few monitors located next to large polluters.

Equation (2) restricts the effect of wind direction to be constant within each of the four $WINDDIR$ bins. We employ only four bins because it is computationally burdensome to increase the number of instruments. The specification presented in (2) includes hundreds of instruments and tens of thousands of control variables and fixed effects, and is estimated using over one million observations.[17] The main cost of restricting the number of bins is the loss of potentially useful variation in wind direction. We have investigated the effect of increasing the number of $WINDIR$ bins on our estimates; those results, shown in the robustness section, are very similar to our preferred specification.

The large number of instruments employed in our analysis raises the concern that our two-stage least squares (2SLS) IV estimates may suffer from weak instrument bias. However, as illustrated by Figures 2 and 3, wind direction is a strong predictor of air pollution levels, and this is confirmed by the large first-stage F statistics presented in our tables.[18] Moreover, estimating our IV model using the limited information maximum likelihood (LIML) estimator, which is approximately median-unbiased, yields results similar to

[17] We partial out the fixed effects and controls using the algorithm developed by Correia (2017), which automatically performs the necessary degrees-of-freedom adjustments. The algorithm does not partial out instruments, making it more computationally expensive to increase the number of instruments than the number of control variables.

[18] Our tables present first-stage F statistics that are computed assuming errors are homoskedastic. This means they can be compared to the well-known Stock and Yogo (2005) critical values, which are valid only under homoskedasticity. We have also computed first-stage F statistics assuming serially correlated errors. In every specification we have run, those statistics are significantly larger than the first-stage F statistics computed assuming homoskedastic errors.

2SLS. As a robustness check, we also estimate our model using placebo wind instruments and obtain very small F-statistics.

B. Effect of PM 2.5 on life-years lost

The previous section detailed how we estimate the effect of PM 2.5 on the number of lives lost, as measured by the mortality rate. To monetize the social cost of this mortality, one could multiply the estimated number of lost lives by the value of a statistical life (VSL), which is the approach currently utilized by the EPA. However, economic models of the value of life equate the value of statistical life with the net present value of future *life-years* (Murphy and Topel, 2006). Using the same VSL for all deaths regardless of life-years lost may overstate the economic cost if the individuals who die as a result of pollution exposure are relatively sick and have short life expectancies, a concern that may be particularly relevant for the elderly. In line with this reasoning, an alternative approach commonly used in the empirical literature is to estimate life-years lost (LYL) rather than lives lost, and then to monetize this using the value of a statistical life-year (VSLY) (Cutler 2004; Finkelstein and McKnight 2008; Huh and Reif 2017).

In practice, estimating life-years lost is challenging because counterfactual life expectancy is unobserved. The standard approach used throughout the literature is to approximate counterfactual life expectancy per life lost using population life tables (Gardner and Sanborn 1990; Fontaine et al. 2003; Steenland and Armstrong 2006; CDC 2008; Deschenes and Greenstone 2011; Rapsomaniki et al. 2014) or using estimated changes in cause- and age-specific mortality over time (Finkelstein and McKnight 2008; Huh and Reif 2017). All previous studies that we are aware of use only age, or age and sex, to calculate counterfactual life expectancies. However, this approach will still overstate life-years lost if individuals who are more affected by pollution also have shorter life expectancies than average, conditional on these variables (Deschenes and Greenstone 2011). For example, frail individuals with advanced heart or lung disease may be more susceptible to the adverse effects of air pollution and generally have lower life expectancies than observationally similar individuals who do not have such conditions.

We propose a new methodology that exploits detailed Medicare claims data to estimate life-years lost and that is less prone to bias than previous methods. We first present a framework illustrating why the traditional method of estimating life-years lost is likely to produce upwardly biased estimates. We highlight an additional assumption required to eliminate this bias and explain how our approach is more likely to meet this assumption than prior approaches. We then use our new methodology to estimate the number of life-years saved due to a reduction in fine particulate matter.

Let $c_{it} > 0$ be the expected number of remaining life-years for individual i in period t. Let d_{it} be an indicator equal to 1 if the individual dies from a pollution shock in period t and 0 otherwise. Define $L_{it} \equiv c_{it} d_{it}$ to be the number of life-years lost due to the death of this individual. For simplicity, we assume that exposure to PM 2.5 is assigned randomly and begin with the case where PM 2.5 affects all individuals equally.[19] If c_{it} were observable, then the researcher can estimate the effect of PM 2.5 on the number of life-years lost in period t with the following regression equation:

$$L_{it} = \alpha + \gamma \text{PM2.5}_{it} + e_{it}. \tag{3}$$

The error term e_{it} represents factors other than pollution that affect mortality and, by assumption, is uncorrelated with PM2.5_{it}. Thus, $Cov(\text{PM2.5}_{it}, e_{it}) = 0$, and the identifying assumption necessary for γ to represent the causal effect of PM2.5_{it} on life-years lost is satisfied.

In practice, a researcher only observes whether an individual dies, and must estimate counterfactual life expectancy c_{it}. For example, one could model it as a function of age, which is a strong predictor of remaining life expectancy. Let $\hat{L}_{it} = \hat{c}(Z_{it}^0)d_{it}$ be the prediction of counterfactual life expectancy generated by some model using covariates Z_{it}^0 (e.g., age and sex), and let $u_{it} \equiv \hat{L}_{it} - L_{it}$ describe the measurement error in this estimate, so that $u_{it} > 0$ indicates the model has overestimated a decedent's true counterfactual life expectancy, L_{it}. Although $E[u_{it}] = 0$ by design, this measurement error is non-classical: by definition, $u_{it} = 0$ when $L_{it} = 0$ and is thus correlated with the outcome variable. Now the analog of equation (3), which the researcher can estimate with observable data, is

$$\hat{L}_{it} = \alpha + \gamma \text{PM2.5}_{it} + u_{it} + e_{it}. \tag{4}$$

Equation (4) requires an additional identifying assumption for γ to represent the causal effect of PM2.5_{it} on life-years lost: $Cov(\text{PM2.5}_{it}, u_{it}) = 0$. This assumption is trivially satisfied if PM2.5_{it} is both randomly assigned and has a constant treatment effect. Then, $E[\hat{\gamma}] = \gamma + \frac{Cov(PM2.5_{it}, u_{it} + e_{it})}{Var(PM2.5_{it})} = \gamma$.

However, bias can arise when estimating equation (4) in the presence of heterogeneous treatment effects, such as if pollution exposure is deadlier for those who are sicker, even if PM 2.5 exposure is as good as randomly assigned. To see this, suppose that the effect of PM 2.5 depends on unobserved characteristics, Z_{it}^1, such that the true relationship between L_{it} and PM2.5_{it} is $L_{it} = \alpha + \gamma \text{PM2.5}_{it} + \phi Z_{it}^1 \times \text{PM2.5}_{it} + e_{it}$, where $E[Z_{it}^2] = 0$. Equation (4) can then be written as

$$\hat{L}_{it} = \alpha + \gamma \text{PM2.5}_{it} + u(Z_{it}^0, Z_{it}^1 \times \text{PM2.5}_{it}) + e_{it}, \tag{5}$$

[19] As in other studies, we focus on estimating the *immediate* effects of pollution exposure on life-years lost. It is also possible that exposure reduces an individual's remaining life expectancy, c_{it}, without killing her during the time window we focus on. In that case, our life-years lost estimates can be interpreted as lower bounds.

where we emphasize that the measurement error $u(Z_{it}^0, Z_{it}^1 \times \text{PM2.5}_{it})$ is now a function of both observed characteristics Z_{it}^0 and unobserved characteristics Z_{it}^1. If the identifying assumption $Cov(\text{PM2.5}_{it}, u_{it}) = 0$ continues to hold, then estimation will yield an unbiased estimate of the average treatment effect: $E[\hat{\gamma}] = \gamma + \phi E[Z_{it}^1] = \gamma$. This is the case if Z_{it}^1 is uncorrelated with expected life years remaining, c_{it}, for example. However, in general the identifying assumption $Cov(\text{PM2.5}_{it}, u_{it}) = 0$ is no longer trivial. For example, suppose that the researcher does not include pre-existing heart conditions, z_{it}, when estimating $\hat{L}_{it}$, i.e., $z_{it} \notin Z_{it}^0$. Then the estimation of γ will be biased upwards if people with heart conditions are both more likely to die following exposure to PM 2.5 (i.e., $z_{it} \in Z_{it}^1$) and to have lower life expectancies than average. Because this characteristic is not included in the model of life expectancy, it results in positive measurement error (i.e., $\hat{L}_{it} > L_{it}$) when levels of PM 2.5 are high. Consequently, $E[\hat{\gamma}] = (\gamma + \phi E[Z_{it}^1]) + \frac{Cov(PM2.5_{it}, u_{it})}{Var(PM2.5_{it})} = \gamma + \frac{Cov(PM2.5_{it}, u_{it})}{Var(PM2.5_{it})} > \gamma$.

Equation (5) summarizes a key challenge that researchers face when estimating the effect of pollution exposure on life-years lost. Even if pollution is as good as randomly assigned, estimation of γ is biased upward in the presence of an unobserved factor (e.g., heart disease) that is negatively correlated with remaining life expectancy and positively correlated with the probability of dying following exposure to pollution. This is problematic because populations with low levels of remaining life expectancy, such as the elderly, are often more vulnerable in general and may be more susceptible to dying from pollution exposure than populations with high levels of remaining life expectancy. The individual-level framework above also holds for more aggregated levels of analysis, such as the county level.

We address this challenge by harnessing the comprehensive health and demographic information available in the Medicare dataset to generate relatively precise predictions of counterfactual life expectancy. In other words, we minimize the magnitude of the measurement error represented by u_{it} in equation (5). Note that it is not necessary to eliminate all measurement error to remove the bias in estimating life-years lost; it suffices to eliminate just the portion of the measurement error that is correlated with the heterogeneous treatment effect. To our knowledge, no previous study has addressed this bias by using variables other than age and sex to predict life expectancy (e.g., Deschenes and Greenstone 2011). By contrast, we incorporate information on chronic conditions, medical spending, health care utilization, and geographic location, among others. We show that this matters: using average life expectancy or estimating life expectancy using only basic demographic variables such as age and sex causes significant upward bias in regression estimates of life-years lost due to air pollution.

A challenge with estimating counterfactual life expectancy is that not everybody dies during the period we observe them. We therefore employ the semi-parametric Cox proportional hazards survival model, which assumes that the hazard rate of death for individual i can be factored into two functions:[20]

$$h(t_i|x_i,\beta) = h_0(t_i)\exp[x_i'\beta]$$

The hazard rate at time t_i, $h(t_i|x_i,\beta)$, depends on the baseline hazard rate, $h_0(t_i)$, and on a vector of individual characteristics, x_i. The parameter vector β is estimated by maximizing the log partial likelihood function:

$$\ln L(\beta) = \sum_{i=1}^{N} \delta_i \left[x_i'\beta - \ln \sum_{j\in R(t_i)} \exp[x_j'\beta]\right], \tag{6}$$

where δ_i is equal to one for individuals whose deaths we observe (uncensored observations) and equal to zero otherwise. The risk set $R(t_k) = \{l: t_l \geq t_k\}$ is the set of observations at risk of death at time t_k. We nonparametrically estimate the baseline hazard function, $h_0(t_i)$, following Breslow (1972). See the appendix for details.

We estimate this Cox proportional hazards model using the 2002 cohort of Medicare beneficiaries.[21] We observe all deaths that occur among this cohort between January 1, 2002 and December 31, 2011. During this 10-year time period, 50 percent of the sample dies. To ensure accurate measures of beneficiaries' chronic conditions, we limit the sample to Medicare beneficiaries who as of January 1, 2002 had been continuously enrolled in fee-for-service Medicare for at least two years.[22] For computational ease, we further limit the analysis to a random 5 percent sample of these beneficiaries. The final estimation sample for our survival analysis includes 1,211,585 individuals.

