THE MORTALITY AND MEDICAL COSTS OF AIR POLLUTION: 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

August 2016

Abstract. We estimate the effect of acute air pollution exposure on mortality, life-years lost, and health care utilization among the US elderly. We address endogeneity and measurement error using a novel instrument for pollution exposure: changes in the regional wind direction, which strongly predict changes in fine particulate matter (PM 2.5) pollution concentrations. Using detailed administrative data on the universe of Medicare beneficiaries, we find that an increase in daily PM 2.5 exposure significantly increases mortality, hospitalizations, and inpatient spending over the next three days, and that these effects are not explained by co-transported pollutants like ozone and carbon monoxide. Finally, we develop a new method for estimating the years of life lost due to pollution using detailed data on health and healthcare use. The mortality benefit of a reduction in PM 2.5 is much smaller than a calculation that controls only for age and gender or ignores heterogeneity altogether.

* We thank Allen C. Basala, Alex Hollingsworth, Christos Makridis, and participants at the University of Illinois Research Lunch, the Midwestern Health Economics Conference, and the AERE Summer Conference for helpful comments. 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.

I. INTRODUCTION

Guided by the widely accepted belief that air pollution negatively affects human health, many countries have gradually tightened air pollution standards. Precisely quantifying 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. However, estimating the causal effect of pollution on health is complicated due to well-documented challenges, including separately identifying the effects of different pollutants, endogeneity and measurement error. Moreover, quasi-experimental studies that identify a plausibly exogenous source of pollution variation are typically confined to a small geographic and temporal scale, 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 typically focus on morbidity (e.g., hospitalizations) or infant mortality.[2]

This paper estimates the causal impact of pollution on mortality and medical cost using a dataset that matches detailed data on the universe of Medicare beneficiaries with pollution and atmospheric data over a 13 year period. We overcome the identification challenges described above by exploiting variation in fine particulate matter (PM 2.5) caused by changes in daily wind direction. A key innovation of our study relative to other quasi-experimental approaches is demonstrating that it is not necessary to isolate a particular pollution source to estimate the causal impact of pollution on health. Because the wind carries pollutants over long distances and is plausibly exogenous, it is a valid instrument regardless of where the pollution is coming from. This approach allows us to analyze a long panel for a large group of beneficiaries living over a diverse set of geographies, increasing statistical power to detect rare events, and it naturally accommodates multiple sources, which allows us to separately identify the effects of different pollutants. [3]

Our identification assumption is simple and intuitive: changes in wind direction are unrelated to mortality except through pollution. Since the pattern of pollution transport varies spatially, we allow the relationship between pollution and wind direction to vary across 100 groups of nearby pollution monitors. To increase the plausibility of our identification assumption, we flexibly control for other climatic factors,

---

[2] See, for example, Chay and Greenstone (2003), Currie and Neidell (2005), Currie et al. (2009), Deschenes et al. (2016).

[3] For a review of the literature on short-run pollution fluctuations and mortality in non-experimental settings, see Pope (2000). Chen et. al. (2013) and Anderson (2016) study the effect of long-run pollution variation on mortality in quasi-experimental settings.

including minimum and maximum temperature, precipitation, and wind speed, as well as spatial and temporal fixed effects. Although we focus on fine particulate matter (PM 2.5), our approach allows us to instrument for multiple pollutants simultaneously and thereby check whether our estimates are driven by other pollutants.

We use the exogenous variation in PM 2.5 due to changes in wind direction to estimate the causal effect of daily fluctuations in pollution on three-day elderly mortality, hospitalizations, and health care spending, using Medicare administrative data for the years 1999 to 2011. These data span the entire United States and include over 97% of adult residents aged 65 and older. Our estimates are 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).

Estimating the effects of short-run changes in pollution on non-infant mortality is important for two key reasons. First, it is uncertain how vulnerable the adult human body is to short-run changes in pollution; it is easier to make a theoretical case for effects of long-run exposure. Laboratory studies have shown that healthy volunteers exposed to ambient pollution concentrations for a relatively short time have worse cardiovascular performance than those exposed to very clean air (Brook et al. 2009, Langrish et al. 2013). However, whether this translates into increased mortality in, for example, the general elderly population is unclear. Second, in the event mortality effects are found, it is important to translate these into life-years lost, as those who die may have had shorter life expectancies than the average member of the relevant population. Failure to do so correctly may lead one to overstate the benefits of pollution reduction.

We find that a 1 $\mu g/m^3$ increase in PM 2.5 exposure for one day causes 0.61 additional deaths per million elderly individuals on that day and the following two days. 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. These estimates are significantly larger than the corresponding OLS results, demonstrating the potential for substantial bias from measurement error and endogeneity. We also find that increases in PM 2.5 lead to more hospitalizations and higher inpatient spending, driven almost entirely by admissions via the emergency room (ER). Each unit increase in PM 2.5 increases three-day ER spending by over $15,000 per million beneficiaries. These estimates imply that, for every death caused by PM 2.5, there are more than 3 additional ER visits. OLS estimates are again much smaller and, in the cases of total inpatient spending and the total admission rate, *negative*, suggesting

that these naïve estimates suffer from significant bias. The causal impact of PM 2.5 is robust to simultaneously instrumenting for PM 2.5, CO, and O_3.[4]

Another central concern in the study of pollution's health effects is whether those who die because of pollution exposure would have died in the near future even in the absence of exposure. Whether this "mortality displacement" or "harvesting" effect exists is central for policy. If the mortality effect of pollution is concentrated among the oldest and sickest individuals, then the mortality benefit from reducing pollution, in terms of life-years lost, is likely to be much smaller than if the effect was concentrated among individuals randomly chosen from the population. The harvesting effect may be particularly important when it comes to short-run, high-frequency pollution exposure.

In this paper, we develop a novel, direct approach to the harvesting issue. Using the detailed data on health and health utilization available in the Medicare claims dataset, we predict the expected life-years remaining for each individual in our sample, which we then use to estimate the life-years lost due to pollution exposure. This generates an estimate that is less prone to bias than prior methods.[5]

We find 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 age nor gender. To shed light on whether there is any remaining bias, we employ machine learning techniques that allow us to incorporate information on a rich set of more than 1,000 individual-level variables. Making use of all available information reduces the life-years lost estimate by an additional 45 percent. Nevertheless, that estimate remains significantly above zero. While pollution has larger effects on relatively old and sick individuals, many of them would have been expected to live for years in the absence of a pollution shock.

During our study period, 1999 – 2011, the national average daily PM 2.5 level fell by 3.65-$\mu g/m^3$. As an illustration, using our results on the impact of short-run reductions in PM2.5 exposure, we estimate that this decrease has reduced the number of elderly deaths by 55,000 and life-years lost by about 150,000 annually. Assuming a standard value of $100,000 per statistical life-year implies a corresponding annual

[4] Our estimate of the mortality effect of short-term PM 2.5 exposure is similar to that found in the epidemiological literature (Dockery and Pope 1999). However, our study corrects for several biases that the epidemiological literature does not address, such as endogeneity of pollution, measurement error, and sample selection. The difference between our OLS and IV results suggests that these biases are important but may have tended to cancel each other out in previous work.

[5] Traditional analyses of the question of harvesting or short-term mortality displacement have used 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). We employ this approach as well, although we note that such an analysis cannot address any harvesting that occurs outside of the time window spanned by the leads and lags.

benefit of $15 billion. By comparison, estimating life-years lost using an average life expectancy for the population—commonly done in the health and environmental literatures—increases this estimate by 123 percent. This contrast demonstrates the value of our technique for addressing harvesting and for placing an accurate value on the social benefits of reducing mortality.

In addition to developing a novel identification strategy to estimate pollution's causal impacts across large geographic areas and temporal scales, our study is also the first to look at the impacts of short-run pollution fluctuations on non-infant mortality in a quasi-experimental setting. 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).[6] The best-identified studies often focus on infants (e.g., Currie and Walker 2011, Knittel et al. 2016), leaving a large gap in the literature regarding the impact of pollution on adults and on the elderly in particular.[7]

A handful of recent papers use wind as a source of quasi-experimental variation in pollution, but none considers short-run mortality. Anderson (2016) identifies areas that are upwind and downwind of a highway in Los Angeles and estimates average mortality differences between them.[8] In this case, the validity of the estimates require there to be no sorting across the highway. Schlenker and Walker (2016) find daily morbidity impacts of wind-transported carbon monoxide (CO) from California airports in populations living nearby. However, their study does not consider mortality, and the study's reliance on airports as a pollution source limits its geographic scope.

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

---

[6] The biological mechanisms by which particulate matter affects health are just beginning to be understood. Brook et al. (2009) conduct an experiment exposing humans to PM 2.5 and find an effect on blood pressure and vascular function.

[7] Pope and Dockery (1999) provide a survey of the epidemiology literature on both acute and chronic exposure studies. Their Table 31.1 (p. 677-679) lists more than 54 studies specifically looking at effects of acute exposure, from the 1930s to the 1990s. Pope (2000) cites 55 papers that use daily mortality data to estimate the short-term effect of particulate matter on mortality (see his Table 1, p. 714). There is also a growing literature on other effects of acute pollution exposure, for example, on worker productivity (Chang et al. 2016), which we do not review here.

