THE EFFECT OF POLLUTION ON HEALTH AND HEALTH CARE UTILIZATION: EVIDENCE FROM CHANGES IN WIND DIRECTION*

Tatyana Deryugina
University of Illinois at Urbana-Champaign and NBER

Garth Heutel
Georgia State University and NBER

Nolan H. Miller
University of Illinois at Urbana-Champaign and NBER

David Molitor
University of Illinois at Urbana-Champaign and NBER

Julian Reif
University of Illinois at Urbana-Champaign

March 2016

Abstract. Air pollution is believed to have large impacts on health and on health care utilization and costs. However, estimating the causal impact of pollution exposure is complicated by endogeneity and measurement error concerns. We address these challenges by employing a novel exogenous instrument for pollution exposure: changes in the local wind direction. We show that local wind direction strongly predicts fine particulate matter (PM 2.5) pollution concentrations. Using detailed administrative data on the universe of Medicare beneficiaries, we find that particulate matter significantly increases daily mortality, but has no effect on the hospitalization rate or total hospital spending. Regressions with multiple endogenous air pollutants indicate that the mortality effects of PM 2.5 are not explained by co-transported pollutants like ozone, carbon monoxide, sulfur dioxide, and nitrogen dioxide. Finally, we develop a new methodology for estimating the years of life lost due to pollution using detailed data on health status and estimate that the mortality benefit of the reduction in PM 2.5 since 1999 is worth $13 billion annually.

* We thank Christos Makridis and participants at the University of Illinois Research Lunch for helpful comments.

I. INTRODUCTION

Guided by the widely accepted belief that air pollution negatively affects health, many countries have tightened air pollution standards in the past several decades. Precisely how much pollution affects mortality, morbidity, and health care spending matters greatly for determining optimal environmental policy. This is especially true for countries like the United States, where current pollution levels are relatively low and further reductions may be very costly. More generally, optimal Pigouvian pricing of emissions depends on their marginal damages. Similarly, cost-benefit analysis of command-and-control regulations must quantify and monetize the health benefits of any pollution reductions, along with the accompanying reduction in health care costs.

Estimating the causal effect of pollution on health has proven challenging. A primary concern has been the likely endogeneity of pollution. For example, traffic congestion may both affect pollution and increase drivers' stress, which in turn could have negative health consequences. In such a circumstance, naïve estimation could falsely attribute the health consequences of stress to pollution. A second concern is the possibility that pollution is measured with error. For example, pollution levels over a wide geographic area are often imputed using sparsely located pollution monitors. To the extent that monitor readings are noisy signals of pollution levels away from the monitors, this gives rise to the possibility of attenuation bias due to measurement error when using ordinary least squares. Estimating the impact of pollution using an instrumental variables approach would address both of these issues, but to do so one must find an exogenous source of variation in pollution, which is a challenging task.

We use a novel source of variation in local air pollution—local wind direction—to estimate the effect of fine particulate matter (PM 2.5) on daily county-level mortality, hospitalizations, and health care spending among the U.S. elderly population. Our main identifying assumption is that changes in wind direction are uncorrelated with other, unobserved factors that may affect health or health care utilization. We control for temperature, precipitation, wind speed, and a large number of fixed effects to increase the plausibility of this assumption. An attractive feature of our identification strategy is the availability of national data on historical wind direction are available. Our estimates are therefore not limited to a particular geographic area or time, unlike many other studies (Chay and Greenstone 2003; Currie and Neidell 2005; Schlenker and Walker 2016).

Our empirical strategy consists of two stages. In the first stage, we estimate how local wind direction affects daily PM 2.5 concentrations, conditional on a host of fixed effects and other climatic controls. To ensure that we are picking up long-range transport rather than local pollution sources, where downwind or upwind sorting may be an issue, our instrument uses the relationship between wind direction

and pollution for *groups* of monitors rather than single monitors. The second stage then uses this exogenous variation to estimate the effect of pollution levels on mortality and health care utilization, using over a decade of data for the entire elderly Medicare population, which captures over 97% of adults aged 65 and older.

Using this approach, we estimate that a 1-unit increase in PM 2.5 exposure (measured in micrograms per cubic meter) for one day causes 0.33 additional deaths per million elderly individuals on that day. The increase is larger for older beneficiaries in absolute terms, but similar in relative terms because older beneficiaries have a higher baseline death rate than younger beneficiaries. Importantly, these estimates are significantly larger that the corresponding OLS results, suggesting the presence of substantial bias in studies that do not employ an exogenous source of pollution variation. By contrast, we find no effect of PM 2.5 on the hospitalization rate or total medical spending, and a small decrease in spending per admit and the total length of stay. OLS estimates, however, indicate a significant positive correlation between PM 2.5, total medical spending, and the hospitalization rate, suggesting that these naïve estimates suffer from omitted variable bias.

Another central concern in the study of the health effects of pollution is whether those who die because of pollution exposure would have died in the near future even in the absence of exposure. This so-called "harvesting hypothesis" is important for policy because conventional estimates of the social cost of mortality try to account for decedents' remaining life expectancy. If the mortality effect of pollution is concentrated among the oldest and sickest individuals, then the mortality benefit from reducing pollution is likely to be much smaller than if the effect was concentrated among individuals randomly chosen from the population. The harvesting hypothesis is particularly important to consider in settings that use high-frequency variation in pollution, such as ours.

Fortunately, the richness of the Medicare data provides us with an extraordinary opportunity to investigate the true mortality cost of pollution.[2] We propose a new methodology that exploits the detailed data available in the Medicare claims dataset to generate an estimate that is less prone to bias than previous methods. We first present a framework that illustrates why the traditional method of estimating life-years lost is likely to produce upward biased estimates. We outline the specific assumptions required to eliminate this bias and explain how the rich Medicare claims data can plausibly meet those assumptions. We then use

[2] Traditional analyses of the question of harvesting or short-term mortality displacement have concentrated on using leads and lags of the independent variables of interest to investigate whether mortality effects moderate as the length of time under consideration increases, as would be the case under harvesting (Schlenker and Walker 2016). Future versions of this paper will include such an analysis as well. However, such an analysis still does not fully account for the fact that those who died may have had shorter life expectancies than those who lived.

individual-level demographic and health history data to estimate counterfactual survival functions. We find that accounting for decedents' age significantly reduces estimates of life-years lost. Accounting for additional factors such as the presence of different types of chronic conditions further decrease estimates, but they are not as important as age. We then compute life-years saved among the elderly population due to a reduction of 3.65 units in daily levels of PM 2.5, which is approximately equal to the decrease in the national average between 1999 and 2011. We estimate that this decrease reduced the number of life-years lost by 130,000. Assuming a value of $100,000 per statistical life-year then implies a total benefit of $13 billion. Ignoring the individual-level data available to us and instead estimating life-years lost using the average life expectancy for the population—as is commonly done in the health and environmental literatures—increases this estimate by 123 percent. This demonstrates the value of our technique for addressing "harvesting" and for placing an accurate value on the social benefits of reducing mortality.

A large literature has documented a relationship between pollution exposure and health outcomes, like mortality, even after controlling for a number of possible confounding factors. However, the vast majority of this literature has been correlational (e.g., Dockery et al. 1993, Pope et al. 1995, Laden et al. 2000, Samet et al. 2000, Currie and Neidell 2005, Pope and Dockery 2006, Currie et al. 2009, EPA 2009). The best-identified studies often focus on infants (e.g., Currie and Walker 2011), leaving a large gap in the literature.

We contribute to the literature in three main ways. First, we know of no other study that has used national data on wind direction to instrument for pollution exposure as we do. Doing so both gives us power to detect effects that are small in magnitude (but that may be economically important) and increases the generalizability of our study. Our approach also does not rely on being able to identify pollution sources, greatly increasing the scope of its applicability. A handful of recent papers use wind direction interacted with a particular pollution source to estimate the effects of pollution from that source. Schlenker and Walker (2016) find morbidity impacts of carbon monoxide in populations living near California airports, while recent working papers by Anderson (2015) and Herrnstadt and Muehlegger (2015) each focus on a single metropolitan area and exploit variation in pollution across interstates due to local wind direction. Interpreting the effects estimated in the latter two papers as the causal effect of a given pollutant is challenging because air pollutants like particulate matter (PM 2.5), ozone (O3), carbon monoxide (CO), sulfur dioxide (SO2), and nitrogen dioxide (NO2) can be co-transported by wind. Our strategy, however, allows us to estimate regressions with multiple endogenous air pollutants because variation in local wind direction changes the mix of air pollutants as well as their overall levels. Moreover, because our approach is not tethered to specific pollution sources and nearby residents, our estimates are representative of the US elderly population residing in counties monitored by the US Environmental Protection Agency (EPA).