To assess the value of accounting for individual characteristics and using machine learning over simpler approaches, we estimate the survival model several times, using increasingly large sets of characteristics. First, we use no individual characteristics, assuming a homogeneous survival function. A second specification controls for age and sex, and then a third specification additionally controls for the presence of 27 pre-existing chronic conditions. Our final and preferred specification incorporates over one thousand variables, derived from individual-level Medicare data and ZIP code-level data from the American

[20] Employing fully parametric models that assume survival rates are governed by either the Gompertz or Weibull distributions yields very similar results.

[21] Although earlier cohorts are observable, we do not use them because the Medicare variables denoting the presence of pre-existing chronic conditions, which are strong predictors of survival, are nonexistent or unreliable in earlier years.

[22] Because our analysis focuses on beneficiaries who are eligible for Medicare based on their age being 65 or older, this restriction means that no one in this sample is under age 67.

Community Survey. This model includes information on prior medical spending; outpatient and inpatient visits; length of stay for inpatient, skilled nursing facility, and hospice events; number of hospital readmissions; and average commute times, median income, median housing values, and employment in the beneficiary's 5-digit ZIP code of residence (see Appendix for details). To avoid reverse causality, all Medicare variables are based on medical histories from the previous year (2001). Including so many control variables creates two challenges. First, some variables may be significant predictors of survival for the 2002 cohort by chance, even if they are not good predictors of survival in general. This may cause bias due to overfitting (Harrell et al. 1996). Second, computational limitations prevent us from including a large set of regressors when performing conventional maximum likelihood estimation on a large sample using standard numerical procedures.

Recent advances in machine learning techniques help us overcome these challenges and use all 1,062 variables when predicting individual-level life expectancies (Athey and Imbens 2016). One popular method is the Least Absolute Shrinkage and Selection Operator (LASSO) estimator (Tibshirani 1997).[23] LASSO can be implemented by maximizing a penalized version of objective function (6):

$$\ln L(\beta) = \left(\sum_{i=1}^{N} \delta_i \left[x_i'\beta - \ln \sum_{j \in R(t_i)} \exp[x_j'\beta] \right] \right) - \lambda \sum_{i=1}^{k} |\beta_i| \tag{7}$$

where $|\beta_i|$ is the absolute value of β_i (where β_i is element i of the vector β) and k is the number of included regressors. We select the optimal penalty parameter λ using 5-fold cross validation.[24] We then use estimates of β and observable characteristics x_i to predict the life expectancy of each Medicare beneficiary who was continuously enrolled in fee-for-service for at least two years at some point during our sample period.

To integrate these estimates into the county-level empirical strategy in Section IV.A, we simply aggregate life-years lost over all individuals in the county and replace the dependent variable in equation (1) with the estimated daily number of life-years lost per capita in county $c, \hat{L}_{cdmy}$. The variable $\hat{L}_{cdmy}$ is equal to the sum of the estimated counterfactual life expectancies for all decedents divided by the total number of beneficiaries in the county, and thus is analogous to how to we calculate the mortality rate.

We now demonstrate the increase in explanatory power that accompanies the inclusion of additional demographic and health variables in estimating life-years lost. We first identify the Medicare beneficiaries who died between 2001 and 2011. We then use our model to predict their counterfactual life expectancy using increasing numbers of explanatory variables, lagged by one year from their death date.

---

[23] We also used other machine learning techniques like ridge regression and elastic net. The results are similar.

[24] See Simon et al. (2011) for a detailed discussion of the algorithm we employ to implement the Cox proportional hazards estimator with a LASSO penalty term.

The results are shown in Figure 5. The green bar, “Medicare FFS average”, reports the average life expectancy for all Medicare fee-for-service beneficiaries (11.56 years) and serves as a baseline. This value would be an accurate measure of counterfactual life expectancy if Medicare beneficiaries died randomly. However, it is well-known that individuals who die are on average older and sicker than the general population. Intuitively, if a model could perfectly predict life expectancy, then the average predicted life expectancy for this sample of individuals (who all died within one year) should be below 1. However, to the extent that there is an idiosyncratic component to mortality (e.g., some individuals who suffer a heart attack survive purely by chance), even the most complete model may produce estimates above 1.

The rest of Figure 5 consider the performance of three increasingly detailed survival models. The red bar, “Cox (age, sex),” adjusts life expectancy for age and sex and predicts an average life expectancy of 7.85 years. This is similar to the age group analysis performed by Deschenes and Greenstone (2011). To our knowledge, no other study has incorporated information beyond age and sex when accounting for life-years lost. The blue bar, “Cox (age, sex, 27 chronic conditions),” additionally controls for 27 different chronic conditions, reducing predicted life expectancy by over two additional years. This happens because decedents are sicker than the average Medicare beneficiary, even after controlling for age and gender. Finally, the black bar, “Cox (LASSO),” in Figure 5 displays average predicted life expectancy based on a model that incorporates data from all 1,062 variables in our dataset using the LASSO method.[25] This reduces counterfactual life expectancy by yet another half year, to 4.86 years per decedent. The LASSO method prediction provides the smallest, and therefore the most accurate, predicted average life expectancy.

Of course, idiosyncratic error means the LASSO method still does not perfectly predict life expectancy. More generally, we stress that precise predictions, while desirable, are not necessary for identification. As shown by our model, the required identifying assumption is not that the measurement error in counterfactual life expectancy be 0; rather, all that is needed is that this measurement error not be correlated with heterogeneous treatment effects. For example, if air pollution kills people at random, then one does not need to have precise individual-level estimates of life expectancy; the population mean will suffice. The only way to know whether it matters is to see how estimates change when using these different predictions. Those results are presented in the next section.

[25] These 1,062 variables are described in the appendix. Listed in descending order of importance, the fifteen biggest predictors of remaining life expectancy (i.e., those with the largest coefficients) are the twelve indicators for the (non-zero) quantiles of hospice spending, the indicator for lung cancer, the indicator for the highest quantile of Medicare Part B drug spending, and the indicator for the highest quantile of dialysis spending.

V. Results

A. Mortality and health care utilization

Panel A of Table 2 reports OLS estimates of the relationship between daily PM 2.5 and 3-day mortality rates for different age groups. As reported in Column (1), each 1-μg/m^3 increase in daily PM 2.5 exposure is associated with 0.098 additional deaths per million elderly over the following three days, or a 0.025 percent increase relative to the average 3-day mortality rate. Columns (2)-(6) report results estimated separately for each of five age groups. The absolute and relative increases in mortality are non-monotonic across age groups, with those aged 70-79 experiencing lower (and insignificant) increases in death rates than those aged 64-69 despite having higher mean death rates.

Panel B of Table 2 presents the corresponding IV estimates of the causal effect of daily PM 2.5 on 3-day mortality. These are substantially (4.8 – 14 times) larger than the estimates in Panel A, suggesting that OLS estimation suffers from significant bias. The IV estimates imply that each 1-μg/m^3 increase in daily PM 2.5 exposure corresponds to 0.61 additional deaths per million elderly over the following three days, or a 0.15 percent increase relative to the average 3-day mortality rate.[26] The corresponding estimate for a one standard deviation increase in daily PM 2.5 is a 1.1 percent increase in 3-day mortality. Columns (2)-(6) show a largely monotonic relationship between the mortality effect of PM 2.5 and age, with each 1-μg/m^3 increase in daily PM 2.5 causing 0.263 additional deaths per million among the 65-69 population but 2.050 additional deaths per million among the 85 and over population. However, because the average mortality rate is also much higher for the older elderly, the *relative* mortality effects across age groups follow a U-shaped pattern: each 1-μg/m^3 increase in daily PM 2.5 exposure increases 3-day mortality by 0.20 percent among ages 65-69, by 0.10 percent among ages 75-79, and by 0.18 percent among ages 85 and over. This pattern is somewhat unexpected: if sicker individuals are more vulnerable to pollution shocks, and if age is a good proxy for health, then we might expect relative mortality to increase monotonically with age. We return to this point when discussing our estimates of life-years lost due to PM 2.5, where we will find that relative mortality does increase monotonically with counterfactual life expectancy.

Next, we estimate the effect of daily PM 2.5 on 3-day hospitalization rates and associated medical spending per million beneficiaries enrolled in fee-for-service Medicare. As discussed earlier, the change in sample from all Medicare beneficiaries to those enrolled in traditional, fee-for-service Medicare is necessary because spending information is only available for this subsample. For reference, we show the

[26] As described in our empirical strategy, we focus on 3-day mortality in order to avoid capturing short-run mortality displacement and to allow for pollution to have lagged effects. The comparable estimate from IV estimation of (1) for 1-day mortality yields a coefficient of 0.382 additional deaths resulting from a 1-μg/m^3 increase in daily PM 2.5 exposure, suggesting that the mortality impact of PM 2.5 exposure grows over time due to lagged effects.

all-age mortality response to PM 2.5 for this population in Column (1) of Table 3; it is very similar to what we find for the overall Medicare population. Panel A of Table 3 shows that the association between PM 2.5, hospitalization, and medical spending is mixed: each 1-$\mu g/m^3$ increase in daily PM 2.5 exposure is associated with significantly *less* inpatient spending and *fewer* hospital admissions, is not associated with spending on ER admissions, and is associated with significantly *more* ER admissions and visits.

A more consistent story emerges from our IV approach (Panel B), which shows that increases in daily PM 2.5 increase both hospitalizations and inpatient spending, driven primarily by encounters that originate in the ER. The IV estimates imply that each 1-$\mu g/m^3$ increase in daily PM 2.5 causes a highly significant increase in ER inpatient spending of over $15 thousand per million beneficiaries (relative to a mean of $13.7 million). This increase is almost as large as the increase in total inpatient spending, and we cannot reject that the latter is driven entirely by increases in ER spending.[27] The overall admissions rate increases by 2.03 per million beneficiaries, an increase which also can be almost entirely explained by the 1.96 additional admissions originating through the ER. We also estimate that PM 2.5 increases total ER visits, including visits that do not result in a hospital admission, by 2.29 per million beneficiaries. Finally, we consider the non-ER admissions rate, which largely consists of planned hospitalizations and thus serves as a natural placebo test. We do not find any significant effects for this category, and the point estimate is very small, further lending credence to our identification strategy.

Comparing the OLS estimates to the IV estimates in Tables 2 and 3 provides strong evidence that observational studies of the relationship between air pollution and health outcomes suffer from significant bias: virtually all our OLS estimates are smaller than the corresponding IV estimates. If the only source of bias were classical measurement error, which causes attenuation, we would not expect to see significantly *negative* OLS estimates. Thus, other biases, such as changes in economic activity that are correlated with both hospitalization patterns and pollution, appear to be a concern even when working with high-frequency daily data.

It may also be of interest to compare the magnitudes of our IV estimates to those from the epidemiological literature, which are often used to calculate the benefits of various environmental policies. However, while there are many epidemiology papers estimating the health effects of acute pollution exposure, few of them focus on the effect of PM 2.5 on the elderly. Furthermore, studies that do estimate the health effects of acute PM 2.5 exposure often focus on specific causes of death or hospitalization, which

[27] Our results are similar if we consider the natural logs of these spending variables.

makes a direct comparison difficult.[28] We are also not aware of any study that uses 3-day mortality as the outcome to capture both short-term harvesting and delayed effects.