[8] Herrnstadt and Muehlegger (2015) also use variation in pollution across highways but consider the effects of pollution on criminal activity rather than health outcomes.

describes our econometric strategy, including how we estimate the life-years lost, in detail. Section V presents results, 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).

PM 2.5 is not unique in its ability to traverse considerable distances; other pollutants, including carbon monoxide, sulfur dioxide, nitrous oxide, and ozone precursors, can also be carried by the wind. The extent of pollution transport depends on a host of factors, including the pollutant, 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. Clearly, the location of the pollution source also affects wind transport patterns. 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 except at a few points in time, 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. Third, even if information on emission sources were available and omitted variable bias were not a concern, 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. Such an approach 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 instrument for local PM 2.5 concentrations using 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. Holding constant the prevailing wind direction and looking at daily deviations largely eliminates this concern. Because we hold the prevailing wind direction constant, our method is most useful for examining acute, rather than chronic, exposure.

We now illustrate the type of variation used to estimate the causal effects of PM 2.5, relegating the details to Section IV. 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 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, as we discuss in Section IV.

In both figures, 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 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, 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 into 100 monitor groups, displayed in Figure 3, 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 local wind direction to vary at the monitor *group* level, but not at the individual monitor level. This restriction eases the computational burden of having many instruments and also avoids 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 reflect long-range transport, making sorting less likely to be a problem.

III: DATA

III.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 two other criteria pollutants: ozone (O_3) and carbon monoxide (CO).[9] As with PM 2.5, past literature has linked these air pollutants to mortality and other adverse health outcomes. 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 4 displays aggregate trends in PM 2.5 over time. Average concentrations of PM 2.5 have been steadily falling from about 13 micrograms per cubic meter ($\mu g/m^3$) in 1999 to about 9 $\mu g/m^3$ in 2011. One unit of PM 2.5 thus represents about 10% of the average concentration during our time period. Figure 4 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. Moreover, Grainger et al. (2016) find evidence that counties strategically place their pollution monitors in relatively clean areas. 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.

III.B. Atmospheric conditions

Wind speed and wind direction data for the years 1999-2011 are obtained from the North American Regional Reanalysis (NARR) daily reanalysis data published by the National Centers for Environmental Information (NCEI).[10] 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 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:

[9] We ignore sulfur dioxide and nitrogen dioxide because they are precursors to PM 2.5. 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.

[10] 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).

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

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 detailed weather map at the daily level using data from PRISM and weather stations.[11] These 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 average the daily measures across all grid points in a particular county.

III.D. 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 leaves us with a sample that represents over 97% of elderly living in the U.S. Respondents' dates of death, age, sex, and county of residence are obtained from the annual Medicare enrollment files for years 1999-2011 and cover 100% of beneficiaries. Health care utilization and costs are derived from the Medicare Provider Analysis and Review (MEDPAR) File, which includes an 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.[12] All county-level daily measures of hospital utilization and costs are aggregates of the MEDPAR records based on the patient's county of residence and the admission date.

Individual-level indicators for the presence of 27 chronic conditions, which we use to estimate life-years lost, are obtained from the Chronic Conditions segment of the Master Beneficiary Summary File. Chronic conditions include heart disease, COPD, diabetes, and depression, among others.[13] 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,

---

[11] 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.

[12] Specifically, our measure of cost is the total allowed charges due to the provider. This amount includes payments made by Medicare, the beneficiary, or another payer.

[13] A full listing of these 27 chronic condition variables, along with the other Medicare variables that we employ in our life-years lost analysis, is available in the appendix.

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. Our sample does not encompass the entire U.S. 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 5). However, because pollution monitors tend to be placed in more populated counties, our main regression estimates capture about 70% of the Medicare population.

Table 1 reports that the mean daily concentration of PM 2.5 in our estimation sample is 10.86 micrograms per cubic meter, with a standard deviation of 7.34. There are 49,500 Medicare beneficiaries in the average 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. 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 in order to ensure well-measured chronic conditions. The average county-day has 27,700 such individuals. 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 that the average Medicare beneficiary (McGuire et al. 2011).

Finally, the 3-day hospital spending for the entire FFS population is about $34 per beneficiary per day, in nominal terms. About 40 percent of this spending is for emergency admissions (those originating in the ER). On average, there are about 3.4 thousand hospital admissions per million FFS beneficiaries per day, and about 45 percent of these happen through the emergency room. There are also many ER visits that do not result in hospitalizations: the overall ER visit rate is over 4 thousand per million FFS beneficiaries.

IV: EMPIRICAL STRATEGY

IV.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 particulate matter on mortality, health, and health spending, net of any confounding factors. This relationship can be represented by the following regression equation:

$$Y_{cdmy} = \beta \text{PM2.5}_{cdmy} + f\left(Temp_{cdmy}, Prcp_{cdmy}, WindSpeed_{cdmy}\right) \quad (1)$$
$$+ \sum_{t=d+1}^{d+2} \left[\gamma_t \text{PM2.5}_{ctmy} + f_t\left(Temp_{ctmy}, Prcp_{ctmy}, WindSpeed_{ctmy}\right)\right]$$
$$\sum_{t=d-1}^{d-2} \gamma_t \text{PM2.5}_{ctmy} + \alpha_c + \alpha_{sm} + \alpha_{my} + \epsilon_{cdmy},$$

where the dependent variable is one of several outcome variables in county c on day d in month m and year y. We first examine the effect of PM 2.5 on the death rate, measured in deaths per one million Medicare beneficiaries. The other outcome variables measure health care utilization: total hospital spending per million beneficiaries, hospital admissions per million beneficiaries. We calculate these measures for all hospital admissions and for emergency room (ER) admissions only. Finally, we also observe outpatient ER visits (those that do not result in hospitalizations), but not the spending associated with such visits. The parameter of interest is β, the coefficient on daily PM 2.5 levels.

The dependent variable Y_{cdmy} is calculated using a three-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 accounts for very short-term harvesting as well as delayed effects – someone who gets sick from the high pollution on January 1 may not die until a day or so later. To ensure that β is not capturing the effects of *future* pollution variation, which may be correlated with contemporaneous pollution variation, we control for PM 2.5 concentrations and weather conditions over the two days following day d. To ensure that β is not capturing any effects from *past* pollution variation, we also include two lags of PM 2.5 concentrations.

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. 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 regressions.[14] Our estimates are robust to less flexible weather controls or omitting weather controls entirely (see Table 8), reinforcing the assumption that our source of identifying variation is exogenous.

The regressions 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 and allows this correlation to vary by state. Finally, month-by-year fixed effects control flexibly for common time-varying shocks, as could be induced by Medicare or environmental policy changes. As with weather controls, our results are very robust to varying the fixed effects (see Table 8).

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:

$$\begin{aligned} \text{PM2.5}_{cdmy} = & \sum_{g=1}^{100}\sum_{b=0}^{2} \beta_b^g WINDDIR_{cdmy}^{90b} + f\left(Temp_{cdmy}, Prcp_{cdmy}, WindSpeed_{cdmy}\right) \\ & + \sum_{t=d+1}^{d+2} \left[g_t\left(WINDDIR_{ctmy}\right) + f_t\left(Temp_{ctmy}, Prcp_{ctmy}, WindSpeed_{ctmy}\right)\right] \\ & + \sum_{t=d-1}^{d-2} g_t(WINDDIR_{ctmy}) + \alpha_c + \alpha_{sm} + \alpha_{my} + \epsilon_{cdmy} \end{aligned} \tag{2}$$

The variable $WINDDIR_{cdmy}^{90b}$ is an indicator variable equal to 1 if the daily average wind direction in county c falls in the 90-degree interval $[90b, 90b + 90)$ and 0 otherwise. The omitted category corresponds to the interval [270,360). The coefficient on this variable, β_b^g, is allowed to vary across 100 different geographic regions, as explained below. Because our outcomes of interest are measured over three

[14] 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).

days, we include two leads of the instrumental variables as controls (represented by $g_t(WINDDIR_{\text{ctmy}}) = \sum_{g=1}^{100}\sum_{b=0}^{2}\gamma_{b,t}^{g}WINDDIR_{ctmy}^{90b}$).[15] To capture any autocorrelation in wind direction, we also control for two lags of the instruments. (Our results are robust to the inclusion of more (or fewer) lags, as shown in Table 9.) The other control variables are defined as in equation (1).

Equation (2) estimates a common effect of wind direction on pollution for all monitors within each of the 100 geographic areas, all 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).[16] The results are displayed in Figure 3. 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 with PM 2.5 readings and 9 counties.

Clustering monitors together in this way improves the plausibility of our key identifying assumption because it causes the specification outlined in equation (2) to capture pollution produced primarily 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.[17] 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. A second benefit of clustering monitors into groups is that is reduces measurement error. Pollution from local sources will spread heterogeneously throughout an area. Thus, when a monitor registers a certain amount of pollution, some people in the region will experience

[15] In principle, we could follow the specification of equation (1) and instrument for leads of PM 2.5 instead. However, instrumenting for PM 2.5 leads significantly raises the computational burden of the estimation, and those coefficients are not the primary outcome of interest.

[16] Specifically, we partition the monitors into 100 groups using the *k*-means cluster algorithm.

[17] 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.

higher pollution levels than others, and this measurement error in experienced pollution will tend to bias estimates of the causal impact of pollution. In contrast, since our clustering approach focuses on pollution that is transported into the region, it is much more likely to uniformly affect the entire region, purging our estimates of bias due to measurement error.