A second key contribution is that our Medicare data on mortality, morbidity, and health care utilization and costs provide the most detailed information to date on the effect of air pollution on the elderly, a particularly vulnerable population. Third, we address the issue of "harvesting"—the extent to which the deaths we attribute to PM 2.5 are deaths that would have occurred in the near future anyway—by using our detailed data on demographics and health histories to estimate a plausible lower bound on life-years lost from pollution exposure. To our knowledge, this paper is the first to account precisely for the effects of harvesting when estimating the effects of pollution on mortality. The methodology we develop is very general and can be applied to a variety of different settings.[3]

The rest of the paper is organized as follows. Section II 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, including how we estimate the life-years lost, in detail. Results are presented in Section V, and Section VI concludes.

II: WIND TRANSPORT OF POLLUTION

Fine particulate matter, PM 2.5, is a mixture of various compounds including nitrates, sulfates, ammonium, and carbon (e.g., Kundu and Stone 2014). In addition to natural sources, PM 2.5 comes from power plant and car emissions and can be carried for hundreds of miles from where it is emitted. Sulfur dioxide and nitrogen dioxide, two other pollutants regulated by the EPA, are also precursors to sulfates and nitrates, which are components of PM 2.5. According to EPA estimates, there are many parts of the country, particularly the East, where regional, rather than local emissions, make up a significant share of local particulate matter (EPA 2004).

The extent of pollution transport depends on a host of factors, including wind direction and speed, precipitation, the height of the planetary boundary layer, and chemical reactions with other airborne molecules. One way to exploit variation in pollution transport is to employ a sophisticated atmospheric science model to simulate daily pollution transport across the United States and use the resulting estimates as instruments. However, this is neither feasible nor desirable for three key reasons. First, the more sophisticated models require the researcher to specify emission sources. With the exception of power plant locations, this information is not readily available, even for the United States. Second, this approach could potentially inject confounding factors into the analysis because changes in nearby emissions are likely to be correlated with changes in nearby economic activity, which could itself affect local health outcomes.

[3] For example, estimating the number of life-years lost matters for evaluating the mortality-reduction benefits of health care insurance programs such as Medicare and Medicaid.

Third, even if information on emission sources were available, the computational burden associated with simulating daily pollution transport for many locations over many years is enormous.

An instrumental variables approach, by contrast, only makes use of some of the factors involved in pollution transport and thus is much simpler to implement. It only requires 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 use this fact to instrument for local PM 2.5 concentrations using only local wind direction, which we shall show is by itself an important determinant of local pollution levels. Importantly, we use daily *variation* in local wind direction rather than the prevailing wind direction. A concern with using prevailing wind directions is that individuals may sort to be upwind or downwind of the pollution, biasing the estimates. By contrast, holding constant the prevailing wind direction and looking at daily deviations largely eliminates this concern.

We now illustrate the type of variation that we will utilize to estimate the causal effects of PM 2.5. Figure 1 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 Bay Area, CA. Figure 2 shows the same relationship for pollution monitors in and around the Boston Area, 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, as will be discussed in Section IV.

In both cases, the local wind direction is a very strong predictor of pollution levels. Moreover, the patterns are consistent with what we would expect given the geographic placement of the monitors. In and around the Bay Area, PM 2.5 levels are highest, all else equal, when the wind is blowing from the East and the Southeast and lowest when the wind is blowing from the South, North, and the West. In other words, more pollution is blown in from Southern California and the East than from the ocean and the Northern states like Oregon and Washington. In and around Boston, MA, on the other hand, 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 are dominant.

We classify the pollution monitors in our data in 100 monitor groups, based on their geographic proximity (the monitors in Figures 1 and 2 represent two such groups). Our first stage, described in Section IV, allows the relationship between pollution and wind direction to vary at the monitor *group* level. Such a restriction will not only ease the computational burden of having many instruments but also avoid picking up within-group idiosyncratic relationships between pollution and wind direction. To the extent that local

sources are not located systematically to one side of pollution monitors in a given cluster, our first stage will then reflect long-range transport, making sorting less likely to be a problem.

III: DATA

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 many pollutants, including the six criteria pollutants regulated by the Clean Air Act. We focus on PM 2.5, but we also obtain data on four other criteria pollutants – ozone (O_3), carbon monoxide (CO), nitrogen dioxide (NO_2), and sulfur dioxide (SO_2). As with PM 2.5, past literature has linked these air pollutants to mortality and other health outcomes. Monitor readings are aggregated up to the daily level by averaging across hourly observations. County level pollution measures are constructed by averaging across all monitors located within the county with a non-missing pollution reading for a given pollutant on a given day. Comprehensive data for PM 2.5 are available beginning in 1999.

Figure 3 displays aggregate trends in PM 2.5 over time. Average concentrations of PM 2.5 have been steadily falling from about 13 micrograms per meter cubed $\left(\frac{\mu g}{m^3}\right)$ in 1999 to about 9 $\frac{\mu g}{m^3}$ in 2011. It is worth noting that 1 unit of PM 2.5 thus represents about 10% of the average concentration during out time period.

The amount of spatial and temporal coverage depends on the pollutant and varies by year because the EPA generally monitors only the dirtiest counties. If a county becomes sufficiently clean then the EPA often ceases to monitor it. Similarly, the EPA will begin monitoring a previously unmonitored county if it becomes dirty. As a result, the set of monitored counties changes over time. This "churn" in monitored counties can make it difficult to interpret aggregate changes in pollution over time such as the one displayed in Figure 3. To avoid this compositional bias problem, we perform all analyses at the county level with county fixed effects. Moreover, our instrumental variables approach will exploit variation in pollution that is independent of monitor placement, resulting in unbiased estimates.

Temperature and precipitation

We obtain daily temperature and precipitation data from Schlenker and Roberts (2009), who produce a detailed weather map at the daily level using data from PRISM and weather stations.[4] These data

[4] 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.

include total daily precipitation, and daily maximum and minimum temperatures for each point on a 2.5 by 2.5 mile grid covering the contiguous United States for the years 1999-2011. To aggregate the gridded data to the county level, we simply average the daily measures across all grid points located in a particular county.

Wind speed and direction

Wind speed and wind direction data are obtained from the North American Regional Reanalysis (NARR) daily reanalysis data published by the National Centers for Environmental Information (NCEI).[5] As with temperature and precipitation, wind variables are reported on a grid, but with a coarser resolution of about 32km. The wind data 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 the location of 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. Specifically, 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 onto a 0-360 scale depending on the signs of u and v. Specifically, given θ, the wind angle, wa, is calculated as follows:

$$wa = \begin{cases} 180 - \theta & \text{if u} < 0 \text{ and v} > 0 \\ \theta + 180 & \text{if u} < 0 \text{ and v} < 0 \\ 360 - \theta & \text{if u} > 0 \text{ and v} < 0 \\ \theta & \text{if u} > 0 \text{ and v} > 0 \end{cases}$$

Finally, we average the monitor-day-level wind direction and speed to the county-day level.

Mortality, morbidity, and health care costs

Our data on mortality, morbidity, and health care costs come from Medicare administrative data. We focus on elderly beneficiaries aged 65 – 100, which covers over 97% of elderly living in the U.S. Detailed demographic variables, including age, sex, dates of death and county of residence, are obtained from the annual Medicare enrollment files for years 1999-2011 and cover 100% of beneficiaries. Individual indicators for the presence of 27 chronic conditions, which we use to estimate life-years lost, are obtained from the Chronic Conditions segment of Master Beneficiary Summary File. Health care utilization and costs are derived from the Medicare Provider Analysis and Review (MEDPAR) File, which includes an

[5] Available from https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/north-american-regional-reanalysis-narr. The NCEI was formerly the National Climatic Data Center (NCDC).

observation for 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 service claims corresponding to that stay, and include the date of admission, length of stay, and total cost of the stay.[6] All county level daily measures of hospital utilization and costs are aggregates of the MEDPAR records based on the county of patient residence and the admission date.

Table 1 presents summary statistics for our main estimation sample, which consists of 2,221,481 observations at the county-day level. Our sample does not encompass the entire United States due to limitations in the EPA's pollution monitor coverage. In particular, PM 2.5 pollution measures are available for only 902 counties during our sample period (see Figure 4). However, because pollution monitors tend to be placed in more populated counties, our main regression estimates capture almost 70% of the Medicare population.

Table 1 reports that the mean daily concentration of PM 2.5 in our estimation sample is 10.9 micrograms per cubic meter, with a standard deviation of 7.34. There are 43.6 thousand Medicare beneficiaries in the average county, with about half of these aged between 65 and 74. Because we focus on the elderly, the daily death rate in our sample is fairly high, ranging from 47 per million for those aged 65-69 to 394 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 60% of the population in our sample. The daily hospital spending for this population is $423,000 per county or about $14.5 per beneficiary per day, in nominal terms. On average, there are about 1.5 hospital admissions per thousand FFS beneficiaries per day and each admission costs $9,600. Finally, the daily per-county length of hospital stay averages 427 days.