While we prefer the 3-day mortality measures, we have also estimated the effect of PM 2.5 on 1-day mortality and hospitalizations (see Table A1 in the Online Appendix) to facilitate comparison to two studies from the epidemiological literature with settings similar to ours. Using data from 27 large US cities from 1997-2002, Franklin et al. (2007) report that a 10 $\mu g/m^3$ increase in daily PM 2.5 exposure increases all-cause mortality for those aged 75 and above by 1.66 percent. Our 1-day IV estimate for 75+ year olds (column 1 of Table A1) is an increase of 2.73 percent (an additional 5.6 deaths per million beneficiaries for a 10 $\mu g/m^3$ increase in daily PM 2.5), which is over 60 percent larger. Allowing for short-run delayed effects and harvesting by considering 3-day mortality for this age group (column 2 of Table A1) further raises our estimated effect of PM 2.5 by over 65 percent (9.3 deaths per million). On the hospitalization side, Dominici et al. (2006) use Medicare claims data from US urban counties from 1999-2002 and find an increase in elderly hospitalization rates associated with a 10 $\mu g/m^3$ increase in daily PM 2.5 exposure ranging from 0.44 percent (for ischemic heart disease hospitalizations) to 1.28 percent (for heart failure hospitalizations). Our estimated increase in 1-day all-cause hospitalizations from a 10 $\mu g/m^3$ increase in daily PM 2.5 is 1.92 percent (column 3 of Table A1), which is 50 percent larger than the heart failure estimate and over four times larger than the ischemic heart disease estimate. As with mortality, allowing for short-run delayed effects and harvesting further raises our estimated hospitalization effect. Because our hospitalization rate includes causes almost surely unrelated to pollution (e.g., hip fractures), restricting our sample of hospitalizations to those plausibly affected by PM 2.5 would likely magnify this difference further. While we can only make an apples-to-apples comparison of our results to these two comparable studies, the comparison suggests that the observational studies common in the epidemiological literature may be underestimating the effect of acute pollution exposure on mortality and health outcomes.

B. Life-years lost and the value of mortality reductions

We next turn to estimating the mortality cost of acute PM 2.5 exposure, using the method of estimating life-years lost described in Section IV.B. Table 4 displays estimates of equation (1) when the outcome variable is the estimated 3-day life-years lost per million beneficiaries ($\hat{L}_{cdmy}$). As discussed in Section III, this estimation sample is limited to those beneficiaries continuously enrolled for at least two years in fee-for-service (FFS) Medicare so that chronic conditions are well-measured. For reference,

[28] For example, Zanobetti and Schwartz (2006) estimate the effect of short-run fluctuations in PM 2.5 on ER hospitalizations for myocardial infarction and pneumonia only. See Bentayeb et al. (2012) for a review of the epidemiological literature on the effects of air pollution on elderly health.

Column (1) shows the estimated effect of PM 2.5 on the 3-day mortality rate among the 2-year FFS population. This estimate is slightly larger in absolute terms than the IV estimate from Table 2 in part because the 2-year FFS restriction mechanically excludes individuals aged 65-66, causing this sample to be older on average. The effects relative to average 3-day mortality are very similar for both populations.

Column (2) displays results when every decedent's counterfactual life expectancy is set equal to the mean for the 2-year FFS population (11.6 years). This estimate implies that each 1-$\mu g/m^3$ increase in daily PM 2.5 increases life-years lost by 8.6 years per million beneficiaries. This same effect can also be obtained directly by multiplying the mortality effect of 0.746 in Column (1) by the mean life expectancy of 11.6. However, this estimate is accurate only if beneficiaries killed by PM 2.5 are representative of the overall 2-year FFS population. If decedents have a lower counterfactual life expectancy than those who remain alive, then the estimate in Column (2) will be biased upward.

Columns (3)-(5) of Table 4 illustrate this bias by progressively expanding the set of covariates used to predict counterfactual life expectancy. Those covariates are reported in the column headers. Column (3) displays estimates when counterfactual life expectancy is modeled solely as a function of age and sex. This approach is comparable to studies that estimate age- and sex-specific mortality effects and multiply them by the corresponding life expectancies from population life tables (e.g., Deschenes and Greenstone 2011). In our setting, accounting for decedents' age and sex reduces the estimated impact of PM 2.5 on life-years lost by 31 percent, to 5.9 life-years per million beneficiaries. This decrease is consistent with the results presented in Table 2: older beneficiaries, who have lower life expectancies, are also more likely to be killed by PM 2.5. The estimate decreases by another 40 percent when the counterfactual life-years estimates account for previously diagnosed chronic conditions (Column 4), implying significant heterogeneity in the mortality effect of PM 2.5 even among individuals of the same age and sex.

Finally, we estimate counterfactual life expectancy using the LASSO machine learning algorithm, optimally incorporating over 1,000 additional predictors, as described earlier. This estimate, reported in Column (5), is 24 percent smaller than the estimate that accounts for age, sex, and 27 chronic conditions and implies that each 1-$\mu g/m^3$ increase in daily PM 2.5 increases life-years lost by 2.7 years per million beneficiaries. The richness of our controls suggests this final estimate is about as good an approximation to the true value as can be obtained empirically. Importantly, Table 4 demonstrates that common methods for calculating life-years lost that control only for basic demographic characteristics significantly overstate the effect of pollution on life-years lost.

The sequential reductions in the estimate of life-years lost reported in Table 4 occur for two reasons. First, better survival models should predict lower remaining life expectancy for decedents *on average*. Table 4 (and Figure 5) report that the mean life-years lost per decedent ("LYL per decedent") decreases from 11.56 in the model with no predictors to 4.86 in the LASSO model. Second, a better survival model should also predict a more accurate *distribution* of predicted life expectancies among decedents. This matters if air pollution selectively kills individuals in this population who are systematically healthier (or sicker) than the average decedent. Indeed, Table 4 demonstrates that this second channel also plays a role in reducing the estimated life-years lost from improved survival modeling. While the average LYL per decedent decreases by only 0.43 per million when moving from LYL estimates based on age, sex, and chronic conditions to those based on the LASSO model, the estimated effect of PM 2.5 on LYL drops by nearly twice as much (0.85 per million). This indicates that the mortality effects of PM 2.5 tend to be larger among individuals with characteristics that LASSO associates with lower life expectancy, even after conditioning on age, sex, and chronic conditions.

The estimates in Table 4 can also be used to describe the estimated counterfactual life-years lost among "compliers": those individuals who died because of increases in wind-driven PM 2.5. Comparing this estimate to the average life-years lost among all decedents can shed light on whether those dying from increased pollution are differentially healthy or frail compared to those who die on a typical day. The LYL per complier is calculated by dividing the estimated effect of increased PM 2.5 on life-years lost by the estimated mortality effect (the coefficient reported in Column 1).[29] When life expectancy is modeled as a function of age and sex alone, those dying from pollution appear to have slightly longer life expectancies (7.9 years) compared to the average decedent (7.8 years). However, estimates that rely on chronic conditions or the LASSO model show the opposite pattern. In Column (5), those dying from pollution appear to have somewhat shorter life expectancies (3.6 years) compared to the average decedent (4.9 years).

Pursuing this point further, we next estimate the effects of PM 2.5 on mortality and life-years lost separately for groups of beneficiaries with different life expectancies. Because the results from Table 4 showed that the typical complier has a low life expectancy, we focus this exercise on individuals in the bottom half of the life expectancy distribution (i.e., less than 10 years), although we show estimates for those in the top half of the distribution as well. This analysis sheds light on whether it is only the very sick who are killed by air pollution or whether people in fair health are vulnerable as well.

Table 5 reports mortality rate and life-years lost estimates separately for five groups of beneficiaries: those with a predicted life expectancy of less than 1 year, 1-2 years, 2-5 years, 5-10 years, and over 10 years. The column headers report the percent of all beneficiaries falling into each group: 55 percent of our sample has a life expectancy that exceeds 10 years, while only 0.7 percent has a life expectancy of less than 1 year. Panel A of Table 5 illustrates that the mortality rate effect of PM 2.5 decreases monotonically with life expectancy. A 1-$\mu g/m^3$ increase in daily PM 2.5 increases deaths among

[29] For example, in Column (3), a one-unit pollution increase causes 0.746 deaths per million and causes an increase in LYL of 5.925 years per million. Thus, the LYL per person killed by pollution is 5.925/0.746 = 7.94.

those with life expectancy of less than one year by 18.9 per million. By contrast, the effect on those with life expectancies of 5-10 years is only 0.53 deaths per million, and the mortality rate effect for those with life expectancies exceeding 10 years is a small and insignificant 0.04. This pattern parallels the estimates reported in Table 2, which showed that the mortality effect is largest among the oldest beneficiaries (who generally have low life expectancies). However, unlike the pattern in Table 2, we find that *relative* mortality also decreases monotonically with life expectancy, which is consistent with the notion that the sickest individuals are most vulnerable to pollution shocks.

Table 5 shows that, although beneficiaries with a life expectancy of less than one year are the most likely to be killed by air pollution, beneficiaries with a life expectancy of up to 10 years are also vulnerable. Beneficiaries with a life expectancy of less than one year make up less than 1 percent of our sample, while those with life expectancies of 5-10 years make up almost 30 percent. Therefore, the absolute number of deaths caused by PM 2.5 varies less across these two groups than does the relative number of deaths.

Panel B shows the effects of PM 2.5 on life-years lost in each of these five groups. Although beneficiaries in Column (1) have less than one year of life expectancy, their high mortality rate causes their number of life-years lost due to pollution to exceed that of any other group: 11.3 life years per million beneficiaries. Thus, even among the group where the "harvesting critique" is most likely to apply, there are still significant benefits to reducing pollution on a per-capita basis. By contrast, among beneficiaries with a life expectancy of 5-10 years (Column 4), the life-years lost from pollution is only equal to 3.7. Although their life expectancy is high relative to those in Column (1), their mortality rate is much lower, resulting in a smaller loss of life years. Those with 1-2 or 2-5 years of life expectancy (Columns 2 and 3) fall somewhere in between, losing 8.2 and 6.7 life years per million beneficiaries, respectively, when PM 2.5 increases by 1 $\mu g/m^3$.

Weighting the life-years lost coefficients from Table 5 by the respective sizes of the groups, we see that the largest portion of the social cost of pollution in terms of life-years lost is borne by those with a life expectancy of 5-10 years (30 percent of sample, 43 percent of burden), followed by those with a life expectancy of 2-5 years (12.7 percent of sample, 33 percent of burden). While the per capita burden is highest for those with the lowest life expectancy, the majority of the *aggregate* social burden falls on those with intermediate life expectancy (2 to 10 additional years).

We also note that our approach, which involves predicting life expectancy, identifies vulnerable populations better than an approach that uses age alone. In Table 2, we found significant mortality effects of pollution across all age groups, but there was no clear relationship between age and the relative mortality effect. In contrast, our life expectancy based approach has identified a group of beneficiaries (i.e., those with life expectancies of more than 10 years) who do not appear to be vulnerable to pollution shocks (i.e., the mortality impact is a precisely estimated zero). In addition, the health-vulnerability gradient is much stronger in both absolute and relative terms when health is measured by life-years remaining rather than life-years lived (i.e., age). Overall, this suggests that a precise measure of life expectancy may be useful not only for characterizing mortality costs, but also in identifying which populations are particularly vulnerable to those shocks.