Equation (2) also restricts the effect of wind direction on pollution levels to be constant within each of the four $WINDDIR$ bins. This makes it even more likely that the pollution variation we use 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 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 hundreds of instruments, including controls for their leads, thousands of control variables and fixed effects, 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 wind direction. 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.

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.[18] 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.

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.

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

IV.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 in deaths per million Medicare beneficiaries. We can then use this estimate to calculate the number of deaths averted because of a decline in the levels of PM 2.5, which can be combined with an estimate of the value of a statistical life (VSL) measure to monetize the benefit. However, such an exercise will overstate the benefits of pollution reduction if the individuals who die as a result of pollution are sicker than the general population and would not have lived for much longer anyway. While this is a concern with any population, including infants, it may be particularly relevant for the elderly.

An alternative economic framework calculates the social value of a mortality reduction based on the number of *life-years* lost rather than *lives* lost. According to this metric, extending an individual's life by four years is twice as valuable as extending her life by two years. A common estimate of the value of a statistical life-year is $100,000 (Cutler 2004). However, since the mortality cost scales linearly in the value of a statistical life year, the key conceptual challenge is accurately estimating the number of life-years lost.

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 counterfactual life expectancy per life lost. This counterfactual life expectancy is typically derived from population life tables (Deschenes and Greenstone 2011) or prior research (Finkelstein and McKnight 2008). A general concern with this approach is that it will overstate estimates of life-years lost when individuals affected by pollution 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 have lower life expectancies than similar individuals who do not have such conditions, even absent a pollution event.

We propose a new methodology that exploits the detailed data derived from Medicare claims 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 upwardly 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.[19]

[19] We present an abbreviated discussion of the method in the main text. Additional details are contained in the Appendix.

Let L_{it} be the number of statistical life-years lost due to death by individual *i* in period *t*.[20] For individuals who do not die, $L_{it} = 0$, while for individuals who do die at time *t*, L_{it} is equal to the number of years that that individual *i* would have lived conditional on being alive at the start of period *t*. For simplicity, we first assume that exposure to PM 2.5 is assigned randomly and affects all individuals equally. If L_{it} was observable, then the researcher could estimate the effect of PM 2.5 on the number of life-years lost in period t by estimating 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 life-years lost and, by assumption, is uncorrelated with PM2.5_{it}. Under these conditions, equation (3) consistently estimates γ, the causal effect of PM2.5_{it} on life-years lost.[21]

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

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

Bias arises when estimating equation (4) in the presence of heterogeneous treatment effects. 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 can then be written as

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

The error term in (5) contains a third component, $v_i \text{PM2.5}_{it}$, which represents the portion of the individual's treatment effect not accounted for by the average treatment effect, γ. If the heterogeneous treatment effect, v_i, is correlated with the measurement error in counterfactual life expectancy, so that $Cov(v_i \text{PM2.5}_{it}, u_{it}) \neq 0$, then bias will arise. For example, suppose that the researcher does not account

---

[20] 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 are lower bounds.

[21] 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^*_{ct}, where $L^*_{ct} = \frac{\sum_{i=1}^{N_{ct}} L_{ict}}{N_{ct}}$ is the average number of life-years lost in county c at time t, and N_{ct} is the number of individuals living in county c. It is straightforward to extend the framework to accommodate this case.

for sex 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) 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 may be more susceptible to dying from pollution exposure than population with high levels of remaining life expectancy.

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 (4). 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, v_i. To our knowledge, no previous study has addressed this bias by using variables other than age to predict life expectancy (e.g., Deschenes and Greenstone 2011). By contrast, we incorporate a rich set of data that includes information on chronic conditions, medical spending, healthcare utilization, and geographic location. We shall show that this matters: using an unconditional life expectancy or estimating life expectancy using only basic demographic variables such as age and gender causes significant upward bias in regression estimates of life-years lost due to air pollution.

Integrating this framework into the county-level empirical strategy presented earlier is straightforward: we simply aggregate estimates of life-years lost over all individuals in the county. 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.

A challenge with estimating counterfactual life expectancy is that not everybody dies during the time period we observe them. We therefore employ a Cox proportional hazards model, a commonly used

survival model.[22] This model assumes that the hazard rate of death for individual i can be factored into two functions:

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

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 the indicator variable δ_i is equal to one for individuals whose deaths we observe (uncensored observations) and equal to zero otherwise. We then nonparametrically estimate the baseline hazard function, $h_0(t_i)$, following Breslow (1972). See the Online Appendix for details.

We estimate this Cox proportional hazards model using data from the 2002 cohort of Medicare beneficiaries.[23] 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 our sample dies; the remaining deaths are censored. To ensure that we have 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.[24] For computational ease, we further limit the analysis to a random 5 percent sample of these beneficiaries. The final estimation sample consists of 1,211,585 individuals.

To assess the importance of accounting for harvesting, we estimate the survival model several times, using increasingly large sets of characteristics. First, we use no individual characteristics; that is, we assume a homogeneous survival function. Then we control for age and sex, and then we add indicators for the presence of different chronic conditions. Our final and preferred specification incorporates additional individual-level data from variables in the Medicare dataset and zipcode-level data from the American

[22] Fully parametric models that assume survival rates are governed by either the Gompertz or Weibull distributions yield very similar results.

[23] Although earlier cohorts are observable for a longer period of time, 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.

[24] Because we do not include beneficiaries who qualify for Medicare prior to age 65, this restriction means that nobody in this sample is under age 67.

Community Survey. These include 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.[25] There are two challenges with including so many control variables. First, it raises the concern that 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. 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).[26] 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.[27] We then use estimates of β and observable characteristics x_i of 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.

Figure 6 demonstrates the increase in explanatory power that accompanies the inclusion of these additional demographic and health variables. This figure plots the average predicted life expectancy, by calendar year, for Medicare beneficiaries who died in that year. Figure 6 reports that the unconditional average life expectancy (i.e., the estimated life expectancy for the average Medicare beneficiary) is slightly under 12 years. The blue line adjusts the predicted life expectancy based on the age and gender of the Medicare beneficiaries who die in that calendar year. Because these decedents are older than the average Medicare beneficiary, this reduces predicted life expectancy by about three years. The red line additionally controls for 27 different chronic conditions, reducing predicted life expectancy by yet another year. This

[25] There are 1,062 variables in total, detailed in the Online Appendix.

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

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

happens because decedents are sicker than the average Medicare beneficiary, even after controlling for age and gender.

Finally, the black line in Figure 6 displays average predicted life expectancy based on a model that incorporates data from all 1,062 variables in our dataset. This reduces predicted life expectancy by yet another half year, on average. Note that this result by itself does not necessarily mean that it is important to control for these additional variables. Although these decedents on average have low predicted life expectancies, it could be the case that air pollution selectively kills beneficiaries with (relatively) high predicted life expectancies. The only way to know whether it matters is to see how results differ when using these different predictions. Those results are presented in the next section.

V. RESULTS

V.A. Mortality counts and health care utilization

Panel A of Table 2 reports OLS estimates from equation (1) describing the relationship between daily PM 2.5 and 3-day mortality rates per million beneficiaries in the relevant age group. As reported in Column (1), each 1-$\mu g/m^3$ increase in daily PM 2.5 exposure corresponds to 0.098 additional deaths per million elderly over the following three days, or a 0.025% increase relative to the average 3-day mortality rate. This implies that a *one standard deviation* (7.34-$\mu g/m^3$) increase in daily PM 2.5 exposure corresponds to 0.72 additional deaths per million, or a 0.18% increase in 3-day mortality. 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 the elderly 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. The magnitudes of the IV estimates are substantially larger than the corresponding OLS estimates, suggesting that OLS estimation may suffer from measurement error bias. The IV estimates imply that each 1-$\mu g/m^3$ increase in daily PM 2.5 exposure corresponds to 0.605 additional deaths per million elderly over the following three days, or a 0.15% increase relative to the average 3-day mortality rate.[28] The corresponding estimate for a one standard deviation increase in daily PM 2.5 is a 1.1% increase in 3-day mortality. Columns (2) – (6) show a monotonic relationship between the mortality effect of PM 2.5 and

[28] We focus on 3-day mortality in order to allow for pollution to have lagged effects. The comparable estimate resulting from IV estimation of (1) for 1-day mortality yields a coefficient of 0.382 additional deaths resulting from a 1-$\mu g/m^3$ increase in daily PM 2.5 exposure, suggesting that the mortality impact of PM 2.5 exposure grows over time.

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 are relatively stable and follow a U-shaped pattern. Each 1-μg/m^3 increase in daily PM 2.5 exposure increases 3-day mortality by 0.19% among ages 65-69, by 0.09% among ages 75-79, and by 0.18% among ages 85 and over. We return to this point when discussing our estimates of life-years-lost due to PM 2.5.