IV: EMPIRICAL STRATEGY

Effects of PM 2.5 on mortality and health care utilization

The key causal relationship we would like to estimate is

$$Y_{cdmy} = \beta \mathrm{PM2.5}_{cdmy} + f\left(Temp_{cdmy}, Prcp_{cdmy}, WindSpeed_{cdmy}\right) + \alpha_c + \alpha_{sm} + \alpha_{my} + \epsilon_{cdmy}, \quad (1)$$

[6] Specifically, our measure of cost is the total allowed charges due to the provider. This amount includes payments made by Medicare, the beneficiary, or a primary payer. Payments for which the patient is responsible, such as deductibles and co-insurance, may be paid directly out of pocket or may be covered by a Medicare Supplement Insurance (Medigap) policy or Medicaid.

where the dependent variable is either the death rate or some measure of health care costs/utilization 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.

The high granularity and comprehensive scope of our data allow us to estimate this regression with multiple sets of high-dimensional fixed effects. This allows us to control for weather, geography, time and seasonality far more flexibly than previous studies have done. To control for weather, we include a large set of indicator variables corresponding to county-level daily maximum and minimum temperature, precipitation, and wind speed. 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 each bin in between spanning 3 degrees Celsius (5.4°F). We do the same for minimum temperatures. 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 include it in all our daily regressions. While our estimates are very robust to less flexible weather controls, including the controls above ensures that our source of variation is truly exogenous.

The regression also includes 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 that vary by geography. State-by-month fixed effects control for potential seasonal correlation between wind direction and population health that may vary by region. Finally, month-by-year fixed effects control flexibly for common time-varying shocks, as could be induced by Medicare or environmental policies. Our results are very robust to varying the fixed effects.

OLS estimates of equation (1) are prone to bias because PM 2.5 levels are not randomly assigned. 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:

$$\mathrm{PM2.5}_{cdmy} = \sum_{g=1}^{100}\sum_{b=0}^{4} \beta_b^g WINDDIR_{cdmy}^{60b} + f(Temp_{cdmy}, Prcp_{cdmy}, WindSpeed_{cdmy}) + \alpha_c + \alpha_{sm} + \alpha_{my} + \epsilon_{cdmy} \quad (2)$$

The variable $WINDDIR_{cdmy}^{60b}$ is an indicator variable equal to 1 if the daily average wind direction in county c falls in the 60-degree interval $[60b, 60b + 60)$ and 0 otherwise. The omitted category corresponds to the interval [300,360). The coefficient on this variable, β_b^g, is allowed to vary across 100 different geographic regions, as explained below. The other control variables are defined as in equation (1).

In order to increase statistical power and reduce the computational burden, equation (2) estimates a common effect of wind direction on pollution for all monitors within each of the 100 geographic areas, most of which span multiple counties. We define these areas by using cluster analysis to classify all the pollution monitors in our data into 100 spatial groups based on their location. 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). Intuitively, monitors that are close to each other are very likely to be assigned to the same group. On average, each geographic area (group) contains 21 monitors and 9 counties.

In addition to reducing the computational burden associated with hundreds of instruments, clustering monitors in this way provides another benefit. The specification of equation (2) means that our identification strategy puts greater weight on pollution produced by distant sources that are systematically located to one side or another of the entire monitor *group*. If we instead allowed the relationship between the wind direction and pollution to vary for each monitor, we would likely pick up the influence of very local pollution sources, which is unappealing. For example, if individuals living upwind and downwind of these local sources were different, our estimates would be valid for the downwind individuals but not for the entire population. Equation (2), however, only utilizes the variation between wind direction and pollution that is common to *all* monitors in a monitor group. This makes it very likely that the source of our variation will be long-range transport rather than local sources.[7] Because the entire geographic area is affected similarly by these sources, this reduces the possibility that individuals will choose where to live *within* the geographic area based on which parts are likely to be most affected by these sources.

Equation (2) also restricts the effect of wind direction on pollution levels to be constant within each of the six $WINDDIR$ bins. This makes it even more likely that the variation in pollution we are employing is driven by long-range transport because pollution disperses across space as it is carried by the wind. For example, the amount of pollution transported from the Midwest to the East Coast is not likely to be affected

[7] We cannot test for this directly. However, in order for this approach to pick up the influence of local sources, the location of the local sources relative to the monitors would have to be correlated across monitors in a monitor group. Because the monitors in a group are fairly dispersed geographically, we think this is highly unlikely to be the case.

by 5-degree differences in local wind direction, whereas pollution from closer sources may be. Another advantage of restricting the number of bins is computational feasibility. It is computationally burdensome to increase the number of $WINDDIR$ bins; the specification presented in (2) includes 500 instruments, several hundred other control variables, and is estimated using over two million observations. The main downside of this coarse binning is that we are potentially failing to exploit some useful variation in pollution. We have investigated this possibility by experimenting with increasing the number of $WINDIR$ bins in our estimations; those results, available upon request, are very similar to the ones we present in our tables, which leads us to conclude that including only six $WINDDIR$ bins is reasonable.

The large number of instruments employed in our analysis raises the concern that our estimates may suffer from weak instrument bias. However, as illustrated by Figures 1 and 2, 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.[8] Moreover, estimating our model using the limited information maximum likelihood (LIML) estimator, which is approximately median-unbiased, rather than 2SLS, yields similar results. As a robustness check, we also estimate our model using placebo instruments and do not obtain large F-statistics or significant estimates.

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. For example, if the dependent variable is the elderly mortality rate then we weight by the county-level elderly population; if the dependent variable is the mortality rate for those 85 and older, then we weight by the county-level population that is 85 and older.

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. We can then use this estimate to calculate the number of deaths averted as a result of a decline in the levels of PM 2.5. However, in most economic frameworks, the social value of this mortality reduction depends on the number of *life-years* lost rather than just on the number of lives lost. A common estimate of the value of a statistical life-year is $100,000 (Cutler 2004). According to this metric, preventing the deaths of individuals who then go on to live for another four years is twice as valuable as preventing the deaths of individuals who live for only another two years.

[8] 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 larger than the first-stage F statistics computed assuming homoskedastic errors.

In practice, estimating life-years lost is challenging because counterfactual life expectancy is unobserved. The standard in the health and environmental literatures is to multiply the estimated number of lives lost by an assumed value of residual life expectancy per life lost. This residual life expectancy is typically derived from population life tables (Deschenes and Greenstone 2011) or estimated elsewhere (Finkelstein and McKnight 2008). A general concern with this approach is that it is likely to overstate estimates of life-years lost, because affected individuals likely have shorter life expectancies than average (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, but would live fewer years on average than similar individuals who do not have such conditions, even absent a pollution event.

We propose a new methodology that exploits the detailed data available in the Medicare claims dataset to generate an estimate that is less prone to bias than previous methods. We first present a framework that illustrates why the traditional method of estimating life-years lost is likely to produce upward biased estimates. We outline the specific assumptions required to eliminate this bias and explain how the rich Medicare claims data can plausibly meet those assumptions. We then produce estimates of the number of life-years saved due to a reduction in PM 2.5 pollution using our new methodology.

Suppose that individual i dies in period t. Let L_{it} be equal to the number of years a deceased individual would have lived had death been averted, and equal to 0 for individuals who do not die.[9] Then, assuming that exposure to PM 2.5 is assigned randomly to individuals, one can estimate the effect of PM 2.5 on the number of years of life lost in period t by estimating the following regression equation:

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

The error term e_{it} represents factors other than pollution that affect life-years lost, and by assumption is uncorrelated with $\mathrm{PM2.5}_{it}$. Thus, equation (3) consistently estimates the causal effect γ of $\mathrm{PM2.5}_{it}$ on life-years lost.[10]

Of course, in practice a researcher does not observe L_{it}, but observes only whether an individual dies. Counterfactual life expectancy must be estimated. For example, one can model it as a function of age,

[9] 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 without immediately killing her. In that case, our estimates can be regarded as lower bounds.

[10] In practice, we estimate life-years lost due to PM 2.5 at the county level, not the individual level. That is, our outcome variable is L^*_{jt}, where $L^*_{jt} = \frac{\sum_{i=1}^{N_{jt}} L_{ijt}}{N_{jt}}$ is the average number of life-years lost in county j at time t, and N_{jt} is the number of individuals living in county j. It is straightforward to extend the framework to accommodate this case.

which is a strong predictor of remaining life expectancy. Let $\hat{L}_{it}$ be the estimate of life expectancy generated by that model, and let $u_{it} \equiv L_{it} - \hat{L}_{it}$ describe the measurement error in this estimate. Then the analog of equation (3), which the researcher can estimate with observable data, is

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

The measurement error present in equation (4) can bias the estimation of γ when heterogeneous treatment effects are present. To see this, let the effect of air pollution on individual mortality be equal to γ_i, and decompose this effect into $\gamma_i = \gamma + v_i$. The estimating equation now becomes

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

The error term now contains a third component, $v_i \text{PM2.5}_{it}$, which contains the portion of the individual's treatment effect not accounted for by the average treatment effect, γ. From equation (5), if the heterogeneous treatment effect, v_i, is correlated with the measurement error in counterfactual life expectancy (i.e., $Cov(v_i, u_{it}) \neq 0$), then the estimate of γ will be biased. For example, suppose that the researcher does not account for gender when estimating $\hat{L}_{it}$ and does not include it as a control variable in (5). Then the estimation of γ will be biased if women and men have both different life expectancies and different probabilities of dying following exposure to PM 2.5.