A simple back-of-the-envelope numerical exercise helps to illustrate the policy implications of our results. The average level of PM 2.5 decreased by 3.65-$\mu g/m^3$ nationwide between 1999 and 2011, as shown in Figure 1. The estimate reported in Column (5) of Table 4 implies that such a decrease saved 147,098 life-years annually among the 41 million Medicare beneficiaries alive in 2011.[30] If we assign each life year a standard value of $100,000 each (Cutler, 2004), the mortality reduction benefits of this decrease added up to about $15 billion in 2011. The EPA's calculation of the annual costs of meeting the 1990 Clean Air Act Amendment air quality standard increased from $19.9 billion to $43.9 billion between 2000 and 2010 (EPA 2011). Thus, the estimated $15 billion in annual mortality benefits represents a large fraction of the estimated annual costs of complying with air pollution standards during this period. By contrast, the reduction in hospitalization costs implied by our estimates is an order of magnitude smaller – about $0.93 billion annually.

Our estimate of the mortality reduction benefits is nearly 70 percent lower than the estimate of $47 billion obtained from ignoring heterogeneity in the effect of pollution on elderly mortality. Estimated benefits that account for age and sex are $32 billion, still more than double the $15 billion estimate based on our most comprehensive model. This contrast demonstrates the importance of properly accounting for counterfactual life expectancy when calculating the mortality benefits of reductions in air pollution.

C. Other pollutants and robustness checks

One concern with interpreting our estimates as the causal effects of PM 2.5 is that other air pollutants like ozone (O_3) and carbon monoxide (CO) can be co-transported with fine particulate matter (PM 2.5). However, because these pollutants are not perfectly co-transported (they can be produced by sources located in different places and are carried differently by the wind), our empirical strategy allows us to instrument separately for each pollutant. Two other pollutants, sulfur dioxide (SO_2) and nitrogen dioxide (NO_2) are precursors to PM 2.5 and are also thought to have independent health effects. SO_2 converts to SO_4^{2-}, an important component of particulate matter, on the order of several percent per hour (Luria et al. 2001). NO_2 converts to particulate nitrate at a similar rate (Lin and Cheng 2007). Recall that we are

---

[30] The exact calculation is $2.693 \times 365 \times 41 \times 3.65$. This calculation assumes that our daily mortality effects can be linearly scaled to the annual level. The epidemiological literature generally finds larger effects from long-run exposure than from short-run exposure (Pope and Dockery 1999), suggesting that linear scaling is a conservative assumption.

considering the effect of a 1-day change in average pollution concentrations on 3-day outcomes. Because the majority of SO_2 and NO_2 converts to particulate matter within 2-3 days, it is impossible to distinguish their effects from those of PM 2.5 with a 3-day specification.[31] We therefore focus our investigation on whether our previous estimates change after controlling for CO and O_3, and provide separate evidence in the Online Appendix that it is not driven by SO_2 or NO_2 by instrumenting for all five pollutants simultaneously in the context of 1-day mortality (see Table A2).

We restrict the sample to county-days where readings for CO, O_3, and PM 2.5 are simultaneously available, and then sequentially add the endogenous variables CO and O_3 to our main estimating equation. The results are shown in Table 6. The estimated effects of PM 2.5 are always significant and fairly stable across the different specifications. This suggests that the mortality effects we found are indeed primarily attributable to PM 2.5 and not these other pollutants. The coefficient on ozone is negative in column (3), reflecting the well-known finding that it is negatively correlated with other pollutants that affect mortality (Currie and Neidell 2005), such as carbon monoxide. When we also add carbon monoxide as a control (column 4), we get slightly different O_3 and CO results depending on whether we consider the entire Medicare population or just fee-for-service beneficiaries (Panel A versus B). Nevertheless, our conclusions about the impacts of PM 2.5 are the same for both populations.

Our main empirical specification employs 300 instruments. Although our reported F-statistics are generally large, and our IV and OLS estimates are quite different, we nevertheless undertake two sets of robustness exercises to ensure that our estimates do not suffer from weak instrument bias. First, we estimate our IV model using LIML, which is approximately median unbiased even in the presence of weak instruments. Those estimates, presented in Table 7, are very similar to the 2SLS estimates presented in Table 2. Second, we conduct a placebo exercise where we generate a set of random wind directions and use those in our first stage instead of the actual wind direction. Those results, shown in Table 8, are largely insignificant. Moreover, the first-stage F-statistics for those estimates are very small, which provides strong evidence that our wind direction instrument is picking up meaningful rather than spurious variation in PM 2.5 levels.

To minimize measurement error, our empirical strategy employs variation in pollution that affects all pollution monitors within a geographic area in the same way, by restricting the coefficient on wind direction to be the same for all monitors in the same group. We argued that this approach filters out local

[31] For example, a conservative conversion rate of 3 percent per hour implies that over half (three-quarters) of the SO_2 and NO_2 would have converted to particulate matter after 24 (48) hours. A 4 percent hourly conversion rate implies that over 60 (85) percent is converted after 24 (48) hours. While it is true that some of the SO_2 and NO_2 is removed from the county by wind currents before conversion, some amount (depending on wind speeds and other climatic factors) will remain in the county.

variation in pollution, leaving primarily variation in pollution that is transported into a county from other regions. If this is the case, our wind direction instrument should be weakest on days with low wind speeds. By contrast, if our approach primarily yielded variation in local pollution emissions, our instrument would be *strongest* on days with low wind speeds, when pollution does not travel far from where it was emitted. Thus, we can test the validity of our approach by estimating the first stage separately by deciles of daily wind speed. Those results are reported in Figure 6. Because each regression uses only 10 percent of our sample, the F-statistics are lower than the ones reported in our main specification. Nevertheless, it is clear that the F-statistics are largest for samples that employ days with high wind speeds. This provides strong evidence that the variation we employ comes primarily from pollution that is transported into counties by the wind rather than generated locally.

Table 9 shows that our estimates are not sensitive to how we parameterize our wind instruments. Column (1) decreases the size of the wind angle bins from 90 to 60 degrees. Column (2) reduces the number of monitor groups from 100 to 50, thereby increasing the geographic area covered by each group and making it more difficult for local (within-group) variation to drive our first-stage results. Column (3) increases the number of groups to 200, which increases the flexibility of the first stage but also risks the introduction of local pollution transport into the first stage. All of these estimates are very similar to our preferred specification. This provides strong evidence that our first stage does not suffer from measurement error as a result of picking up local pollution transport. The Online Appendix presents further evidence that we are primarily capturing non-local pollution transport, by showing that the pollution-wind direction relationship does not vary substantially within a monitor group and that neighboring monitor groups exhibit very similar pollution-wind direction relationships.

Tables 10 and 11 show the robustness of our primary empirical specification to including more or fewer instrument lags and to including different types of fixed effects and weather controls. Table 10 demonstrates that our results are robust to varying the weather controls, suggesting that unobserved interactions between wind direction, weather, and mortality are not driving our results. It also illustrates the invariance of our estimates to more or less stringent fixed effects, providing evidence that our results cannot be explained by seasonal or regional patterns. Table 11 demonstrates that our estimates are not driven by lagged effects from pollution on preceding days and thus can be properly interpreted as the effect of a 1-unit change in daily pollution levels.

Our baseline mortality specification employs a 3-day window. The effect of employing a longer time window on our mortality estimates is theoretically ambiguous. On the one hand, a longer time window can increase estimates if PM 2.5 reduces people's life expectancy but does not kill them quickly. For example, if acute exposure reduces an individual's life expectancy to five days, that would not be reflected in our 3-day window specification, but could be captured by a 5-day specification. On the other hand, a longer time window will reduce estimates if the 3-day specification suffers from short-term mortality displacement ("harvesting"). We investigate these different possibilities by estimating how a one-day increase in fine particulate matter affects mortality over the next 5, 7, 10, and 14 days (see Online Appendix Figure A7), controlling for the appropriate number of weather and instrument leads. Overall, we find that our estimate *increases* with the length of the time window, which suggests that short-term mortality displacement is not a concern in our setting. This is also consistent with our life-years lost analysis, which found that many of those killed by air pollution had medium rather than short predicted life expectancies.

Finally, because we are instrumenting for PM 2.5 with wind direction, our estimates should not be affected by monitor entry and exit, as these are plausibly unrelated to daily wind direction. Nonetheless, we probe the robustness of our results to imposing various continuity requirements on the sample of included monitors. About a quarter of PM 2.5 monitors in our sample are online during our entire 13-year sample period; the median number of years online is eight. Most monitors do not operate every day, as EPA does not require it.[32] As a robustness check, we restrict our sample of monitors to (a) those that exist for at least five years of our sample period; (b) those that exist for at least seven years; (c) those that are online for at least 120 days per year, on average, regardless of how many years they are online; or (d) those that are online for at least 240 days per year. Our estimates (available upon request) vary little under each of these four restrictions.

VI. Conclusion

Understanding how air pollution affects health and health care spending is essential for crafting efficient environment policy, such as Pigouvian pricing based on health externalities, but endogeneity and measurement error make it challenging to identify the causal effects of pollution. Moreover, it is also difficult to estimate the mortality cost of air pollution because the mortality effects are heterogeneous. If deaths caused by pollution occur disproportionately among the least healthy, then ignoring this heterogeneity will lead to upward bias when estimating the social cost of pollution.

Our paper sheds light on these issues by estimating the causal effect of acute fine particulate matter exposure on mortality, life-years lost, and hospitalizations using a novel identification strategy based on changes in wind direction. This is accomplished by linking daily pollution and climatic variables to detailed administrative records on all Medicare beneficiaries from 1999-2011. We find significant effects of pollution on mortality, health care spending, and hospitalizations.

[32] For example, many monitors operate according to EPA's 3-day or 6-day schedule, as listed here: https://www3.epa.gov/ttn/amtic/calendar.html.

Our life-years lost analysis shows that the least healthy individuals are more vulnerable to pollution than the average elderly person, whether vulnerability is measured by total mortality risk, relative mortality risk, or expected number of life-years lost. Our estimate of the total number of life-years lost, which is based on health information gleaned from detailed Medicare data, is less than half the magnitude of an estimate that accounts for age and sex alone, and less than a third of the magnitude of an estimate based solely on the life expectancy of an average Medicare beneficiary. This suggests that failure to adjust for the health of those who die can result in substantial overvaluation of the mortality benefits of pollution reduction. Nonetheless, we calculate that the reduction in PM 2.5 experienced nationwide between 1999 and 2011 generated $15 billion annually in mortality benefits among the elderly alone by the end of that period. Finally, the methodology we develop to estimate life-years lost in the context of air pollution is very general and could be used to estimate mortality costs or identify vulnerable populations in other studies that estimate health effects.

REFERENCES

Anderson, Michael (2015). "As the Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality." NBER Working Paper #21578.

Athey, Susan, and Guido Imbens (2016). "The State of Applied Econometrics-Causality and Policy Evaluation." arXiv preprint arXiv:1607.00699.

Bentayeb, M., M. Simoni, N. Baiz, D. Norback, S. Baldacci, S. Maio, G. Viegi, I. Annesi-Maesano, and Geriatric Study in Europe on Health Effects of Air Quality in Nursing Homes (GERIE) Group (2012). "Adverse respiratory effects of outdoor air pollution in the elderly [Review article]." *The International Journal of Tuberculosis and Lung Disease* 16 (9): 1149-1161.

Breslow, Norman E (1972). "Discussion of Professor Cox's paper." *Journal of the Royal Statistical Society B*, 34: 216-217.

Brook, Robert, Bruce Urch, J. Timothy Dvonch, Robert Bard, Mary Speck, Gerald Keeler, Masako Morishita et al (2009). "Insights into the Mechanisms and Mediators of the Effects of Air Pollution Exposure on Blood Pressure and Vascular Function in Healthy Humans." *Hypertension*, 54(3): 659-667.