Numerous papers have estimated the effects of pollution on mortality using epidemiological methods, and these estimates are often used for calculating the benefits of various environmental policies.[29] It is worth comparing our point estimates of the mortality effects of PM 2.5 to these previous estimates. Many of these studies that are well-cited and used in cost-benefit analysis focus on the effects of long-term (chronic) exposure, rather than short-term (acute) exposure, which is the focus of our study. One such study is Pope et al. (2002), which uses a Cox proportional hazards survival model to estimate the relative risk associated with PM 2.5 exposure. They find that a 10-μg/m^3 increase in long-run PM 2.5 exposure is associated with a 4% increase in the annual risk of all-cause mortality. Our coefficient implies that the same magnitude *daily* increase in PM 2.5 increases *daily* mortality by 1.5% of the average daily mortality rate. Roman et al. (2008) present a study of expert elicitation of the mortality effects of long-term PM 2.5 exposure and report that the effect of a 1-μg/m^3 decrease in PM 2.5 ranges from a 0.6% to 1.25% reduction in mortality. Besides the fact that these papers estimate the effect of long-term, not short-term, PM 2.5 exposure, there are other important differences between these two studies and ours.[30] Pope et al. (2002) was until recently used in EPA regulatory analysis, but now the EPA employs a larger estimate of the mortality effect: a 10-μg/m^3 PM 2.5 increase raises all pollution-related mortality by 10.6% (Chapter 5, US EPA 2011; see also Parry et al. 2014, p. 75).[31] Overall, our estimated (short-term) PM 2.5 pollution effects are 14-38% as large as these (long-term) PM 2.5 pollution effects.

A more direct comparison is between our results and other studies that specifically look at effects of *short-term* particulate exposure. In a literature review incorporating more than 50 studies examining the

---

[29] For example, Holland et al. (2015) use estimates of mortality effects of conventional pollutants to calculate the benefits of electric vehicles, and Coady et al. (2016) use them to calculate the costs of global energy subsidies.

[30] Both Pope et al. (2002) and Roman et al. (2008) look at overall adult mortality, not just elderly mortality. Pope et al. (2002) do not use an instrumental variables strategy, and instead they estimate relative risk ratios using a Cox proportional hazards model. Roman et al. (2008) present qualitative judgments of the pollution effects from a panel of experts. The magnitude of the mortality effect from PM 2.5 from both of these papers is used in Holland et al. (2015).

[31] This estimate is used in the impact analysis in Coady et al. (2016).

effects of fine particulates (PM 2.5), Pope (2000) reports that a 5-μg/m^3 increase in acute PM 2.5 exposure is associated with a 1% increase in all-cause mortality (see his Figure 2, p. 717). Our estimates imply a somewhat smaller magnitude: 0.77%. Pope and Dockery (1999), in a review of the literature on short-term effects of particulates, report that a 10-μg/m^3 PM 10 increase raises daily mortality rates by 0.5% to 1.5% (see their Figure 31.3, p. 681). Using a standard conversion of PM 10 to PM 2.5,[32] our estimated effect from PM 2.5 is within this range.[33] Although the estimates from the literature reviews reported in Pope (2000) and Pope and Dockery (1999) are for effects of short-term (daily) particulate exposure, they are not specifically about elderly mortality. A small number of papers specifically estimate the effect of short-term particulate exposure on elderly mortality.[34] Franklin et al. (2007) report that a 10-μg/m^3 PM 2.5 increase is associated with a 1.66% increase in mortality among the 75+ population. Salvina et al. (1995) examine the effect of PM 10 on daily elderly mortality in Sao Paolo, Brazil. They report that a 100-μg/m^3 PM 10 increase is associated with a 13% increase in elderly mortality. These are nearly identical to the magnitude of our estimate.[35]

Next, we explore the relationship between daily PM 2.5 and 3-day hospitalization rates and associated medical spending per million beneficiaries enrolled in fee-for-service Medicare. Panel A of Table 3 shows that the OLS relationship estimated by equation (1) is mixed: each 1-μg/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. However, when we instrument for PM 2.5, we estimate that increases in daily PM 2.5 consistently increase both hospitalizations and inpatient spending. The IV estimates reported in Panel B imply that each 1-μg/m^3 increase in daily PM 2.5 causes a marginally significant increase in total inpatient spending and a similarly sized but highly significant increase in ER inpatient spending by over $15 thousand per million beneficiaries (relative to a mean of $13.7 million). The overall admissions rate increases by 2.03 per million

---

[32] We use the EPA's guidelines suggesting that one unit of PM 2.5 is equivalent to 1.67 units of PM 10 (Dockery and Pope 1994).

[33] Some more recent studies are generally consistent with the studies reported in these literature reviews. For example, Zanobetti and Schwartz (2009) find a 0.98% increase and Dai et al. (2014) find a 1.18% increase in total mortality from a 10-μg/m^3 PM 2.5 increase.

[34] Liao et al. (1999) examine the effect of daily PM 2.5 exposure on elderly cardiac control, not elderly mortality. They find a link between the two, suggestive of the role that acute PM 2.5 exposure may play in elderly cardio-vascular-related mortality.

[35] 100 $\frac{\mu g}{m^3}$ PM 10 is roughly equivalent to 60 $\frac{\mu g}{m^3}$ PM 2.5, so their estimated effect implies that a 1 $\frac{\mu g}{m^3}$ PM 2.5 increase is associated with roughly a 0.22% elderly mortality increase, compared to our estimate of 0.26%.

beneficiaries, an increase which appears to be almost entirely explained by the 1.96 additional admissions originating through the ER. Finally, we 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. Whereas existing studies on pollution and hospitalizations have primarily examined more narrow categories of hospitalization, such as ER admissions or cause-specific admissions, a limitation of this approach is that the results may partly reflect a substitution effect of pollution on the types of admissions individuals experience. In contrast, our results demonstrate that PM 2.5 increases *on net* both the number of individuals visiting the ER and the number admitted to the hospital.

The results are consistent with a large epidemiological literature that generally concludes, using observational studies, that acute particulate exposure leads to an increase in hospital admissions.[36] For example, Schwartz (1994a) finds that a 100 $\frac{\mu g}{m^3}$ increase in daily PM10 concentrations in Birmingham, Alabama, increases the relative risk of pneumonia admissions by 19% and of chronic obstructive pulmonary disease admissions by 27%.[37] Dai et al. (2014) find that a 10-μg/m^3 PM 2.5 increase is associated with a 1.03%-2.35% increase in hospital admissions. Zanobetti et al. (2009) find that the same PM 2.5 increase is associated with a 1.85% to 2.74% increase in ER admissions. Our results imply that the same increase causes an increase of about 1.2% in ER admissions.

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

Table 4 displays estimates of our baseline specification, equation (1), when the outcome variable is L_{cdmy}, the estimated 3-day life-years lost per million beneficiaries in county c. The estimation sample in this section is limited to those beneficiaries continuously enrolled for at least two years in fee-for-service (FFS) Medicare because some of the specifications require data on prior year's medical spending and utilization and on chronic conditions coded using a look-back window of FFS claims to estimate counterfactual life expectancy. For reference, Column (1) shows the estimated effect of PM 2.5 on 3-day mortality per million among the 2-year FFS population. This estimate is slightly larger in absolute terms than the IV estimate from Table 2 for the 65+ population in part because the 2-year FFS restriction mechanically excludes individuals ages 65-66 and is therefore older. The effects relative to average 3-day mortality are very similar for both populations.

---

[36] For a survey, see Pope and Dockery (1999, p. 681-684).

[37] A smaller effect is found in Detroit, Michigan (Schwartz 1994b), and in Minneapolis-St. Paul, Minnesota, there is a similarly-sized effect for pneumonia admissions but an even larger effect for COPD admissions (Schwartz 1994c). All three Schwartz studies use daily variation in pollution levels and hospital admissions.

Column (2) displays results when every decedent's counterfactual life expectancy is set equal to the mean for the 2-year FFS population (11.7 years). This estimate implies that each 1-μg/m^3 increase in daily PM 2.5 increases life-years lost by 8.7 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 11.7. However, this estimate is accurate only if beneficiaries killed by PM 2.5 represent a random sample of the 2-year FFS population. By contrast, if those decedents have a lower counterfactual life expectancy than those who remain alive, then the estimate in Column (2) will be biased upward.

Columns (3), (4), and (5) of Table 4 illustrate this bias by progressively increasing the number of covariates used to predict counterfactual life expectancy. Those covariates are reported in the column headers. Column (3) displays estimates when a decedent's counterfactual life expectancy is modeled solely as a function of age and sex. This approach is comparable to studies that calculate age- and sex-specific mortality effects and multiply by the corresponding life expectancies from population life tables (e.g., Deschenes and Greenstone 2011). In our setting, accounting for age and sex in the life expectancy calculations reduces the coefficient on PM 2.5 by about 30 percent, to 6.0 years per million beneficiaries. This is consistent with the result from Table 2 that older beneficiaries, who have lower life expectancies, are 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 impact 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, which allows us to optimally incorporate data among the more than 1,000 additional predictors as described earlier. This final estimate, reported in Column (5), is about 20 percent smaller than estimates that account only for age, sex, and chronic conditions and implies that each 1-μg/m^3 increase in daily PM 2.5 increases life-years lost by 2.7 years per million beneficiaries. The fact that the estimate from Column (5), with over 1,000 control variables, is only modestly smaller than the prior estimate, based only on age, sex, and chronic conditions, suggests that this final estimate is close to the true value.