Equation (5) succinctly summarizes the key challenge researchers face when estimating the effect of pollution exposure on life-years lost. Any unobserved factor that is positively correlated with both remaining life expectancy and the probability of dying following exposure to pollution will cause upward bias in the estimate of γ. This is problematic because populations with low levels of remaining life expectancy, such as the elderly, are often more vulnerable in general and thus also likely to be more susceptible to dying from pollution exposure.

We address this challenge by harnessing the comprehensive health data available in the Medicare dataset to generate very precise predictions of counterfactual life expectancy. In other words, we minimize the magnitude of the measurement error represented by u_{it} in (4). Note that it is not necessary to eliminate all measurement error to eliminate the bias; it suffices to eliminate just the portion of the measurement error that is correlated with the heterogeneous treatment effect, v_i. While this is ultimately untestable, if sequentially including additional predictors of mortality has no effect on the estimate of life-years lost, then one can plausibly argue that adding other, potentially unobserved, predictors will have no effect either.

This line of reasoning is similar to using observables to make inferences about unobservables (see, for example, Altonji, Elder, and Taber 2005). For example, consider estimating the OLS model

$$y = \alpha + \beta_1 x_1 + e$$

A researcher may be concerned that the estimate of β_1 suffers from omitted variable bias. Suppose the researcher sequentially adds covariates x_2, x_3 and x_4 to the model and that doing so significantly increases the explanatory power of the model (as measured by R^2), but has no effect on the estimate $\hat{\beta}_1$. Then for another potential covariate x_5 to substantially affect $\hat{\beta}_1$, x_5 would have to be uncorrelated with previously added covariates x_2, x_3, and x_4 but correlated with x_1. If it seems unlikely that such an additional covariate exists, the researcher may plausibly argue that $\hat{\beta}_1$ provides a consistent estimate of β_1.

To test for and eliminate omitted variable bias in our setting, we generate five sets of estimates of remaining life expectancy using an increasingly detailed set of controls. The sets of covariates are as follows: (1) none; (2) age; (3) age and sex; (4) age, sex, and chronic conditions that are likely to be exacerbated by air pollution; and (5) an even more comprehensive set of covariates, including historic medical spending and the beneficiaries' location. We then estimate an analogue of equation (4) using each of these five estimates as the dependent variable. Importantly, because the fifth set of estimates includes a large set of covariates, we argue that it is unlikely that further, potentially unobserved, factors would affect the results.

Integrating this framework into the empirical strategy presented earlier is straightforward. The only difference is that the dependent variable in (1) is now $\hat{L}_{cdmy}$, the estimated daily number of life-years lost per capita in county c. The outcome 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.

In practice, we estimate counterfactual life expectancies for decedents with a survival analysis model. The model is estimated using the 2002 Medicare cohort, which we observe from 2002-2011.[11] Deaths are therefore censored after 10 years. We assume that survival rates are governed by the Gompertz distribution, which is commonly used to model mortality (Cameron and Trivedi, 2005):

$$S_j(t) = \exp\left[-\lambda_j \gamma^{-1} \exp[\gamma t_j - 1]\right]$$

We parametrize this distribution by assuming that λ_j depends on a vector of individual characteristics, x_j:

[11] We do not use earlier cohorts because the Medicare variables denoting the presence of chronic conditions, which are strong predictors of mortality, become increasingly unreliable in earlier years.

$$\lambda_j = \exp[x_j'\delta]$$

We employ five different sets of characteristics as outlined above: (1) none; (2) age; (3) age and sex; (4) age, sex, and chronic conditions that are likely to be exacerbated by air pollution; and (5) an even more comprehensive set of covariates, including historic medical spending and the beneficiaries' location. (See the appendix for additional details.) We then use the estimate $\hat{\delta}$ to predict life expectancies for all Medicare FFS decedents in the 2001-2011 time period based on their reported characteristics.

V. RESULTS

Mortality and health care utilization

Table 2 reports OLS estimates from equation (1) for different age groups. The dependent variable is the daily mortality rate for different age groups in the Medicare population. Column (1) estimates that a 1 microgram per cubic meter increase in PM 2.5 levels increases the daily mortality rate by 0.099 per million for Medicare beneficiaries over the age of 65. This corresponds to a statistically significant increase of 0.075 percent $(= 0.099/132)$. The corresponding estimate for a *one standard deviation* ($7.34\ \frac{\mu g}{m^3}$) increase in PM 2.5 is 0.73 deaths per million or 0.55 percent of the daily mean. Columns (2)-(5) illustrate how our estimated effect varies by age. The absolute increase in the death rate rises with the age group, with the oldest elderly experiencing 0.343 additional deaths per million when PM 2.5 increases by $1\ \frac{\mu g}{m^3}$.

Table 3 presents the corresponding IV estimates for the mortality specification (Panel A) and also shows the mortality impacts on the fee-for-service population in years 2001-2011 (Panel B), which is our sample for looking at health utilization and life years lost. Compared to the OLS results, the magnitudes of the estimated effects for IV are substantially larger, suggesting that OLS estimation may suffer from measurement error bias. According to the IV estimates, a $1\ \frac{\mu g}{m^3}$ increase in PM-2.5 increases daily mortality by an estimated 0.332 per million, or 0.25% of the daily mortality rate. The corresponding estimate for a one standard deviation increase is 1.8% of the average daily death rate. In Columns (2)-(6), we see that the absolute increases in mortality are much larger for older age groups. For example, 85+ year-olds see increases in mortality of 1.07 deaths per million while the youngest elderly (65-69 year-olds) experience an increase on 0.10 deaths per million, about ten times smaller. However, because the average mortality rate is also much higher for the older elderly, the *relative* mortality rates across age groups are nearly constant, ranging from 0.22% for the 65-69 and 80-84 year olds to 0.27% for the oldest group. We will return to this point when we estimate the number of life-years-lost rather than just lives lost.

Panel B shows the IV mortality estimates for the fee-for-service beneficiaries in the years 2001-2011. Across the board, the per-unit impact of PM 2.5 on mortality is larger for these individuals. This is consistent with the conventional wisdom and empirical evidence that the fee-for-service population is generally sicker that the average Medicare beneficiary. Notably, the FFS mean mortality rate is also higher for each age group and overall.

Next, we proceed to look at hospitalizations and hospital spending. Columns (1)-(5) of Table 4 show the daily OLS estimates for log total spending, hospital admissions per 1,000 beneficiaries, spending per admission, spending per beneficiary, and log total length of stay, respectively. Higher PM 2.5 concentrations are associated with higher total and per-beneficiary spending, a higher admission rate, and a higher total length of stay. However, Table 5, which shows the corresponding IV estimates, contradicts the OLS results. When we instrument for PM 2.5, we find no significant increase in total spending, the admission rate, spending per beneficiary or the total length of stay. There is a small but significant *decline* in the amount spent per admission, suggesting that the marginal admit may be healthier. These results suggests that deaths caused by PM 2.5 may be offsetting some hospital admissions and demonstrates the importance of instrumenting for pollution concentrations.

The main concern with interpreting our estimates above as the causal effects of PM 2.5 is that other air pollutants like ozone (O_3), carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen dioxide (NO_2) can be co-transported with fine particulate matter (PM 2.5). In principle, we can use our set of instruments to estimate a regression with five endogenous variables (PM 2.5, O_3, CO, SO_2, and NO_2). Unfortunately, a key data limitation is that many counties only track a subset of these pollutants, resulting in substantial sample size reductions for each additional pollutant added to the specification. Specifically, including all five pollutants would reduce our sample size from over 2 million observations to slightly over 350 thousand observations. Thus, while there is independent variation in how wind directions predict concentrations of these different pollutants, we lose a lot of power and (potentially) generalizability when we estimate an equation that includes all five as endogenous variables.

Nonetheless, it is useful to consider how our estimates of the impact of PM 2.5 would be affected if we instrumented for and included additional pollutants in the regression. We do this by restricting the sample to county-days where readings for all five pollutants are available and then add one pollutant at a time. The results are shown in Table 6. Despite a significant reduction in the F-statistics, 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.