Card, David, Carlos Dobkin, and Nicole Maestas (2009). "Does Medicare Save Lives?" *The Quarterly Journal of Economics*, 124(2): 597-636.

Centers for Disease Control and Prevention (CDC) (2008). "Smoking-attributable mortality, years of potential life lost, and productivity losses--United States, 2000-2004." *Morbidity and Mortality Weekly Report* 57(45): 1226.

Chay, Kenneth, Carlos Dobkin, and Michael Greenstone (2003). "The Clean Air Act of 1970 and Adult Mortality." *Journal of Risk and Uncertainty* 27(3): 279-300.

Chay, Kenneth, and Michael Greenstone (2003). "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *The Quarterly Journal of Economics*, 118(3): 1121-1167.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li (2013). "Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences*, 110(32): 12936-12941.

Correia, Sergio (2017). "Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator." Working Paper. http://scorreia.com/research/hdfe.pdf

Currie, Janet, and Matthew Neidell (2005). "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *The Quarterly Journal of Economics*, 120(3): 1003-1030.

Currie, Janet, Matthew Neidell, and Johannes Schmieder (2009). "Air Pollution and Infant Health: Lessons from New Jersey." *Journal of Health Economics*, 28(3): 688-703.

Cutler, David (2004). *Your Money or Your Life: Strong Medicine for America's Health Care System*. New York: Oxford University Press.

Deryugina, Tatyana, Julian Reif, Nicole Riemer, and Matthew West (2016). "The Wind's Fatal Blow: the Effect of Transported Air Pollution on Infant and Adult Mortality." mimeo.

Deschenes, Olivier, Michael Greenstone, and Joseph Shapiro (2016). "Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program." Mimeo.

Deschenes, Olivier, and Michael Greenstone (2011). "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US." *American Economic Journal: Applied Economics*, 3(4): 152-85.

Dockery, Douglas, and C. Arden Pope (1994). "Acute Respiratory Effects of Particulate Air Pollution." *Annual Review of Public Health*, 15(1): 107-132.

Dockery, Douglas, C. Arden Pope, Xiping Xu, John Spengler, James Ware, Martha Fay, Benjamin Ferris Jr, and Frank Speizer (1993). "An Association Between Air Pollution and Mortality in Six U.S. Cities." *New England Journal of Medicine*, 329(24): 1753-1759.

Dominici, Francesca, Michael Greenstone, and Cass Sunstein (2014). "Particulate Matter Matters." *Science*, 344(6181): 257.

Dominici, Francesca, Roger D. Peng, Michelle L. Bell, Luu Pham, Aidan McDermott, Scott L. Zeger, and Jonathan M. Samet (2006). "Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases." *JAMA* 295 (10): 1127-1134.

EPA (2004). "The Particle Pollution Report. Current Understanding of Air Quality and Emissions through 2003." US Environmental Protection Agency Washington, DC.

EPA (2009). "Integrated science assessment for particulate matter." US Environmental Protection Agency Washington, DC.

EPA (2011). "The Benefits and Costs of the Clean Air Act from 1990-2020." US Environmental Protection Agency Washington, DC.

Fontaine, Kevin R., David T. Redden, Chenxi Wang, Andrew O. Westfall, and David B. Allison (2003). "Years of life lost due to obesity." *JAMA* 289(2): 187-193.

Finkelstein, Amy and Robin McKnight (2008). "What Did Medicare Do? The Initial Impact of Medicare on Mortality and Out of Pocket Medical Spending." *Journal of Public Economics*, 92(7): 1644-1669.

Franklin, Meredith, Ariana Zeka, and Joel Schwartz (2007). "Association between PM2. 5 and all-cause and specific-cause mortality in 27 US communities." *Journal of Exposure Science and Environmental Epidemiology* 17(3): 279-287.

Gardner, John W., and Jill S. Sanborn. "Years of Potential Life Lost (YPLL)-What Does it Measure?." Epidemiology 1, no. 4 (1990): 322-329.

Grainger, Corbett, Andrew Schreiber, and Wonjun Chang (2016). "How States Comply with Federal Regulations: Strategic Ambient Pollution Monitoring." mimeo

Harrell, Fran, Kerry Lee, and Daniel Mark (1996). "Tutorial in Biostatistics Multivariable Prognostic Models: Issues in Developing Models, Evaluating Assumptions and Adequacy, and Measuring and Reducing Errors." *Statistics in medicine*, 15: 361-387.

Huh, Jason, and Julian Reif (2017). "Did Medicare Part D Reduce Mortality?" forthcoming, *Journal of Health Economics*.

Knittel, Christopher, Douglas Miller, and Nicholas Sanders (2016). "Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health." *Review of Economics and Statistics*, 98(2): 350-366.

Kundu, Shuvashish, and Elizabeth Stone (2014). "Composition and Sources of Fine Particulate Matter across Urban and Rural Sites in the Midwestern United States." *Environmental Science: Processes & Impacts*, 16(6): 1360-1370.

Laden, Francine, Lucas Neas, Douglas Dockery, and Joel Schwartz (2000). "Association of Fine Particulate Matter From Different Sources With Daily Mortality in Six U.S. Cities." *Environmental Health Perspectives*, 108(10): 941-947.

Langrish, Jeremy et al (2013). “Altered Nitric Oxide Bioavailability Contributes to Diesel Exhaust Inhalation-Induced Cardiovascular Dysfunction in Man.” *Journal of the American Heart Association*, 2(1).

Lin, Y-C. and M-T. Cheng (2007). “Evaluation of formation rates of NO_2 to gaseous and particulate nitrate in the urban atmosphere.” *Atmospheric Environment*, 41(9): 1903-1910.

Luria, Menachem, Robert Imhoff, Ralph Valente, William Parkhurst, and Roger Tanner (2001). “Rates of Conversion of Sulfur Dioxide to Sulfate in a Scrubbed Power Plant Plume.” *Journal of the Air & Waste Management Association,* 51(10): 1408-1413.

McGuire, Thomas, Joseph Newhouse, and Anna Sinaiko (2011). “An Economic History of Medicare Part C.” *Milbank Quarterly*, 89(2): 289-332.

Moretti, Enrico, and Matthew Neidell (2011). “Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles.” *Journal of Human Resources*, 46(1): 154-175.

Muller, Nicholas Z., and Robert Mendelsohn (2007). "Measuring the damages of air pollution in the United States." *Journal of Environmental Economics and Management*, 54(1): 1-14.

Murphy, Kevin M., and Robert H. Topel (2006). “The Value of Health and Longevity.” *Journal of Political Economy* 114(5): 871-904.

Office of Management and Budget (2012). “2012 Report to Congress on the Benefits and Costs of Federal Regulation and Unfunded Mandates on State, Local, and Tribal Entities” (OMB, Washington, DC); www.whitehouse.gov/sites/default/files/omb/inforeg/2012_cb/2012_cost_benefit_report.pdf.

Pope 3rd, C. A (2000). “Epidemiology of fine particulate air pollution and human health: biologic mechanisms and who's at risk?.” *Environmental Health Perspectives* 108(4): 713.

Pope, C. Arden., and Douglas Dockery (1999). “Epidemiology of Particle Effects.” *Air pollution and health*, 31: 673-705.

Pope, C. Arden, and Douglas Dockery (2006). “Health Effects of Fine Particulate Air Pollution: Lines that Connect.” *Journal of the Air & Waste Management Association,* 56(6): 709-742.

Pope, C. Arden, Michael Thun, M M Namboodiri, Douglas Dockery, John Evans, Frank Speizer, and Clark Heath Jr (1995). “Particulate Air Pollution as a Predictor of Mortality in a Prospective Study of U.S. Adults.” *American Journal of Respiratory and Critical Care Medicine*, 151(3): 669-674.

Pope, C. Arden, Richard Burnett, Michael Thun, Eugenia Calle, Daniel Krewski, Kazuhiko Ito, and George Thurston (2002). “Lung Cancer, Cardiopulmonary Mortality, and Long-Term Exposure to Fine Particulate Air Pollution.” *JAMA*, 287(9): 1132-1141.

Samet, Jonathan, Francesca Dominici, Frank Curriero, Ivan Coursac, and Scott Zeger (2000). “Fine Particulate Air Pollution and Mortality in 20 US Cities, 1987–1994.” *New England journal of medicine*, 343(24): 1742-1749.

Schlenker, Wolfram and Michael Roberts (2009). “Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields under Climate Change.” *PNAS,* 106(37): 15594-15598.

Schlenker, Wolfram and W. Reed Walker (2016). “Airports, Air Pollution, and Contemporaneous Health.” *Review of Economic Studies,* 83 (2): 768-809.

Simon, Noah, Jerome Friedman, Trevor Hastie, and Rob Tibshirani (2011). “Regularization Paths for Cox’s Proportional Hazards Model via Coordinate Descent.” *Journal of Statistical Software,* 39(5): 1-13.

Steenland, Kyle, and Ben Armstrong (2006). “An overview of methods for calculating the burden of disease due to specific risk factors.” *Epidemiology* 17(5): 512-519.

Stock, James and Motohiro Yogo (2005). “Testing for Weak Instruments in Linear IV Regression.” Ch.5 in J.H. Stock and D.W.K. Andrews (eds), *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, Cambridge University Press.

Tibshirani, Robert (1997). “The Lasso Method for Variable Selection in the Cox Model.” *Statistics in medicine*, 16(4): 385-395.

Ward, Courtney (2015). “It’s an Ill Wind: the Effect of Fine Particulate Air Pollution on Respiratory Hospitalizations.” *Canadian Journal of Economics*, 48(5): 1694-1732.

Zanobetti, Antonella, and Joel Schwartz (2006). “Air pollution and emergency admissions in Boston, MA.” *Journal of Epidemiology and Community Health* 60(10): 890-895.

Zhang, Q., Jiang, X., Tong, D., Davis, S. J., Zhao, H., Geng, G. & Ni, R. (2017). "Transboundary health impacts of transported global air pollution and international trade." *Nature*, *543*(7647), 705.

Figures

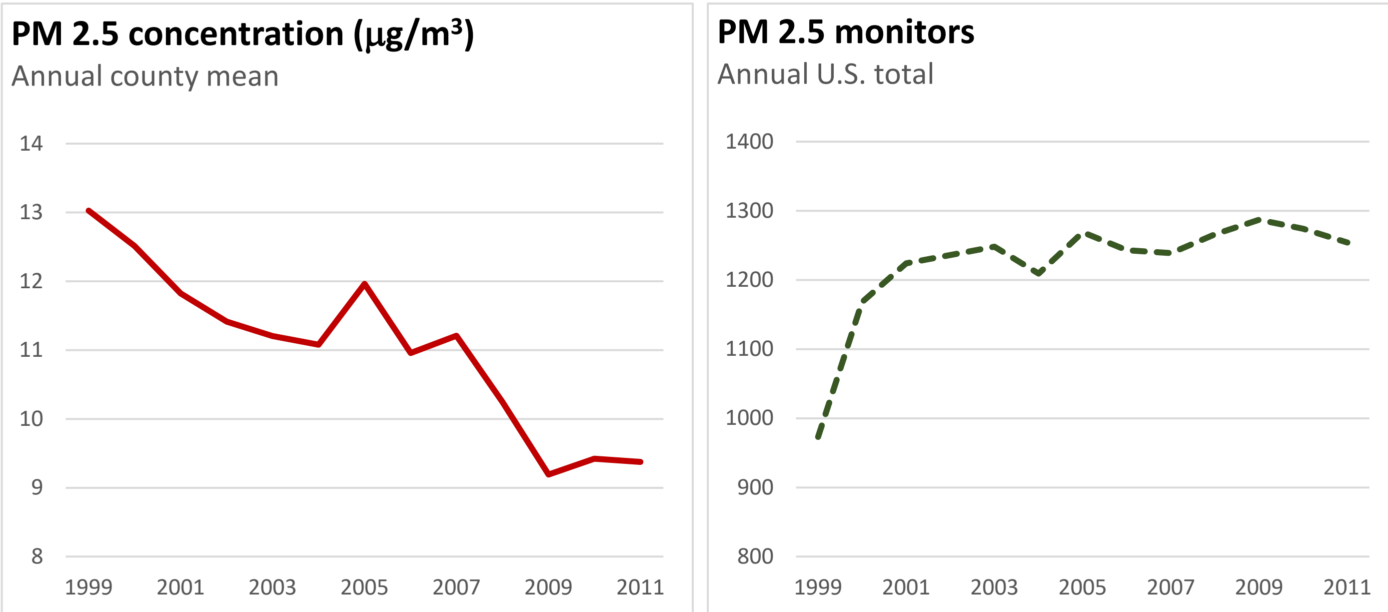

Figure 1. Trends in PM 2.5 air pollution and monitoring, 1999-2011. Figure displays annual county means for PM 2.5 concentration (left panel), and the nationwide total number of PM 2.5 monitors (right panel).