Table 4 illustrates that adding additional predictors when estimating life expectancy can substantially reduce the estimate of life years lost due to pollution. This reduction can occur for two reasons. First, better survival models should predict lower remaining life expectancy for decedents *on average*. Table 4 reports that the mean life years lost per decedent ("LYL per decedent") decreases from 11.5 in the model with no predictors to 4.9 in the LASSO model. Second, better survival models generate more accurate predictions of counterfactual life expectancy even among individuals observed to die. Thus, a more accurate model may not only reduce the *average* predicted life expectancy among decedents, but may also

increase the variance in the individual-level distribution of those predictions. 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 (0.83 per million). This indicates that the mortality impacts 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 increased PM 2.5 driven by the wind. This estimate can be compared to the average life years lost among all decedents to shed light on whether those dying from increased pollution appear to be differentially healthy or frail compared to those who die on a typical day. The LYL per complier is calculated by dividing the estimated impact of increased PM 2.5 on life years lost by the mortality impact (the coefficient in Column (1)).[38] When life expectancy is modeled by age and sex alone, those dying from pollution appear to have slightly longer life expectancies (8.1 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 (3.7 years) compared to the average decedent (4.9 years).

Next, we estimate the value of a 3.65-$\mu g/m^3$ decrease in PM 2.5, which roughly corresponds to the average decrease experienced nationwide between 1999 and 2011 as shown in Figure 3. The estimate in column (5) of Table 4 implies that such a decrease saved 148,791 life-years annually among the 41 million Medicare beneficiaries alive in 2011.[39] If we assign each life-year a standard value of $100,000 each, this implies annual health benefits of about $15 billion. Perhaps not surprisingly, this number is over 75 percent lower than the estimate of $64 billion obtained from ignoring heterogeneity in the effect of pollution on mortality. More importantly, the estimated benefits based on age and sex alone are about $33 billion, more than double our estimated benefit of $15 billion based on our most comprehensive model. This

[38] For example, in Column (3), a one-unit pollution increase causes 0.746 deaths per million and causes an increase in LYL of 6.013 years per million. Thus, the LYL per person killed by pollution is 6.013/0.746 = 8.057.

[39] The exact calculation is $2.724 \times 365 \times 41 \times 3.65$

demonstrates the importance of properly accounting for "harvesting" when calculating the mortality benefits of reductions in air pollution.[40]

Importantly, this calculation assumes that the estimates of the effects of *short-run* pollution exposure that we estimate can be mapped into the *long-run* decrease in pollution exposure from 1999-2011. As documented earlier, the epidemiological literature generally finds larger effects from long-run exposure than from short-run exposure, on a per-unit basis. Finally, our calculation of $15 billion ignores any other benefits associated with a reduction in air pollution, e.g., a reduction in non-lethal asthma attacks, or any health or other benefits to the non-Medicare population.

V.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). Two other pollutants, sulfur dioxide (SO_2) and nitrogen oxide (NO_2) are precursors to PM 2.5, but as also thought to have independent health effects. 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, many counties only track a subset of these pollutants, and pollutants are often monitored on different days. Including all five pollutants reduces our sample size from over 1.6 million observations to slightly over 300 thousand observations. Thus, although our empirical approach can be applied to a setting with multiple pollutants, the inconsistent measurements of those pollutants causes us to lose a lot of power, and possibly generalizability, when we estimate an equation that includes all five as endogenous variables.

Moreover, SO_2 quickly 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 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.[41]

---

[40] A calculation of the costs of this reduction in PM2.5 is unavailable. But for comparison, the EPA's calculation of the annual costs of meeting the CAAA air quality standards (which includes standards for all criteria pollutants, not just PM2.5) is $43.9 billion in 2010. Their calculation of total benefits is $1.3 trillion, of which $1.2 trillion comes from PM-caused adult mortality (EPA 2011). This is based on a mean value for VSL of $7.4 million and 160,000 annual avoided deaths. The EPA does not use age-adjusted VSLs or the value of a statistical life-year (VSLY).

[41] For example, a conservative conversion rate of 3% 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% hourly conversion rate implies that over 60 (85) percent is converted after 24 (48) hours.

Nonetheless, it is useful to investigate whether our estimates of the impact of PM 2.5 change after controlling for CO and O_3. We do this by restricting the sample to county-days where readings for these pollutants and PM 2.5 are available, and then sequentially adding additional pollutants to our main estimating equation. The results are shown in Table 5. 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 300 instruments. Although our reported F-statistics are generally quite large, we nevertheless 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 6, 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, $WINDDIR\_PLACEBO_{cdmy}^{60b}$, and use those in our first stage instead of the actual wind direction, $WINDDIR_{cdmy}^{60b}$. Those results, shown in Table 7, 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.

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 burden of high and increasing Medicare spending, it is especially important to understand this effect among the elderly. Causal effects of pollution are difficult to identify because of endogeneity and measurement error. Valuation of mortality reductions is difficult because of the heterogeneity of pollution effects – if pollution-induced deaths occur among the least healthy of the population, then ignoring this heterogeneity will lead to upward-biased calculations of the value of pollution reductions.

This paper sheds light on these issues by estimating the effect of short-term pollution exposure on mortality counts, life-years lost, and other health outcomes using a novel identification strategy based on

exogenous variation in wind direction.[42] This is accomplished through a novel data effort linking daily pollution and climatic variables to detailed administrative Medicare records on all Medicare beneficiaries from 1999-2011. We find significant effects of pollution on mortality, health care spending, and hospitalizations. Finally, we use the rich Medicare data to estimate counterfactual survival functions and estimate that the reduction in PM 2.5 experienced nationwide since 1999 generates $15 billion annually in mortality benefits among the elderly alone.

[42] While our empirical strategy is based on daily wind fluctuations and thus designed to estimate the effect of short-term pollution exposure, other studies estimate the effects of long-term exposure (Chay et al. 2003, Chen et al. 2013, Anderson 2015).

REFERENCES

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

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

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

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.

Breslow, Norman E. "Discussion of Professor Cox's paper." J Royal Stat Soc B 34 (1972): 216-217.

Brook, Robert D., Bruce Urch, J. Timothy Dvonch, Robert L. Bard, Mary Speck, Gerald Keeler, Masako Morishita et al. "Insights into the mechanisms and mediators of the effects of air pollution exposure on blood pressure and vascular function in healthy humans." Hypertension 54, no. 3 (2009): 659-667.

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

Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell. "The Effect of Pollution on Worker Productivity: Evidence from Call-Center Workers in China." National Bureau of Economic Research Working Paper no. 22328, June 2016.

Chay, Kenneth, Carlos Dobkin, and Michael Greenstone. "The Clean Air Act of 1970 and adult mortality." Journal of Risk and Uncertainty 27, no. 3 (2003): 279-300.

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

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. "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, no. 32 (2013): 12936-12941.

Coady, David P. and Parry, Ian and Sears, Louis and Shang, Baoping, How Large are Global Energy Subsidies? (March 17, 2016). CESifo Working Paper Series No. 5814. Available at SSRN: http://ssrn.com/abstract=2766224

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.

Dai, Lingzhen, Antonella Zanobetti, Petros Koutrakis, and Joel D. Schwartz. "Associations of fine particulate matter species with mortality in the United States: a multicity time-series analysis." Environmental Health Perspectives (Online) 122, no. 8 (2014): 837.

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, Douglas W., and C. Arden Pope. "Acute respiratory effects of particulate air pollution." Annual review of public health 15, no. 1 (1994): 107-132.

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

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

Franklin, Meredith, Ariana Zeka, and Joel Schwartz. "Association between PM2. 5 and all-cause and specific-cause mortality in 27 US communities." Journal of Exposure Science and Environmental Epidemiology 17, no. 3 (2007): 279-287.

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

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

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.

Harrell, Frank E., Kerry L. Lee, and Daniel B. Mark. "Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors." Statistics in medicine 15 (1996): 361-387.

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

Holland, Stephen P., Erin T. Mansur, Nicholas Z. Muller, and Andrew J. Yates. Environmental benefits from driving electric vehicles?. No. w21291. National Bureau of Economic Research, 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.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders. "Caution, drivers! Children present: Traffic, pollution, and infant health." Review of Economics and Statistics 98(2) (May 2016), 350-366.

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.

Liao, Duanping, John Creason, Carl Shy, Ron Williams, Randall Watts, and Roy Zweidinger. "Daily variation of particulate air pollution and poor cardiac autonomic control in the elderly." Environmental Health Perspectives 107, no. 7 (1999): 521.

Lin, Y-C. and M-T. Cheng. “Evaluation of formation rates of NO_2 to gaseous and particulate nitrate in the urban atmosphere.” *Atmospheric Environment* 41, no. 9 (2007).

Luria, Menachem, Robert E. Imhoff, Ralph J. Valente, William J. Parkhurst, and Roger L. Tanner. “Rates of conversion of sulfur dioxide to sulfate in a scrubbed power plant plume.” *Journal of the Air & Waste Management Association* 51, no. 10 (2001).

McGuire, Thomas G., Joseph P. Newhouse, and Anna D. Sinaiko. "An economic history of Medicare Part C." *Milbank Quarterly* 89, no. 2 (2011): 289-332.

Parry, Ian WH, Mr Dirk Heine, Eliza Lis, and Shanjun Li. Getting energy prices right: from principle to practice. International Monetary Fund, 2014.

Pope, C. A. "Epidemiology of fine particulate air pollution and human health: biologic mechanisms and who's at risk?." Environmental health perspectives 108, no. Suppl 4 (2000): 713-723.

Pope, C. A., and D. W. Dockery. "Epidemiology of particle effects." Air pollution and health 31 (1999): 673-705.

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.