Our main empirical specification employs 500 instruments and our F-statistics easily exceed 1000 in many cases. However, because there is no well-accepted threshold for a weak instrument when the number of instruments is in the hundreds, we undertake two different sets of robustness exercises to ensure that our estimates are not driven by weak instrument bias. First, we estimate our model using LIML, which is approximately median unbiased even in the presence of weak instruments, rather than 2SLS. Those estimates, presented in Table 7, are very similar to the 2SLS estimates presented in Table 3. Second, we conduct a placebo exercise where we generate a set of random wind directions, $WINDDIR\_PLACEBO_{cdmy}^{60b}$, and use those in our first stage instead of the actual wind direction, $WINDDIR_{cdmy}^{60b}$. Those results, shown in Table 8, generate very small F-statistics, on the order of 5-6. This provides further evidence that our estimates do not suffer from weak instrument bias.

The social value of reducing mortality

Table 9 displays estimates of our baseline specification, equation (1), when the outcome variable is L_{cdmy}, the estimated daily life-years lost (per 1 million beneficiaries) in county c. The estimation sample in this table is limited to the fee-for-service population because some of the specifications require data on chronic conditions to estimate counterfactual life expectancy. For reference, Column 1 also shows the estimated effects of PM 2.5 on mortality in the fee-for-service population.

In Columns 2-6, we progressively increase the number of covariates that are used to predict counterfactual life expectancy. If people who die as a result of pollution exposure have a lower counterfactual life expectancy than those who remain alive, then we expect these estimates to decrease as we add more and more predictors.

Column 2 displays results when a decedent's counterfactual life expectancy is set equal to the mean for the FFS population (11.4 years). Here, we estimate that a one-unit increase in levels of PM 2.5 reduces life-years lost by 5.2 years per million FFS beneficiaries. Column 3 displays estimates when a decedent's counterfactual life expectancy is modeled as a function of her age. This causes the magnitude of the coefficient on PM 2.5 to decrease by about 30 percent, implying that older beneficiaries, who have lower life expectancies, are also more likely to be killed by PM 2.5.

Columns 4 and 5 of Table 8 sequentially incorporate sex and the presence of chronic conditions, respectively, when estimating counterfactual life expectancy. Accounting for gender has little effect, but accounting for both gender and chronic conditions reduces the estimate by about 20 percent. Nonetheless, the estimate is still large and highly significant, showing that a one-unit increase in daily PM 2.5 leads to a loss of 2.9 life-years per million beneficiaries.

Finally, we estimate a model that incorporates medical spending, the beneficiary's location, and a host of other variables into the life expectancy estimation. We use a machine learning algorithm (LASSO) to select 60 variables that are most predictive of life expectancy, out of 780. (See the appendix for details.) This final estimate, shown in Column 6, is again 20 percent smaller than the preceding estimate. It indicates that a one-unit increase in PM 2.5 leads to a loss of about 2.4 life years per million beneficiaries. This raises the question of whether there are additional (potentially unobservable) characteristics that would further lower the estimated life years lost. We plan to explore this possibility in future iterations of this paper.

We also calculate the mean life years lost per decedent ("LYL per decedent") by dividing the average life years lost by the average daily mortality rate in this population (153 per million). This provides a performance metric for the counterfactual life expectancy prediction models: better models should predict lower remaining life expectancy for decedents. Finally, we also compute the life years lost per *complier* ("LYL per complier") by dividing the coefficients in Columns 2-6 by PM 2.5's estimated effect on the mortality rate from Column 1. For example, Column (6) estimates that individuals dying from pollution exposure lost 5.2 years of life, on average.

Next, we estimate the value of a 3.65 unit decrease in PM 2.5, which roughly corresponds to the average decrease experienced nationwide between 1999 and 2011 as shown in Figure 1. The estimate in column (6) of Table 9 implies that such a decrease saved 129,673 life-years annually among the 41 million Medicare beneficiaries alive in 2011.[12] If we value those life-years at $100,000 each, then this implies annual health benefits of about $13 billion. If we ignore heterogeneity in the effect of pollution on mortality and instead employ the estimate from Column 2 of Table 9, this estimate more than doubles, to $29 billion. This demonstrates the importance of properly accounting for "harvesting" when calculating the mortality benefits of reductions in air pollution.

Finally, we note that our estimate of $13 billion ignores any quality-of-life benefits associated with a reduction in air pollution, e.g., a reduction in non-lethal asthma attacks. We also do not estimate the health benefits that accrue to the non-Medicare population. Accounting for these two benefits would increase our estimate.

VI. CONCLUSION

Understanding how pollution affects health and health care spending is essential for crafting efficient environment policy, such as Pigouvian pricing based on health externalities. Given the public

[12] The exact calculation is $2.374 \times 365 \times 41 \times 3.65$

burden of high and increasing Medicare spending, and it is important for effective health policy to understand how pollution affects both health and health care spending among the elderly, and how these effects differ by demographics and health status.

This paper sheds light on these issues by estimating the effect of pollution on mortality and other health outcomes using a novel identification strategy based on exogenous variation in wind direction. This is accomplished through a novel data effort linking daily pollution and other climatic variables to detailed, administrative Medicare records on all Medicare beneficiaries from 1999-2011. We find significant effects of pollution on mortality, but not on health care spending and hospitalizations. Although other pollutants may also matter for these outcomes, we demonstrate that they do not drive our main PM 2.5 mortality results. Finally, we use the rich Medicare data to estimate counterfactual survival and estimate that the reduction in PM 2.5 experienced nationwide since 1999 generates $13 billion annually in mortality benefits alone.

REFERENCES

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

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools." Journal of political economy 113, no. 1 (2005): 151-184.

Angrist, Joshua and Jorn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2008.

Bell, Michelle L., Aidan McDermott, Scott L. Zeger, Jonathan M. Samet, and Francesca Dominici. "Ozone and short-term mortality in 95 US urban communities, 1987-2000." Jama 292, no. 19 (2004): 2372-2378.

Cameron, A. Colin, and Pravin K. Trivedi. "Microeconometrics: methods and applications." Cambridge university press, 2005.

Currie, Janet, and Matthew Neidell. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" The Quarterly Journal of Economics 120, no. 3 (2005): 1003-1030.

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

Currie, J. and R. Walker (2011). "Traffic Congestion and Infant Health: Evidence from E-ZPass." American Economic Journal: Applied Economics, 3(1): 65-90.

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

Deschênes, Olivier, and Michael Greenstone. "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US." American Economic Journal: Applied Economics 3, no. 4 (2011): 152-85.

Dockery, D., C. Pope, X. Xu, J. Spengler, J. Ware, M. Fay, B. Ferris Jr, and F. 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 R. Sunstein. "Particulate matter matters." Science (New York, NY) 344, no. 6181 (2014): 257.

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

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

EPA. “Integrated Science Assessment of Ozone and Related Photochemical Oxidants.” US Environmental Protection Agency, Washington DC (2013).

Graff Zivin, Joshua, and Matthew Neidell. "Environment, Health, and Human Capital." Journal of Economic Literature 51, no. 3 (2013): 689-730.

Gryparis, Alexandros, Bertil Forsberg, Klea Katsouyanni, Antonis Analitis, Giota Touloumi, Joel Schwartz, Evangelia Samoli et al. "Acute effects of ozone on mortality from the “air pollution and health: a European approach” project."American journal of respiratory and critical care medicine 170, no. 10 (2004): 1080-1087.

Heutel, Garth, Nolan H. Miller, and David Molitor. "The Impact of Air Temperature on Morbidity, Mortality, and Health Care Cost in the Medicare Population." Working Paper, 2015.

Herrnstadt, Evan, and Erich Muehlegger. *Air Pollution and Criminal Activity: Evidence from Chicago Microdata*. NBER Working Paper #21787, 2015.

Ito, Kazuhiko, Samantha F. De Leon, and Morton Lippmann. "Associations between ozone and daily mortality: analysis and meta-analysis." Epidemiology16, no. 4 (2005): 446-457.

Kundu, S., and Stone, E. A. (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, F., L. M. Neas, D. W. Dockery, and J. 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.

Pope III, C. Arden, and Douglas W. Dockery. "Health effects of fine particulate air pollution: lines that connect." Journal of the Air & Waste Management Association 56, no. 6 (2006): 709-742.

Pope, A. C. I., M. J. Thun, M. Namboodiri, D. Dockery, J. Evans, F. Speizer, and C. J. Heath (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.

Ruidavets, Jean-Bernard, Maxime Cournot, Sylvie Cassadou, Michel Giroux, Mariam Meybeck, and Jean Ferrières. "Ozone air pollution is associated with acute myocardial infarction." Circulation 111, no. 5 (2005): 563-569.

Samet, Jonathan M., Francesca Dominici, Frank C. Curriero, Ivan Coursac, and Scott L. Zeger. "Fine particulate air pollution and mortality in 20 US cities, 1987–1994." New England journal of medicine 343, no. 24 (2000): 1742-1749.

Wolfram Schlenker & W. Reed Walker. "Airports, Air Pollution, and Contemporaneous Health," *Review of Economic Studies*, 83, no. 2 (2016): 768-809.