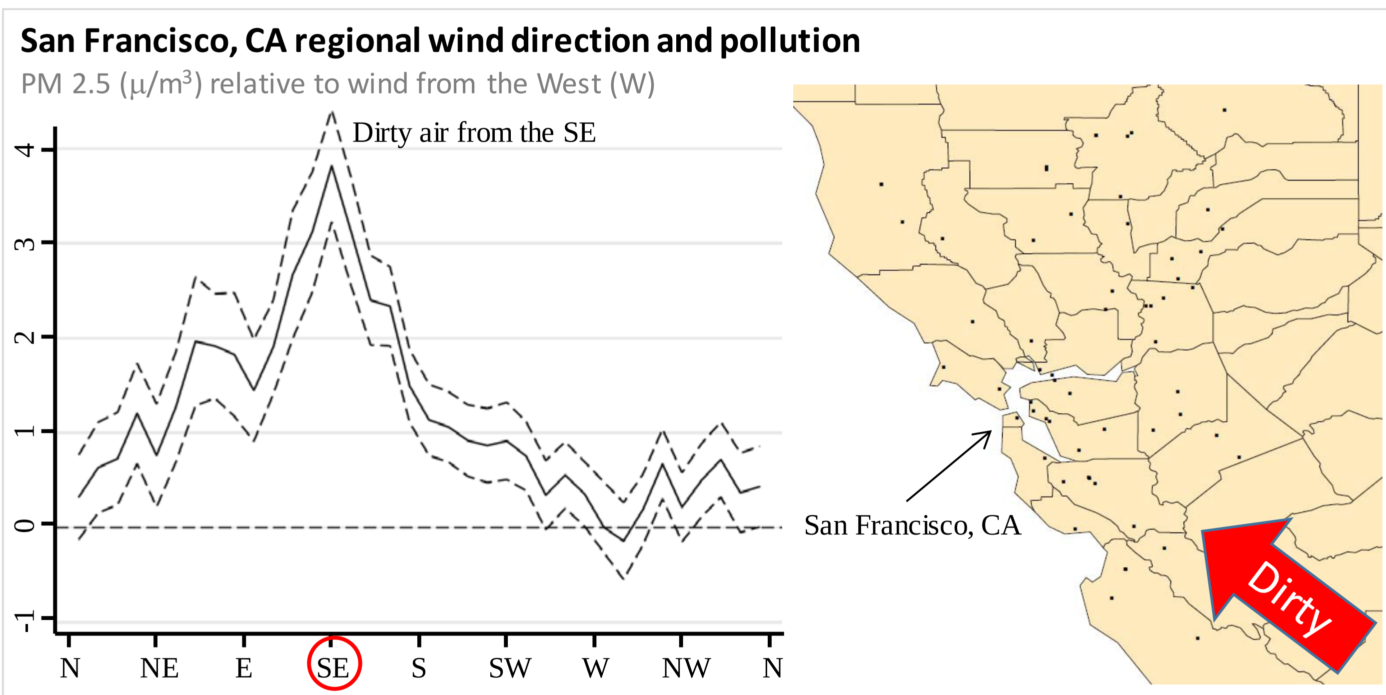

Figure 2. Relationship between daily average wind direction and PM 2.5 concentrations for counties in and around the Bay Area, CA. The left panel shows regression estimates of equation (A1) from the appendix, where the dependent variable is the county average daily PM 2.5 concentration and the key independent variables are a set of indicators for the daily wind direction falling into a particular 10-degree angle bin. Controls include county, month-by-year, and state-by-month fixed effects, as well as a flexible function of maximum and minimum temperatures, precipitation, wind speed, and the interactions between them. The dashed lines represent 95 percent confidence intervals based on robust standard errors. The right panel shows the location of the PM 2.5 pollution monitors (black dots) in the Bay Area that provided the pollution measures for this regression.

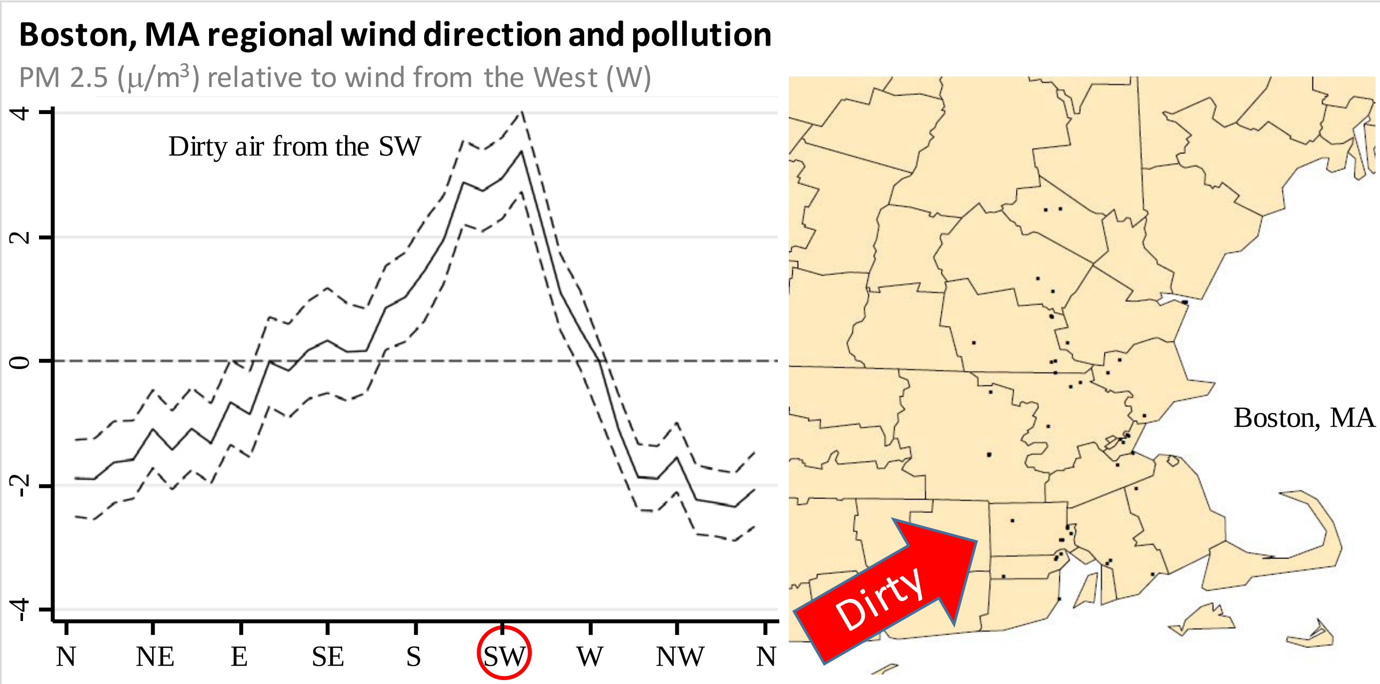

Figure 3. Relationship between daily average wind direction and PM 2.5 concentrations for counties in and around the Boston Area, MA. The left panel shows regression estimates of equation (A1) from the appendix, where the dependent variable is the county average daily PM 2.5 concentration and the key independent variables are a set of indicators for the daily wind direction falling into a particular 10-degree angle bin. Controls include county, month-by-year, and state-by-month fixed effects, as well as a flexible function of maximum and minimum temperatures, precipitation, wind speed, and the interactions between them. The dashed lines represent 95 percent confidence intervals based on robust standard errors. The right panel shows the location of the PM 2.5 pollution monitors (black dots) in the Boston Area that provided the pollution measures for this regression.

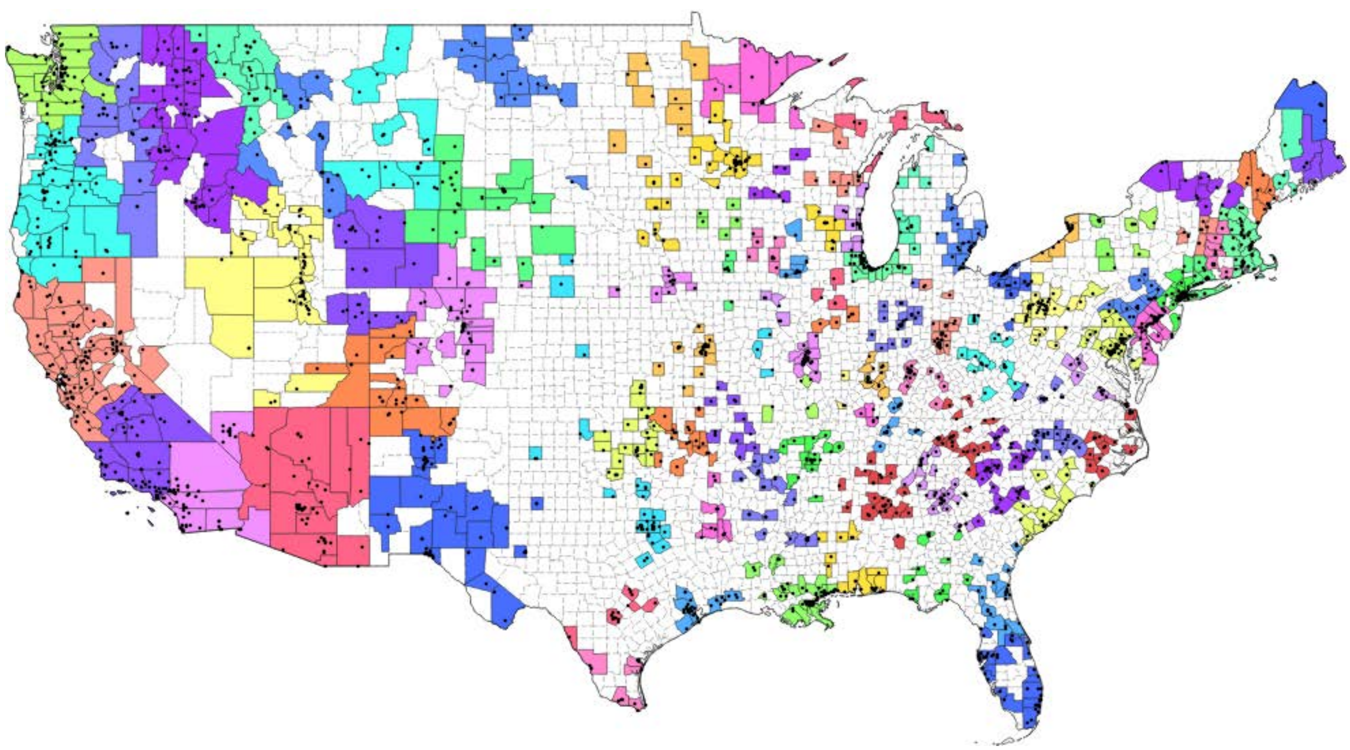

Figure 4. Counties assigned to each monitor group. Different colors correspond to different monitor groups. White corresponds to counties not assigned to any monitor group due to lack of monitors. Black dots represent PM 2.5 pollution monitors.