Pope III, C. Arden, Richard T. Burnett, Michael J. Thun, Eugenia E. Calle, Daniel Krewski, Kazuhiko Ito, and George D. Thurston. "Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution." Jama 287, no. 9 (2002): 1132-1141.

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.

Saldiva, Paulo HN, C. Arden Pope III, Joel Schwartz, Douglas W. Dockery, Ana Julia Lichtenfels, Joao Marcos Salge, Ivana Barone, and Gyorgy Miklos Bohm. "Air pollution and mortality in elderly people: a time-series study in Sao Paulo, Brazil." Archives of Environmental Health: An International Journal 50, no. 2 (1995): 159-163.

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.

Schwartz, Joel. "Air pollution and hospital admissions for the elderly in Birmingham, Alabama." American journal of epidemiology 139, no. 6 (1994a): 589-598.

Schwartz, Joel. "Air pollution and hospital admissions for the elderly in Detroit, Michigan." American journal of respiratory and critical care medicine 150, no. 3 (1994b): 648-655.

Schwartz, Joel. "PM10 ozone, and hospital admissions for the elderly in Minneapolis-St. Paul, Minnesota." Archives of Environmental Health: An International Journal 49, no. 5 (1994c): 366-374.

Tibshirani, Robert. "The lasso method for variable selection in the Cox model." *Statistics in medicine* 16, no. 4 (1997): 385-395.

United States Environmental Protection Agency (US EPA), 2011, The Benefits and Costs of the Clean Air Act from 1990 to 2020 , Report to Congress (Washington: US Environmental Protection Agency).

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

Simon, Noah, Jerome Friedman, Trevor Hastie, and Rob Tibshirani. "Regularization paths for Cox's proportional hazards model via coordinate descent." Journal of statistical software 39, no. 5 (2011): 1-13.

Zanobetti, Antonella, Meredith Franklin, Petros Koutrakis, and Joel Schwartz. "Fine particulate air pollution and its components in association with cause-specific emergency admissions." Environmental Health 8, no. 1 (2009): 1.

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.

Zanobetti, Antonella, and Joel Schwartz. "The effect of fine and coarse particulate air pollution on mortality: a national analysis." Environmental health perspectives 117, no. 6 (2009): 898.

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

FIGURES

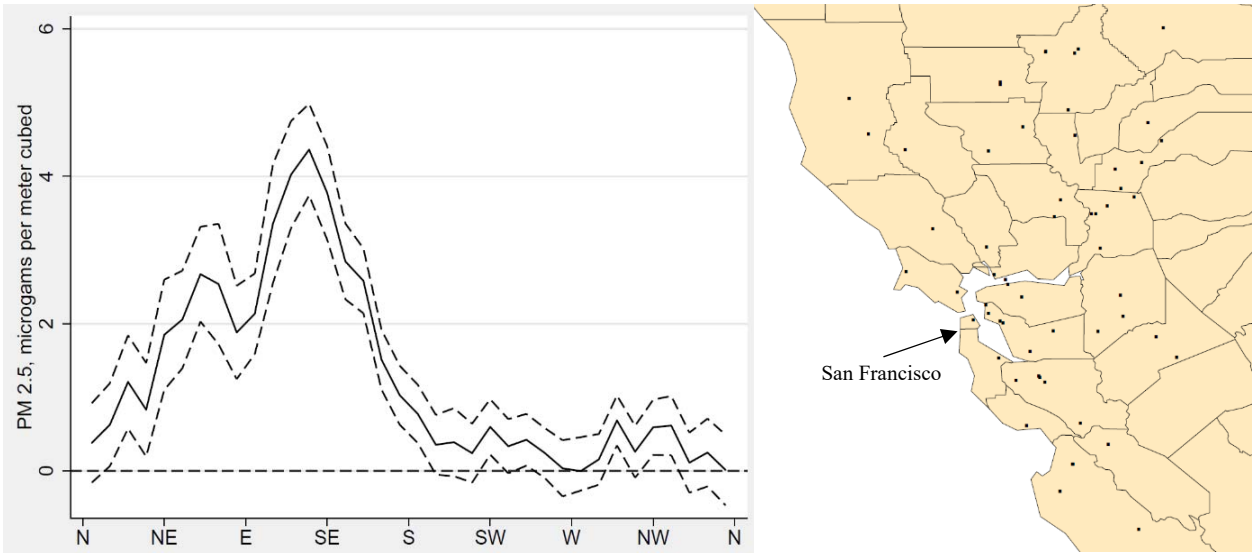


Figure 1. Relationship between daily average wind direction and PM 2.5 concentrations in and around the Bay Area, CA. The left panel shows regression estimates where the dependent variable is the average daily PM 2.5 concentration and the key independent variables are a set of indicators for the daily 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.

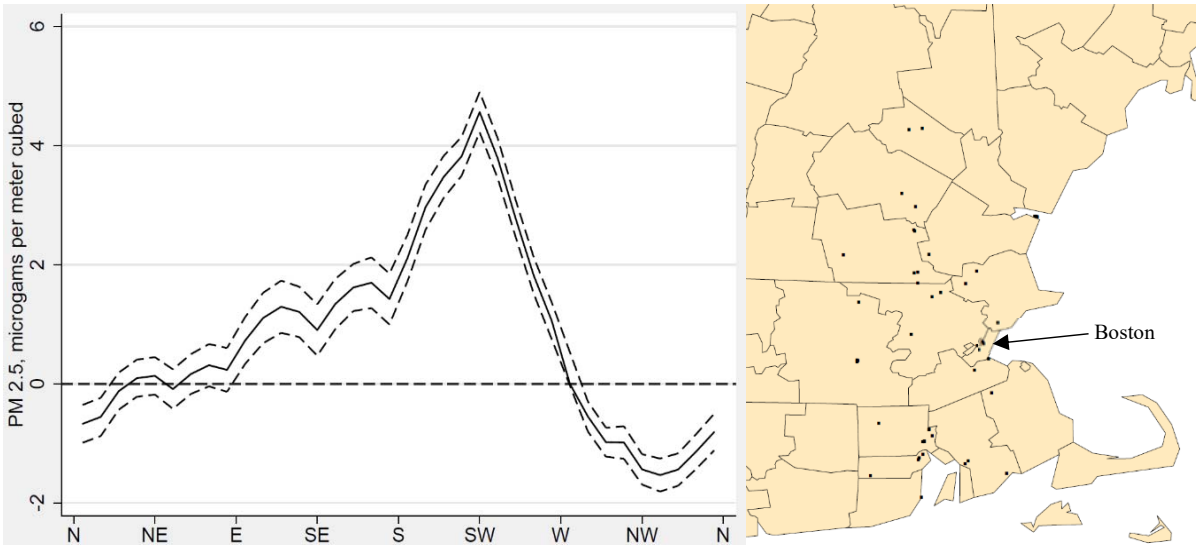


Figure 2. Relationship between daily average wind direction and PM 2.5 concentrations in and around the Boston Area, MA. The left panel shows regression estimates where the dependent variable is the average daily PM 2.5 concentration and the key independent variables are a set of indicators for the daily 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.

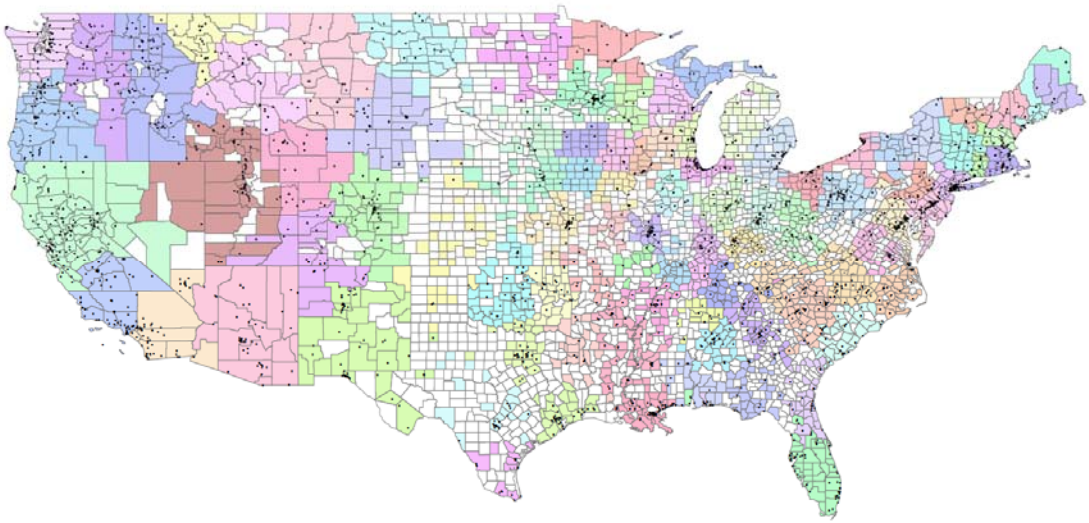


Figure 3. Counties assigned to each monitor group. Different colors correspond to different monitor groups. White corresponds to counties not assigned to any monitor group due to lack of monitors. Black dots represent PM 2.5 pollution monitors.