Zanobetti, Antonella, and Joel Schwartz. "Mortality displacement in the association of ozone with mortality: an analysis of 48 cities in the United States." American Journal of Respiratory and Critical Care Medicine 177, no. 2 (2008): 184-189.

Zeger, Scott L., Francesca Dominici, and Jonathan Samet. "Harvesting-resistant estimates of air pollution effects on mortality." *Epidemiology* (1999): 171-175.

FIGURES

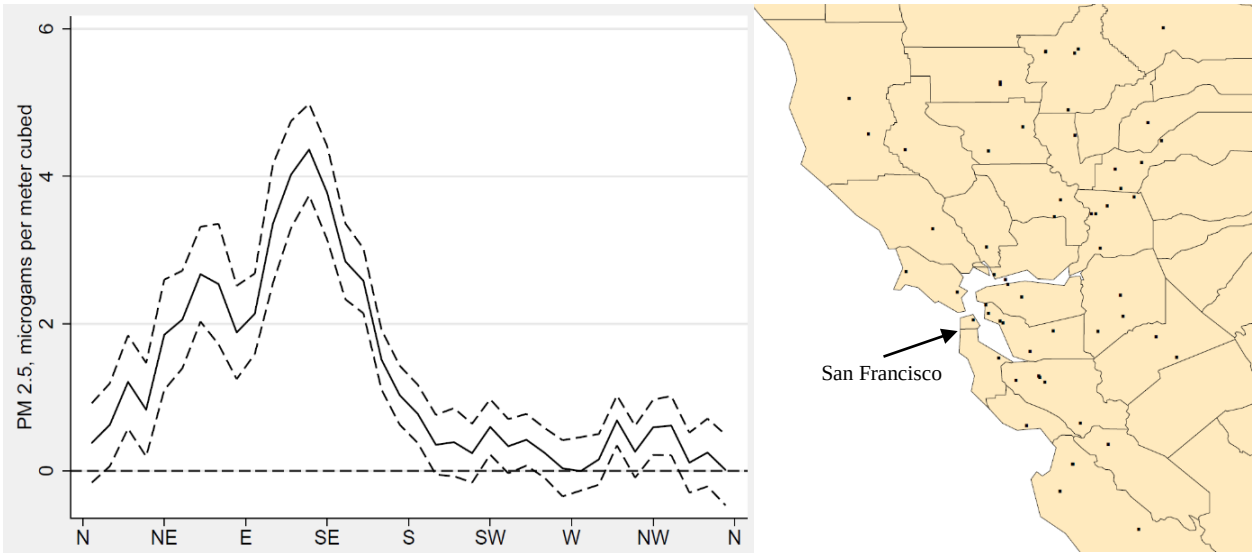

Figure 1. Relationship between daily average wind direction and PM 2.5 concentrations in and around the Bay Area, CA. The left panel shows regression estimates where the dependent variable is the average daily PM 2.5 concentration and the key independent variables are a set of indicators for the daily 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, as discussed in Section IV of the main text. The dashed lines represent 95% confidence intervals based on robust standard errors. The right panel shows the location of the PM 2.5 pollution monitors in the Bay Area that provided the pollution measures for this regression.

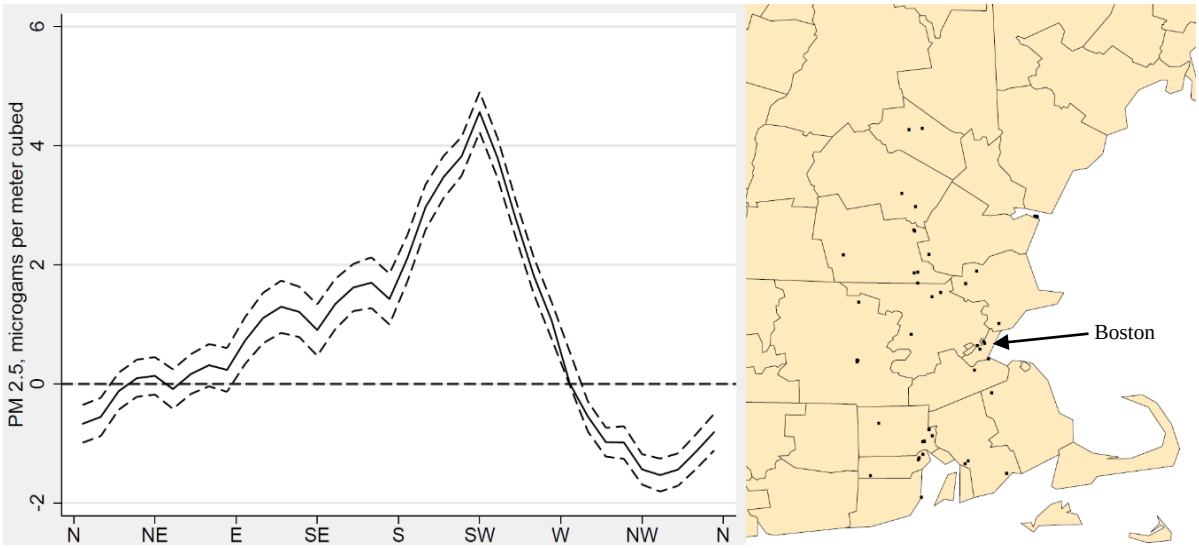

Figure 2. Relationship between daily average wind direction and PM 2.5 concentrations in and around the Boston Area, MA. The left panel shows regression estimates where the dependent variable is the average daily PM 2.5 concentration and the key independent variables are a set of indicators for the daily 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, discussed in Section IV. The dashed lines represent 95% confidence intervals based on robust standard errors. The right panel shows the location of the PM 2.5 pollution monitors in the Boston Area that provided the pollution measures for this regression.

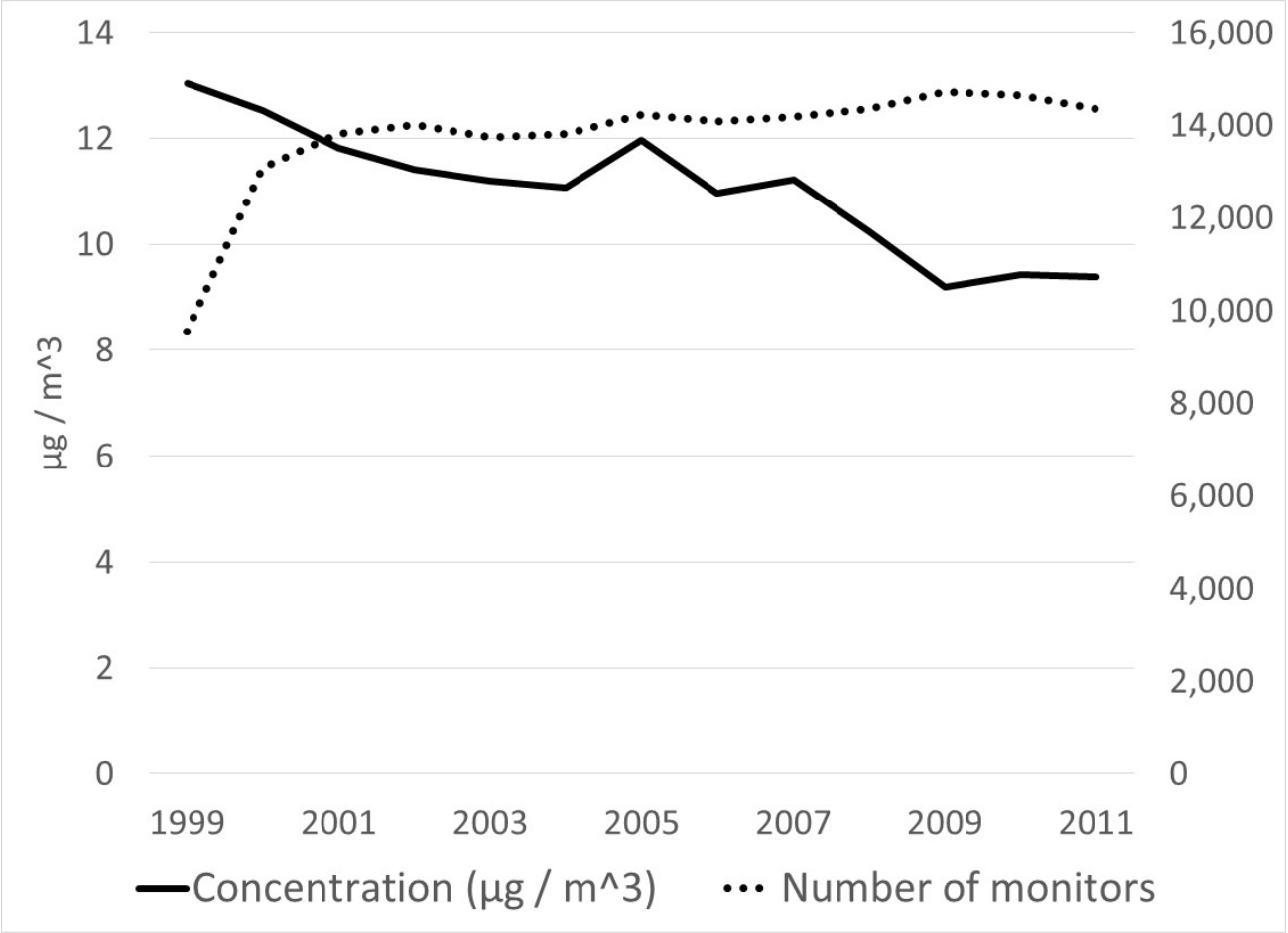

Figure 3. Trends in PM 2.5 air pollution, 1999-2011. Figure displays annual county means.