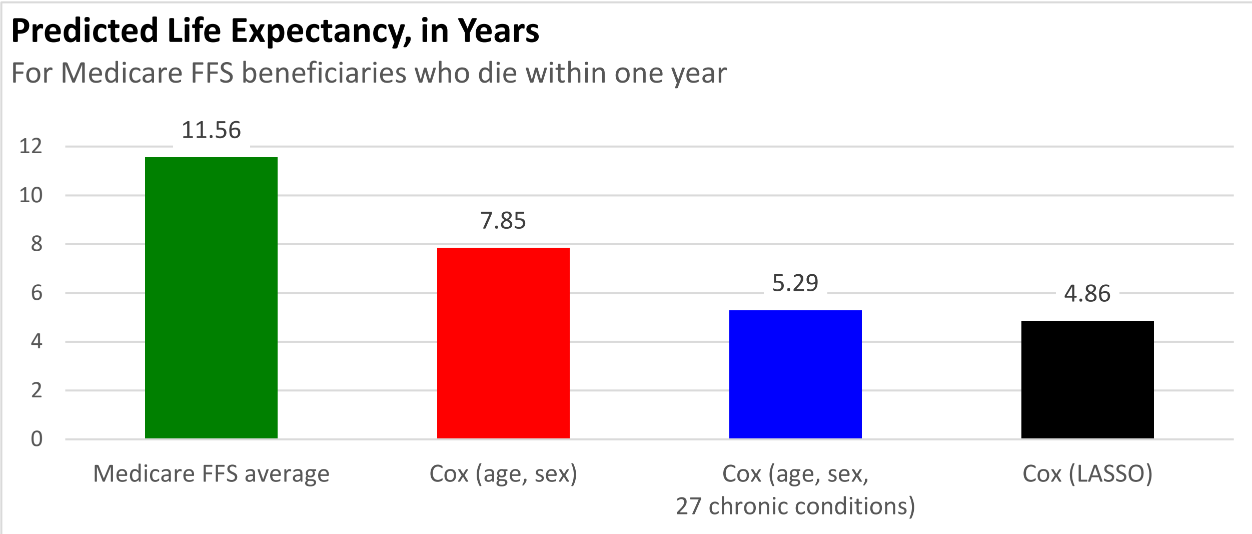

Figure 5. Average life expectancy for continuously enrolled fee-for-service Medicare beneficiaries who later die within one year, 2001-2011. Life expectancy for each beneficiary is estimated as of January 1 of the calendar year of death. Estimates for "Medicare FFS average" are produced by MLE estimation of survival model (6) with no covariates. Estimates for "Cox (age sex)" and "Cox (age sex cc) are produced by estimating the survival model (6) using age and gender, and age, gender and 27 chronic conditions, as predictors, respectively. Estimates for "Cox (LASSO)" are produced by machine learning estimation of the survival model (7) with 1,062 included regressors.

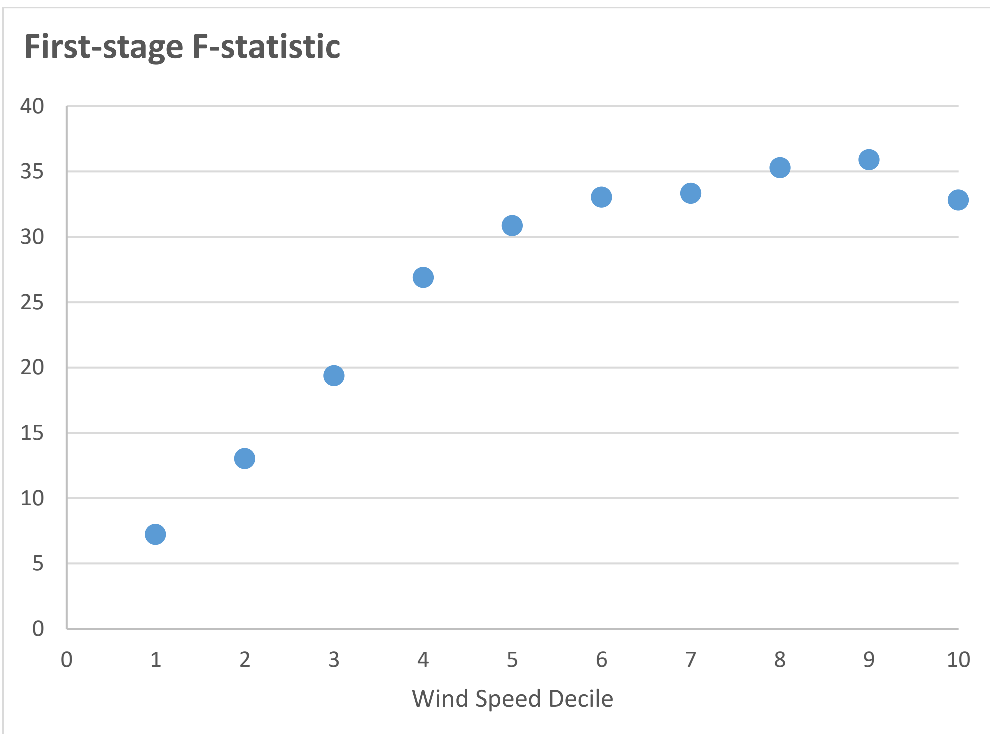

Figure 6. Relationship between the strength of the first stage and wind speed. This figure reports the F-statistic for our first stage (equation 2) for ten different subsamples that each include only days that fall within a particular wind speed decile. The F-statistic is lowest when the sample is limited to days with low wind speeds.

TABLES

Table 1: Summary statistics, 1999-2011

	Mean	Standard deviation	Observations
PM 2.5 ($\mu g/m^3$)	10.86	7.34	1,600,846
Number of beneficiaries, all ages	49,486	78,795	1,600,846
Number of beneficiaries, 65-69	12,923	20,262	1,600,846
Number of beneficiaries, 70-74	11,726	18,731	1,600,846
Number of beneficiaries, 75-79	9,960	16,088	1,600,846
Number of beneficiaries, 80-84	7,695	12,437	1,600,846
Number of beneficiaries, 85+	7,181	11,708	1,600,846
Number of FFS beneficiaries	34,911	52,748	1,518,623
Continuously enrolled FFS beneficiaries	27,716	40,090	1,518,623
3-day mortality rate, all ages	393.49	249.46	1,600,846
3-day mortality rate, 65-69	137.56	269.47	1,600,846
3-day mortality rate, 70-74	205.25	379.71	1,600,846
3-day mortality rate, 75-79	325.52	486.45	1,600,846
3-day mortality rate, 80-84	530.92	742.12	1,600,846
3-day mortality rate, 85+	1,169.86	1,119.82	1,600,846
3-day mortality rate, all FFS	409.02	274.61	1,518,623
3-day mortality rate, continuously enrolled FFS	458.21	315.98	1,518,623
3-day inpatient spending, planned and ER	34,463,288	14,976,401	1,518,623
3-day inpatient ER spending	13,659,622	7,693,555	1,518,623
3-day admissions rate, planned and ER	3,370	1,210	1,518,623
3-day ER admissions rate	1,579	709	1,518,623
3-day ER (inpatient and outpatient) visit rate	4,159	1,198	1,518,623

Unit of observation is county-day. All rates are per million Medicare beneficiaries in the relevant group. Spending and admissions variables are only available for fee-for-service (FFS) beneficiaries. Life-years lost analysis uses variables only available for continuously enrolled FFS beneficiaries. All FFS samples begin in 2001 instead of 1999.

Table 2: OLS and IV estimates of effect of PM 2.5 on elderly mortality, by age group

	(1) 65+	(2) 65-69	(3) 70-74	(4) 75-79	(5) 80-84	(6) 85+
Panel A: OLS estimates						
PM 2.5 ($\mu g/m^3$)	0.098***	0.042***	0.022	0.033	0.137***	0.423***
	(0.021)	(0.015)	(0.019)	(0.023)	(0.037)	(0.074)
Dep. var. mean	393	138	205	326	531	1,170
Effect relative to mean, percent	0.025	0.030	0.011	0.010	0.026	0.036
Observations	1,600,846	1,600,846	1,600,846	1,600,846	1,600,846	1,600,846
Adjusted R-squared	0.249	0.080	0.086	0.084	0.081	0.115
Panel B: IV estimates						
PM 2.5 ($\mu g/m^3$)	0.605***	0.263***	0.312***	0.307***	0.775***	2.050***
	(0.065)	(0.071)	(0.075)	(0.106)	(0.177)	(0.264)
F-statistic	241.115	232.367	236.416	241.909	247.716	256.311
Dep. var. mean	391	134	201	318	514	1,132
Effect relative to mean, percent	0.155	0.196	0.155	0.097	0.151	0.181
Observations	1,600,846	1,600,846	1,600,846	1,600,846	1,600,846	1,600,846

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Table reports OLS and IV estimates of equation (1) from the main text. Standard errors (in parentheses) clustered by county. Dependent variable is the 3-day mortality rate per million beneficiaries in the relevant age group. All regressions include county, month-by-year, and state-by-month fixed effects; flexible controls for temperatures, precipitation, and wind speed; and two leads of these weather controls. OLS (IV) estimates also include two lags and two leads of PM 2.5 (instruments). Estimates are weighted by the number of beneficiaries in the relevant age group.

Table 3: OLS and IV estimates of effect of PM 2.5 on Medicare hospitalization outcomes

	(1) FFS all-age mortality	(2) All inpatient spending	(3) Inpatient E.R. spending	(4) Inpatient admissions rate	(5) Inpatient E.R. admissions rate	(6) Inpatient + outpatient E.R. rate	(7) Non-E.R. admissions rate (placebo)
Panel A: OLS estimates							
PM 2.5 ($\mu g/m^3$)	0.137***	-8439***	877	-0.560***	0.127**	0.406***	-0.687***
	(0.023)	(1993)	(790)	(0.156)	(0.062)	(0.094)	(0.130)
Dep. var. mean	407	34,463,220	13,659,597	3,370	1,579	4,159	1,791
Effect relative to mean, percent	0.034	-0.024	0.006	-0.017	0.008	0.010	-0.038
Observations	1,518,549	1,518,549	1,518,549	1,518,549	1,518,549	1,518,549	1,518,549
Adjusted R-squared	0.236	0.518	0.685	0.515	0.695	0.651	0.308
Panel B: IV estimates							
PM 2.5 ($\mu g/m^3$)	0.628***	17074*	15446***	2.034***	1.960***	2.290***	0.074
	(0.075)	(10182)	(4151)	(0.714)	(0.336)	(0.394)	(0.486)
F-statistic	237	237	237	237	237	237	237
Dep. var. mean	407	37,861,232	16,645,971	3,463	1,783	3,960	1,680
Effect relative to mean, percent	0.154	0.045	0.093	0.059	0.110	0.058	0.004
Observations	1,518,549	1,518,549	1,518,549	1,518,549	1,518,549	1,518,549	1,518,549

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Table reports OLS and IV estimates of equation (1) from the main text. Standard errors (in parentheses) clustered by county. All dependent variables are 3-day measures per million fee-for-service (FFS) beneficiaries. All regressions include county, month-by-year, and state-by-month fixed effects; flexible controls for temperatures, precipitation, and wind speed; and two leads of these weather controls. OLS (IV) estimates also include two lags and two leads of PM 2.5 (instruments). Estimates are weighted by the number of FFS beneficiaries.