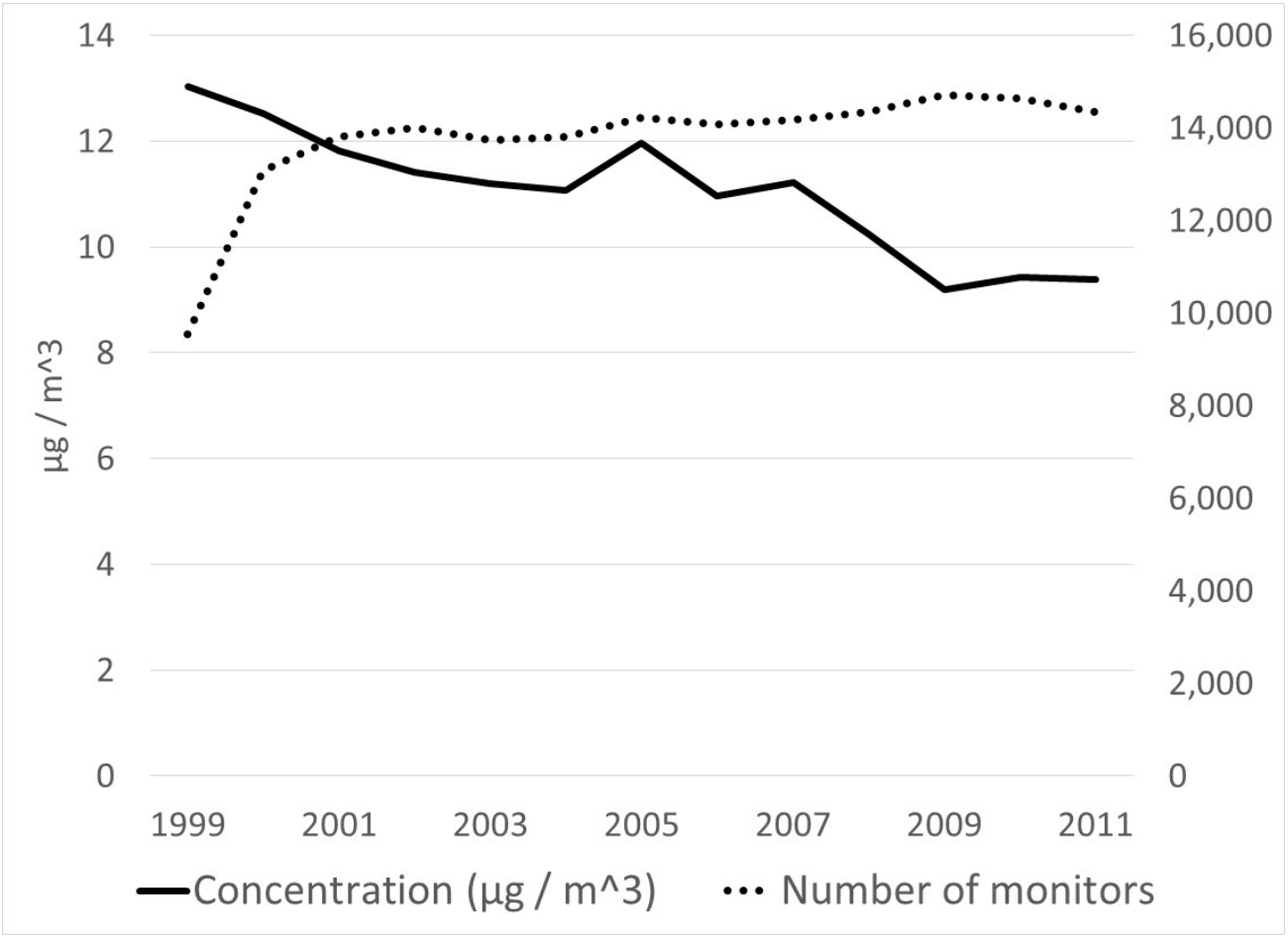


Figure 4. Trends in PM 2.5 air pollution, 1999-2011. Figure displays annual county means for PM 2.5 concentration, and the nationwide total number of monitors.

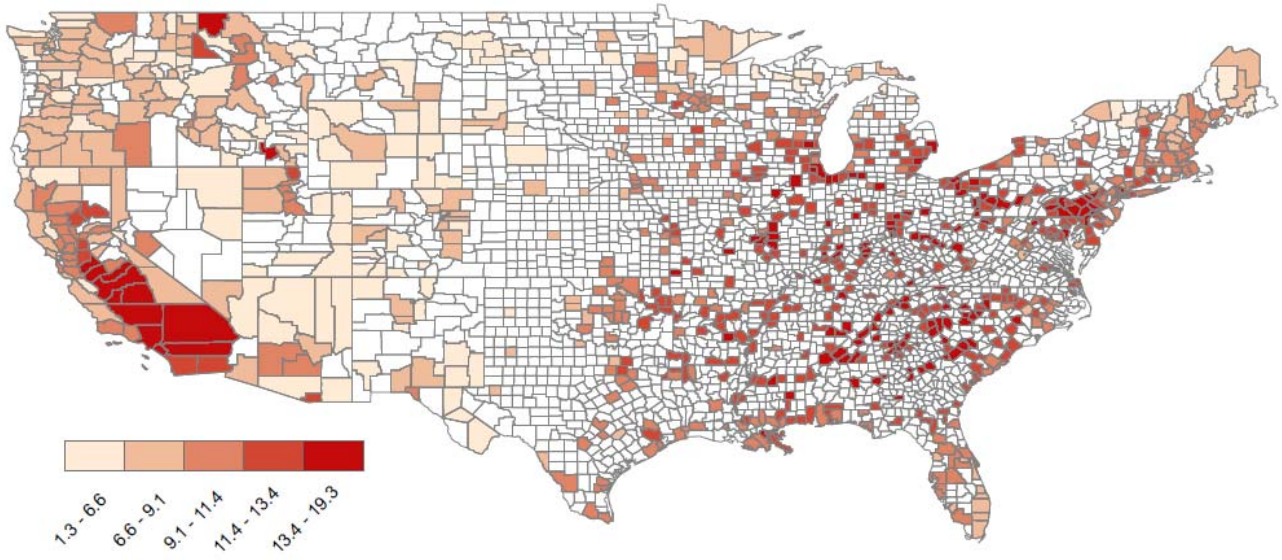


Figure 5. Average PM 2.5 levels in the United States by county, 1999-2011. Units are micrograms per cubic meter. Counties with no data are shaded white.

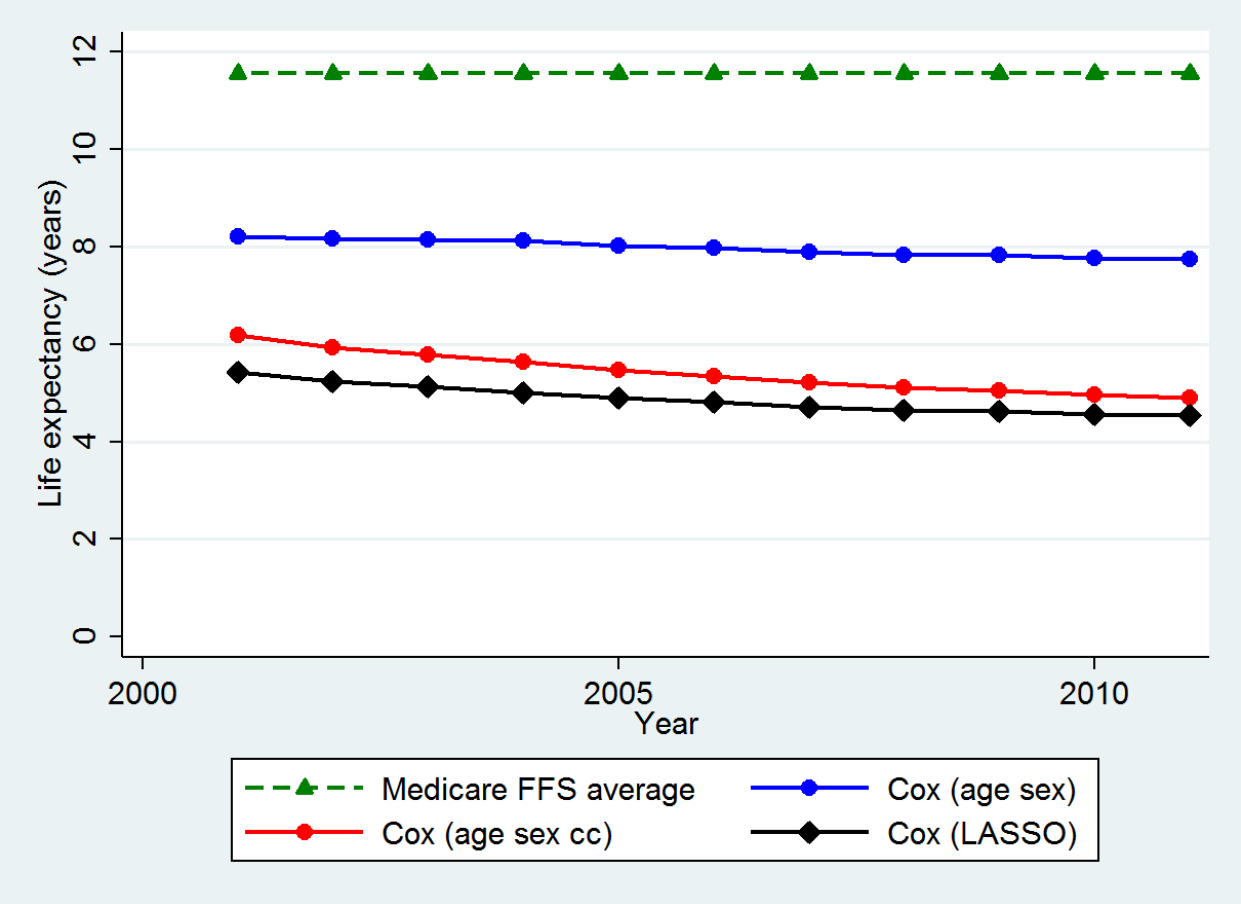


Figure 6. Average ex ante life expectancy for Medicare fee-for-service beneficiaries who later die within one year, by calendar year. Estimates for "Medicare FFS average" are produced by estimating (6) with no covariates. Estimates for "Cox (age sex)" and "Cox (age sex cc) are produced by estimating (6) using age and gender, and age, gender and 27 chronic conditions, as predictors, respectively. Estimates for "Cox (LASSO)" are produced by estimating (7) with 1,062 included regressors.