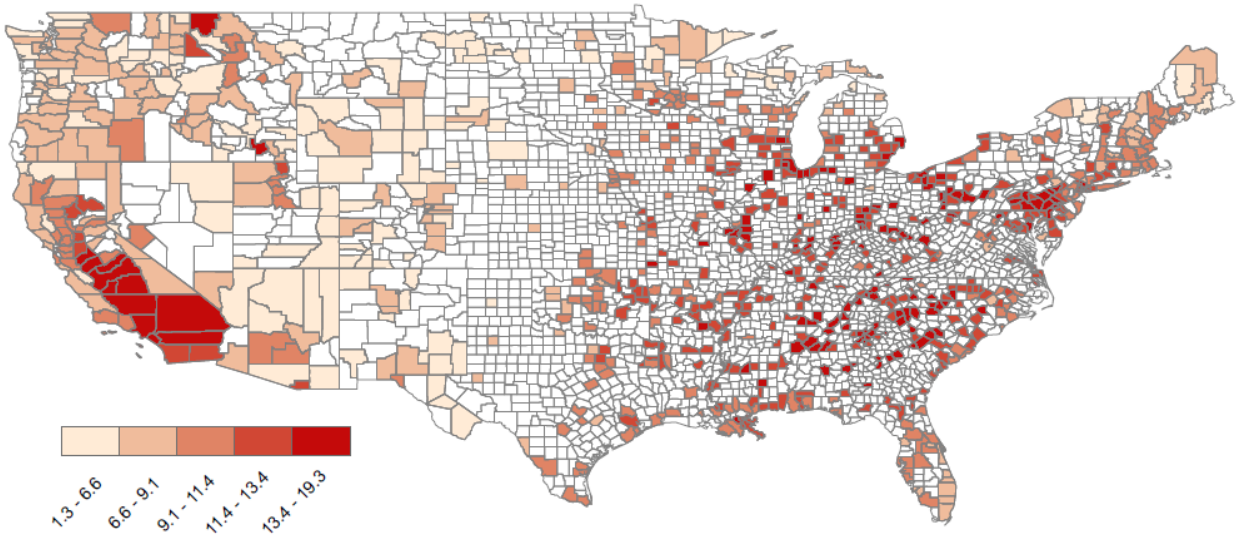

Figure 4. Average PM 2.5 levels in the United States by county, 1999-2011. Units are micrograms per cubic meter. Counties with no data are shaded white.

TABLES

Table 1: Summary statistics, 1999-2011

	Mean	Std. dev.	Obs.
PM 2.5 ($\frac{\mu g}{m^3}$)	10.9	7.34	2,221,481
Number of persons, all ages	43,589	71,984	2,221,481
Number of persons, 65-69	11,424	18,585	2,221,481
Number of persons, 70-74	10,380	17,144	2,221,481
Number of persons, 75-79	8,758	14,637	2,221,481
Number of persons, 80-84	6,728	11,295	2,221,481
Number of persons, 85+	6,299	10,724	2,221,481
1-day mortality rate, all ages	132.18	142.73	2,221,481
1-day mortality rate, 65-69	46.56	155.8	2,221,481
1-day mortality rate, 70-74	69.43	222.86	2,221,481
1-day mortality rate, 75-79	109.7	279.66	2,221,481
1-day mortality rate, 80-84	179.13	443.46	2,221,481
1-day mortality rate, 85+	393.69	657.8	2,221,481
Total daily hospital spending	422,865	790,622	2,092,951
Number of fee-for-service beneficiaries	26,171	37,983	2,092,951
Admissions per 1,000 FFS beneficiaries	1.50	0.74	2,092,951
Spending per admission	9,637	4,410	2,032,365
Spending per beneficiary	14.51	9.56	2,092,951
Total length of stay, days	427	763	2,092,951

Notes: unit of observation is county-day. Mortality rate is number of deaths per 1,000,000 beneficiaries.

Table 2: Daily PM 2.5 concentrations and elderly mortality, OLS

	(1) 65+	(2) 65-69	(3) 70-74	(4) 75-79	(5) 80-84	(6) 85+
PM 2.5 ($\mu g/m^3$)	0.099***	0.037***	0.046***	0.068***	0.132***	0.343***
	(0.013)	(0.008)	(0.010)	(0.014)	(0.022)	(0.045)
Dep. var. mean	132	46.6	69.5	110	179	394
Observations	2,228,410	2,228,410	2,228,410	2,228,410	2,228,410	2,228,410
Adjusted R-squared	0.089	0.024	0.026	0.026	0.025	0.036

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. Mortality is measured as the number of deaths per 1,000,000 people in the relevant age group. All regressions include county, month-by-year, and state-by-month fixed effects, as well as flexible controls for temperatures, precipitation, and wind speed. Estimates are weighted by the relevant population.

Table 3: Daily PM 2.5 concentrations and mortality, IV

	(1) 65+	(2) 65-69	(3) 70-74	(4) 75-79	(5) 80-84	(6) 85+
Panel A: all beneficiaries						
PM 2.5 ($\mu g/m^3$)	0.332***	0.091***	0.165***	0.256***	0.415***	1.069***
	(0.028)	(0.025)	(0.031)	(0.039)	(0.064)	(0.107)
F-statistic	2240	1958	2208	2422	2424	2262
Dep. var. mean	132	46.6	69.4	110	179	394
Observations	2,221,464	2,221,464	2,221,464	2,221,464	2,221,464	2,221,464
Panel B: fee-for-service beneficiaries						
PM 2.5 ($\mu g/m^3$)	0.458***	0.192***	0.216***	0.292***	0.479***	1.193***
	(0.035)	(0.045)	(0.041)	(0.055)	(0.073)	(0.127)
F-statistic	1788	1640	1713	1831	1916	1937
Dep. var. mean	153	52.6	72.2	112	183	404
Observations	2,041,152	2,041,152	2,041,152	2,041,152	2,041,152	2,041,152

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. Mortality is measured as the number of deaths per 1,000,000 people in the relevant age group. All regressions include county, month-by-year, and state-by-month fixed effects, as well as flexible controls for temperatures, precipitation, and wind speed. Estimates are weighted by the relevant population.

Table 4: Daily PM 2.5 concentrations and hospitalization outcomes, OLS

	(1) Log spending	(2) Admit rate	(3) Spending per admit	(4) Spending per bene.	(5) Log length of stay
PM 2.5 ($\mu g/m^3$)	0.001***	0.0017***	1.6**	0.018***	0.002***
	(0.000)	(0.0002)	(0.7)	(0.003)	(0.000)
Dep. var. mean	12.2	1.50	9,637	14.5	5.11
Observations	2,031,719	2,092,917	2,032,294	2,092,917	2,032,294
Adjusted R-squared	0.834	0.273	0.500	0.361	0.783

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. is the number of hospitalizations per 1,000 beneficiaries. All regressions include county, month-by-year, and state-by-month fixed effects, as well as flexible controls for temperatures, precipitation, and wind speed. Per-person estimates are weighted by the relevant population.

Table 5: Daily PM 2.5 concentrations and hospitalization outcomes, IV

	(1) Log spending	(2) Admit rate	(3) Spending per admit	(4) Spending per bene.	(5) Log length of stay
PM 2.5 ($\mu g/m^3$)	-0.0001	0.006	-3.1***	-0.000	-0.0007**
	(0.0003)	(0.004)	(1.1)	(0.005)	(0.0003)
F-statistic	832	1822	1872	1822	834
Dep. var. mean	12.09	15.0	9,637	14.5	5.11
Observations	2,031,702	2,092,903	2,032,277	2,092,903	2,032,277

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. is the number of hospitalizations per 1,000 beneficiaries. All regressions include county, month-by-year, and state-by-month fixed effects, as well as flexible controls for temperatures, precipitation, and wind speed. Per-person estimates are weighted by the relevant population.