Table 4: IV estimates of effect of PM 2.5 on elderly life-years lost, using different survival models

	(1) All-age mortality	(2) None	(3) Age, sex	(4) Age, sex, chronic conditions	(5) LASSO
		Life-years lost regressions			
PM 2.5 ($\mu g/m^3$)	0.746***	8.625***	5.925***	3.539***	2.693***
	(0.085)	(0.978)	(0.757)	(0.562)	(0.521)
F-statistic	239	239	239	239	239
Dep. var. mean	462	5,338	3,624	2,444	2,245
Effect relative to mean, percent	0.162	0.162	0.163	0.145	0.120
LYL per decedent	NA	11.557	7.847	5.292	4.861
LYL per complier	NA	11.557	7.939	4.742	3.608
Observations	1,518,549	1,518,549	1,518,549	1,518,549	1,518,549

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Table reports IV estimates of equation (1) from the main text. Standard errors (in parentheses) clustered by county. The dependent variable in column 1 is the 3-day mortality rate per million continuously enrolled fee-for-service (FFS) Medicare beneficiaries. The dependent variable in columns 2-5 is life-years lost (LYL) over 3 days for the same group. The headings in columns 2-4 display the variables used to predict life expectancy when using a traditional Cox proportional hazards model. Column 5 displays results when life expectancy is predicted using a Cox proportional hazards model that is estimated using a LASSO machine learning algorithm with over one thousand predictors. LYL per decedent is calculated by dividing the average LYL in the sample by the average mortality rate. LYL per complier is calculated by dividing the columns estimate by the mortality effect reported in column 1. All regressions include county, month-by-year, and state-by-month fixed effects, as well as flexible controls for temperatures, precipitation, and wind speed; two leads of the weather controls; and two leads and lags of the instruments. Estimates are weighted by the number of continuously enrolled FFS beneficiaries.

Table 5: IV estimates of effect of PM 2.5 on elderly life-years lost, by remaining life expectancy

	(1) <1 year (0.69%)	(2) 1-2 years (2.24%)	(3) 2-5 years (12.7%)	(4) 5-10 years (29.8%)	(5) >10 years (54.6%)
			Panel A: mortality		
PM 2.5 ($\mu g/m^3$)	18.885***	5.267***	2.160***	0.533***	0.041
	(3.441)	(1.409)	(0.426)	(0.140)	(0.049)
F-statistic	238	254	249	241	234
Dep. var. mean	4,593	2,955	1,425	421	89
Effect relative to mean, percent	0.411	0.178	0.152	0.127	0.047
Observations	1,482,554	1,515,728	1,518,549	1,518,549	1,518,549
			Panel B: life-years lost		
PM 2.5 ($\mu g/m^3$)	11.281***	8.202***	6.727***	3.700***	0.688
	(2.233)	(2.239)	(1.439)	(0.973)	(0.633)
F-statistic	238	254	249	241	234
Dep. var. mean	2,711	4,517	4,773	2,958	1,132
Effect relative to mean, percent	0.416	0.182	0.141	0.125	0.061
Aggregate burden, percent	3.00	7.08	32.9	42.5	14.5
Observations	1,482,554	1,515,728	1,518,549	1,518,549	1,518,549

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Table reports IV estimates of equation (1) from the main text. Standard errors (in parentheses) clustered by county. The dependent variable is either deaths (Panel A) or life-years lost (Panel B) over 3 days per million continuously enrolled fee-for-service beneficiaries for those with remaining life expectancy in the range given by the column heading. Column headings also display (in parentheses) the percent of beneficiaries falling into each range. Life expectancy is predicted using a Cox proportional hazards model that is estimated using a Cox machine learning algorithm with over one thousand predictors. All regressions include county, month-by-year, and state-by-month fixed effects, as well as flexible controls for temperatures, precipitation, and wind speed; two leads of the weather controls; and two leads and lags of the instruments. Estimates are weighted by the number of continuously enrolled FFS beneficiaries.

Table 6: IV estimates of effect of PM 2.5 on elderly mortality when controlling for other pollutants

	(1)	(2)	(3)	(4)
	Panel A: all beneficiaries			
PM 2.5 ($\mu g/m^3$)	0.437***	0.298***	0.568***	0.346***
	(0.101)	(0.098)	(0.097)	(0.123)
CO		0.023***		0.021***
		(0.007)		(0.008)
Ozone			-0.290***	-0.084
			(0.109)	(0.121)
F-statistic	118	33	49	27
Dep. var. mean	391	391	391	391
Observations	552,412	552,412	552,412	552,412
	Panel B: fee-for-service beneficiaries			
PM 2.5 ($\mu g/m^3$)	0.663***	0.568***	0.859***	0.799***
	(0.113)	(0.122)	(0.125)	(0.170)
CO		0.013		0.005
		(0.009)		(0.010)
Ozone			-0.443***	-0.393**
			(0.149)	(0.180)
F-statistic	111	31	45	25
Dep. var. mean	462	462	462	462
Observations	490,413	490,413	490,413	490,413

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. Table reports IV estimates of equation (1) from the main text, with the addition of the endogenous variables CO and/or ozone, which are instrumented for using wind direction. Dependent variable is the 3-day mortality rate per million beneficiaries (Panel A) or per million fee-for-service (FFS) beneficiaries (Panel B). The sample is restricted to county-days where readings for CO, ozone, and PM 2.5 are simultaneously available. All regressions include county, month-by-year, and state-by-month fixed effects; flexible controls for temperatures, precipitation, and wind speed; two leads of the weather controls; and two leads and lags of the instruments. Estimates are weighted by the number of Medicare beneficiaries in Panel A and by the number of FFS beneficiaries in Panel B.

Table 7: LIML IV estimates of effect of PM 2.5 on elderly mortality, by age group

	(1)	(2)	(3)	(4)	(5)	(6)
	65+	65-69	70-74	75-79	80-84	85+
PM 2.5 ($\mu g/m^3$)	0.607***	0.264***	0.313***	0.308***	0.777***	2.055***
	(0.066)	(0.071)	(0.075)	(0.107)	(0.178)	(0.265)
F-statistic	241	232	236	242	248	256
Dep. var. mean	393	138	205	326	531	1,170
Observations	1,600,846	1,600,846	1,600,846	1,600,846	1,600,846	1,600,846

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Table reports IV estimates of equation (1) from the main text when using the LIML estimator instead of the 2SLS estimator. Standard errors (in parentheses) clustered by county. All dependent variables are 3-day mortality rates per million beneficiaries in the relevant age group. All regressions include county, month-by-year, and state-by-month fixed effects; flexible controls for temperatures, precipitation, and wind speed; two leads of these weather controls; and two lags and two leads of the instruments. Estimates are weighted by the number of beneficiaries in the relevant age group.

Table 8: Placebo IV estimates of effect of PM 2.5 on elderly mortality, by age group

	(1) 65+	(2) 65-69	(3) 70-74	(4) 75-79	(5) 80-84	(6) 85+
PM 2.5 ($\mu g/m^3$)	-0.674	0.509	0.697	2.279	-2.736	-7.554**
	(0.672)	(0.840)	(1.056)	(1.384)	(1.898)	(3.260)
F-statistic	1.511	1.455	1.516	1.552	1.548	1.550
Dep. var. mean	393	138	205	326	531	1,170
Observations	1,600,846	1,600,846	1,600,846	1,600,846	1,600,846	1,600,846

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Table reports IV estimates of equation (1) from the main text when using randomly generated placebo instruments. Standard errors (in parentheses) clustered by county. All dependent variables are 3-day mortality rates per million beneficiaries in the relevant age group. All regressions include county, month-by-year, and state-by-month fixed effects; flexible controls for temperatures, precipitation, and wind speed; two leads of these weather controls; and two lags and two leads of the instruments. Estimates are weighted by the number of beneficiaries in the relevant age group.

Table 9: Robustness of mortality IV estimates to different levels of aggregation for pollution monitors and wind angles

	(1)	(2)	(3)
PM 2.5 ($\mu g/m^3$)	0.615***	0.646***	0.612***
	(0.058)	(0.062)	(0.063)
Size of wind angle bins (degrees)	60	90	90
Number of monitor groups	100	50	200
F-statistic	158	452	142
Dep. var. mean	393	393	393
Observations	1,600,846	1,600,846	1,600,844

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Table reports IV estimates of equation (1) from the main text. The baseline specification reported in other tables aggregates pollution monitors into 100 groups and wind angles into 90-degree intervals. This table demonstrates that our estimates are not sensitive to the chosen level of aggregation. Standard errors (in parentheses) clustered by county. The dependent variable is the 3-day mortality rate per million Medicare beneficiaries. Estimates are weighted by the number of beneficiaries.

Table 10: Robustness of mortality IV estimates to including different fixed effects and weather controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PM 2.5 ($\mu g/m^3$)	0.382***	0.571***	0.244***	0.295***	0.372***	0.615***	0.649***	0.583***
	(0.043)	(0.066)	(0.046)	(0.047)	(0.042)	(0.065)	(0.066)	(0.066)
Type of weather controls	None	Separate	None	None	None	Full	Full	Full
County f.e.	X	X	X	X		X	X	
Month f.e.			X			X		
Year f.e.			X			X		
Year-by-month f.e.	X	X		X	X		X	X
State-by-month f.e.	X	X						
County-by-month f.e.					X			X
F-statistic	374	269	355	363	385	228	231	247
Dep. var. mean	394	394	394	394	394	393	393	393
Observations	1,602,889	1,602,889	1,602,889	1,602,889	1,602,860	1,600,846	1,600,846	1,600,817

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Table reports IV estimates of equation (1) from the main text when varying the inclusion of different weather controls and fixed effects. Standard errors (in parentheses) clustered by county. The dependent variable is the 3-day mortality rate per million Medicare beneficiaries. Estimates are weighted by the number of beneficiaries.

Table 11: Robustness of all-age mortality and life-years lost estimates to including fewer or more instrument lags

	(1) No lags	(2) 1 lag	(3) 3 lags	(4) 4 lags	(5) 5 lags
			Panel A: mortality		
PM 2.5 ($\mu g/m^3$)	0.525***	0.655***	0.614***	0.611***	0.616***
	(0.078)	(0.065)	(0.066)	(0.065)	(0.065)
F-statistic	316	247	242	241	239
Dep. var. mean	394	394	393	393	393
Observations	1,624,689	1,612,384	1,590,074	1,579,878	1,570,025
			Panel B: life-years lost		
PM 2.5 ($\mu g/m^3$)	2.589***	3.043***	2.751***	2.713***	2.714***
	(0.536)	(0.535)	(0.554)	(0.548)	(0.546)
F-statistic	309	244	238	237	236
Dep. var. mean	2,211	2,211	2,210	2,210	2,210
Observations	1,537,668	1,527,835	1,509,758	1,501,350	1,493,170

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Table reports IV estimates of equation (1) from the main text. Column headings report the number of instrument lags included in the regression. (The specification reported in other tables includes 2 lags.) Standard errors (in parentheses) clustered by county. Dependent variable in Panel A is the 3-day mortality rate per million beneficiaries. Dependent variable in Panel B is the life-years lost over 3 days per million continuously enrolled fee-for-service (FFS) beneficiaries. Estimates are weighted by the number of Medicare beneficiaries in Panel A and by the number of FFS beneficiaries in Panel B.