Table 6: Daily PM 2.5, other pollutants, and elderly mortality, IV

	(1)	(2)	(3)	(4)	(5)
			Panel A: all beneficiaries		
PM 2.5	0.244***	0.141**	0.115**	0.149***	0.170**
	(0.045)	(0.057)	(0.056)	(0.057)	(0.071)
SO2		0.510***	0.458***	0.456***	0.491***
		(0.162)	(0.166)	(0.170)	(0.178)
CO			0.006***	0.004*	0.005**
			(0.002)	(0.002)	(0.003)
O3				-0.064	-0.067
				(0.051)	(0.051)
NO2					-0.073
					(0.110)
F-statistic	87.2	40.2	26.3	21.7	19.6
Dep. var. mean	131	131	131	131	131
Observations	350,883	350,883	350,883	350,883	350,883
			Panel B: fee-for-service beneficiaries		
PM 2.5	0.364***	0.294***	0.289***	0.373***	0.412***
	(0.060)	(0.073)	(0.074)	(0.080)	(0.101)
SO2		0.355	0.350	0.342	0.427
		(0.235)	(0.238)	(0.245)	(0.272)
CO			0.001	-0.003	-0.000
			(0.003)	(0.004)	(0.004)
O3				-0.149*	-0.158**
				(0.080)	(0.080)
NO2					-0.148
					(0.181)
F-statistic	82.7	36.9	24.2	19.3	16.6
Dep. var. mean	156	156	156	156	156
Observations	310,947	310,947	310,947	310,947	310,947

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. Dependent variable is the all-age daily mortality rate, measured as the number of deaths per 1,000,000 people. All regressions include county, month-by-year, and state-by-month fixed effects, as well as flexible controls for temperatures, precipitation, and wind speed. Estimates are weighted by the county's Medicare population.

Table 7: Daily PM 2.5 concentrations and mortality, LIML IV

	(1) 65+	(2) 65-69	(3) 70-74	(4) 75-79	(5) 80-84	(6) 85+
PM 2.5 ($\mu g/m^3$)	0.334***	0.093***	0.164***	0.260***	0.410***	1.072***
	(0.028)	(0.025)	(0.030)	(0.039)	(0.063)	(0.105)
F-statistic	430	417	424	434	441	446
Dep. var. mean	132	46.6	69.5	110	179	394
Observations	2,228,410	2,228,410	2,228,410	2,228,410	2,228,410	2,228,410

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. Mortality is measured as the number of deaths per 1,000,000 people in the relevant age group. All regressions include county, month-by-year, and state-by-month fixed effects, as well as flexible controls for temperatures, precipitation, and wind speed. Estimates are weighted by the relevant population.

Table 8: Daily PM 2.5 concentrations and mortality, placebo IV

	(1) 65+	(2) 65-69	(3) 70-74	(4) 75-79	(5) 80-84	(6) 85+
PM 2.5 ($\mu g/m^3$)	0.055	0.323	0.065	-0.543**	0.157	0.748
	(0.206)	(0.346)	(0.235)	(0.273)	(0.545)	(0.576)
F-statistic	5.662	5.534	5.899	5.919	5.657	5.130
Dep. var. mean	132	46.6	69.5	110	179	394
Observations	2,228,410	2,228,410	2,228,410	2,228,410	2,228,410	2,228,410

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. Mortality is measured as the number of deaths per 1,000,000 people in the relevant age group. All regressions include county, month-by-year, and state-by-month fixed effects, as well as flexible controls for temperatures, precipitation, and wind speed. Estimates are weighted by the relevant population.

Table 9: Daily PM 2.5 concentrations and life years lost, IV

	(1) All-age mort.	(2) None	(3) Age	(4) Age, sex	(5) Age, sex, chron. cond.	(6) LASSO
PM 2.5 ($\mu g/m^3$)	0.458***	5.242***	3.640***	3.609***	2.948***	2.374***
	(0.035)	(0.405)	(0.309)	(0.306)	(0.256)	(0.224)
Dep. var. mean	153	1750	1254	1234	1007	868
LYL per decedent	.	11.438	8.199	8.064	6.580	5.673
LYL per complier	.	11.438	7.942	7.876	6.433	5.181
F-statistic	415	415	415	415	415	415
Observations	2,041,152	2,041,152	2,041,152	2,041,152	2,041,152	2,041,152

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. Units of the dependent variables in columns 2-6 are life years lost per 1,000,000 fee-for-service beneficiaries. Column 1 shows the raw mortality rate estimates. Column 2-5 headings display what variables were used to predict counterfactual life expectancy. Column 6 uses LASSO to select variables that predict life expectancy. Life years lost per decedent (LYL per decedent) is calculated by dividing the average life years lost in the sample by the average mortality rate. Life years lost per complier (LYL per complier) is calculated by dividing the estimated coefficient on life years lost by the PM 2.5 mortality coefficient 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. Estimates are weighted by the number of fee-for-service beneficiaries.

APPENDIX

Estimating counterfactual life expectancy

We estimate counterfactual life expectancies by estimating a parametric survival analysis model via maximum likelihood. We assume that survival rates are governed by the Gompertz distribution, which is commonly used to model mortality:

$$S(t_i|\lambda_i,\gamma) = \exp[-\lambda_i\gamma^{-1}\exp[\gamma t_i - 1]]$$

The variable t_i represents the length of time until death. This distribution is parameterized by assuming λ_i depends on a vector of individual characteristics, x_i:

$$\lambda_i = \exp[x_i'\beta]$$

Given an estimate $\hat{\beta}$, one can calculate life expectancy using the formula for the mean of the Gompertz distribution:

$$\hat{L}_i = \frac{1}{\gamma}\exp[\hat{\lambda}_i/\gamma]\int_{\hat{\lambda}_i/\gamma}^{\infty}\exp[-x]/x\,dx$$

We estimate this model using data from the 2002 cohort of Medicare beneficiaries. We observe all deaths that occurred among this cohort up through the end of 2011. Deaths are therefore censored after 10 years.[1] To ensure that we have accurate measures of beneficiaries' chronic conditions, we limit the sample to Medicare beneficiaries who as of January 1, 2002 (1) were 67 years of age or older and (2) had been continuously enrolled in fee-for-service Medicare for at least two years. For computational ease, we further limit our analysis to a random 10 percent sample of these beneficiaries. Our final estimation sample consists of 2,421,724 individuals.

Columns (2)-(5) of Table 9 are based on estimations that include the following sets of characteristics as part of x_i: (1) none; (2) age; (3) age and gender; and (4) age, gender, and indicator variables for each of the following chronic conditions: acute myocardial infarction, atrial fibrillation, chronic obstructive pulmonary disease, ischemic heart disease, stroke/transient ischemic attack, and asthma.[2]

The Medicare dataset includes thousands of additional variables, any of which could be significant predictors of mortality. Including all of them is unwise because some will be significant predictors of survival for the 2002 cohort just by chance, even if they are not good predictors of survival in general.

Recent advances in machine learning techniques allow researchers to overcome this challenge. These methods are well-suited to settings with a very large number of potential regressors and where the researcher cares primarily about predictive accuracy. The counterfactual life expectancy that forms the basis

[1] Although earlier cohorts are observable for a longer period of time in our data, we do not use them in this analysis because the Medicare variables denoting the presence of pre-existing chronic conditions, which are strong predictors of mortality, are nonexistent or unreliable in earlier years.

[2] According to the EPA, PM 2.5 "aggravate[s] heart and lung diseases and [has] been linked to effects such as: cardiovascular symptoms; cardiac arrhythmias; heart attacks; respiratory symptoms; asthma attacks; and bronchitis" (https://www3.epa.gov/pm/designations/basicinfo.htm).

of the estimate in Column (6) of Table 9 is based on predictors that were first selected as significant by LASSO, a commonly employed machine learning algorithm.

Our estimation procedure consists of two steps. First, we identify significant predictors of five-year survival by estimating the following LASSO regression with 5-fold cross validation using the 2002 Medicare cohort described above:

$$Y_i = x_i'\delta + e_{it}$$

The outcome variable, Y_i, is an indicator variable equal to 1 if the Medicare beneficiary was alive on January 1, 2007 and 0 otherwise.[3] The vector x_i contains 728 regressors, including:

1. Indicator variables for age group, race, gender, and two-digit zipcode of residence
2. Indicator variables for 27 different chronic conditions and all pairwise interactions
3. Interaction of chronic condition indicator variables with age
4. Interaction of chronic condition indicator variables with race
5. Fifth order polynomials in (normalized) total spending, number of hospital visits, number of hospital ER visits, number of acute stays, number of inpatient stays, number of outpatient stays, number of hospital stays, number of ambulatory surgical events, and number of dialysis events.

More formally, we minimize the following objective function:

$$f(\delta) = (Y_i - x_i'\delta)^2 + \lambda|\delta|$$

where $|\delta|$ is the l_1 norm of δ, the vector of estimated coefficients.

Of these variables, the LASSO regression identified 60 as being statistically significant predictors of five-year survival. These variables were then included as regressors in the estimation of the Gompertz survival model described above.

[3] We use five-year survival because it is half of 10, the length of our observation period for the 2002 cohort. Using survival rates of different lengths does not affect results.