TABLES

Table 1: Summary statistics, 1999-2011

	Mean	Std. dev.	Obs.
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

Notes: unit of observation is county-day. All rates are per million beneficiaries in the relevant group. Spending and admissions variables are per million FFS beneficiaries.

Table 2: Daily PM 2.5 concentrations and mortality

	(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
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	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. 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; and two leads of these weather controls. OLS estimates also include 2 lags and 2 leads of PM 2.5. IV estimates also include 2 lags and 2 leads of the instruments. Estimates are weighted by the relevant population.

Table 3: Daily PM 2.5 concentrations and hospitalization outcomes

	(1) All inpatient spending	(2) E.R. inpatient spending	(3) Total admit rate	(4) E.R. admit rate	(5) All E.R. rate
			Panel A: OLS estimates		
PM 2.5 ($\mu g/m^3$)	-8,439***	877	-0.560***	0.127**	0.406***
	(1993)	(790)	(0.156)	(0.062)	(0.094)
Dep. var. mean	34,463,220	13,659,597	3,370	1,579	4,159
Observations	1,518,549	1,518,549	1,518,549	1,518,549	1,518,549
Adjusted R-squared	0.518	0.685	0.515	0.695	0.651
			Panel B: IV estimates		
PM 2.5 ($\mu g/m^3$)	17,074*	15,446***	2.034***	1.960***	2.290***
	(10182)	(4151)	(0.714)	(0.336)	(0.394)
F-statistic	237	237	237	237	237
Dep. var. mean	34,463,220	13,659,597	3,370	1,579	4,159
Observations	1,518,549	1,518,549	1,518,549	1,518,549	1,518,549

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. All dependent variables are 3-day measures per million 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 estimates also include 2 lags and 2 leads of PM 2.5, while IV estimates also include 2 lags and 2 leads of the instruments. Estimates are weighted by the number of fee-for-service beneficiaries.

Table 4: Daily PM 2.5 concentrations and life-years lost, IV

		Life-years lost regressions			
	(1) All-age mortality	(2) None	(3) +age, sex	(4) +chronic cond.	(5) LASSO
PM 2.5 ($\mu g/m^3$)	0.746*** (0.085)	8.730*** (1.074)	6.013*** (0.811)	3.549*** (0.592)	2.724*** (0.547)
F-statistic	238.7	238.7	238.7	238.7	238.7
Dep. var. mean	462	5,323	3,614	2,439	2,241
LYL per decedent	NA	11.523	7.825	5.281	4.851
LYL per complier	NA	11.698	8.057	4.755	3.650
Observations	1,518,549	1,518,549	1,518,549	1,518,549	1,518,549

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 3-day life-years lost per 1,000,000 fee-for-service beneficiaries. Column 1 shows the raw mortality rate estimates. Column 2-6 headings display what variables were used to predict counterfactual life expectancy (see appendix for details). 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.

Table 5: Daily PM 2.5, other pollutants, and elderly mortality, IV

	(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	117.960	33.022	49.219	26.993
Dep. var. mean	390.642	390.642	390.642	390.642
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	110.893	31.405	45.387	24.763
Dep. var. mean	461.907	461.907	461.907	461.907
Observations	490,413	490,413	490,413	490,413

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. Dependent variable is the all-age 3-day 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; flexible controls for temperatures, precipitation, and wind speed; two leads of the instruments and the weather controls; and two lags of the instruments. Estimates are weighted by the county's Medicare population.

Table 6: 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 ($\mu g/m^3$)	0.607***	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	393.494	137.564	205.253	325.516	530.916	1169.858
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. 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 2 lags and 2 leads of the instruments. Estimates are weighted by the relevant population.

Table 7: 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 ($\mu g/m^3$)	-0.674	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. 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: Robustness of all-age mortality estimates to different fixed effects and weather controls, IV

	(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	373.932	269.053	355.295	363.147	385.380	228.238	231.385	247.395
Dep. var. mean	393.523	393.523	393.523	393.523	393.524	393.494	393.494	393.494
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. Standard errors (in parentheses) clustered by county. Outcome variable is all-age mortality, measured as the number of deaths per 1,000,000 people. Estimates are weighted by the number of Medicare beneficiaries.

Table 9: Robustness of all-age mortality results to more/fewer instrument lags

	(1) No lags	(2) 1 lag	(3) 3 lags	(4) 4 lags	(5) 5 lags
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	316.310	246.501	241.669	240.882	239.136
Dep. var. mean	393.682	393.576	393.422	393.362	393.315
Observations	1,624,689	1,612,384	1,590,074	1,579,878	1,570,025

Significance levels: * 10 percent, ** 5 percent, *** 1 percent. Standard errors (in parentheses) clustered by county. Dependent variable is the all-age 3-day mortality rate per million Medicare beneficiaries. 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 the instruments; and the indicated number of instrument lags. Estimates are weighted by the number of Medicare beneficiaries.

APPENDIX FOR ONLINE PUBLICATION ONLY

Relationship between wind direction and PM 2.5

We illustrate the variation that drives our results by estimating the following regression separately for each of the 100 monitor groups described in the main text:

$$\text{PM2.5}_{cdmy} = \sum_{b=0}^{34} \beta_b WINDDIR_{cdmy}^{10b} + f\left(Temp_{cdmy}, Prcp_{cdmy}, WindSpeed_{cdmy}\right) + \alpha_c + \alpha_{sm} + \alpha_{my} + \epsilon_{cdmy} \quad \text{(A1)}$$

The variables are defined as in equation (2) of the main text. The estimates $\hat{\beta}_b$ are plotted in Appendix Figure 1. The Bay Area in California (cluster 14) and the Boston, MA area (cluster 68) are reproduced in Figures 1 and 2 from the main text.

These regression equation (A1) differs slightly from the first stage of the instrumental variable regressions estimated in the main text: (1) it employs 10-degree bins for $WINDDIR$ instead of 90-degree bins; (2) it does not include leads and lags of the; and (3) it does not employ county weights. In addition, the large number of control variables included in equation (A1) causes estimation to be impossible for a small number of the monitor groups (see notes in Appendix Figure 1). This does not occur in the instrumental variable regressions we estimate in the main text because the first stage in those regressions is estimated simultaneously, not separately, for each monitor group. (This forces the estimated coefficients on the control variables to be constant across all groups, which increases statistical power.)

Medicare sample and mortality data

The baseline sample used in our analysis consists of all Medicare beneficiaries aged 65-100 and is derived from 100% Medicare enrollment information files for years 1999-2011.[1] These annual files include an observation for each beneficiary enrolled in Medicare for at least one day in that calendar year, whether enrolled in Traditional Medicare (fee-for-service) or Medicare Advantage. The enrollment files report a variety of demographic and enrollment variables, including unique beneficiary identifiers that can be used to link individuals over time; monthly indicators for Medicare eligibility; state, county, and ZIP code of residence based on the mailing address for official correspondence; and exact date of birth, date of death, and gender.

[1] The Research Data Assistance Center (ResDAC) provides a helpful overview of the Medicare enrollment information files at http://www.resdac.org/training/workshops/intro-medicare/media/3.

The vast majority of elderly living in the United States are enrolled in Medicare. Panel A of Appendix Figure 2 compares the size of our baseline Medicare sample to Census estimates of the U.S. population age 65 and over. To aid comparison, we use Census estimates of the resident population on July 1 each year and limit the Medicare sample to beneficiaries residing in the 50 states and the District of Columbia and who turned 65 before July 1. Over the period 1999-2011, the Census estimates an average of 37.3 million elderly individuals each year, compared to 36.2 million elderly in Medicare. Thus, the Medicare sample covers over 97% of elderly living in the U.S., a share which remains roughly constant over the sample period.

The mortality variables used in our analysis are based on dates of death recorded in the Medicare enrollment files. Medicare's death data come primarily from the Social Security Administration but are augmented based on reviews triggered by hospitalization claims indicating patient death. The annual mortality rates in the Medicare data align closely with mortality rates based on national vital statistics death records and Census population estimates, as shown in Appendix Figure 2, Panel B. While all recorded deaths in the Medicare data are validated, some death *dates* in the data are not validated and are assigned the last date in the month of death. Because much of our analysis is performed at the daily level, we drop individuals who die at any point in the year and who do not have a validated death date flag. This restriction affects less than 2% of the deaths in our sample, and the share of deaths with unvalidated dates diminishes over time (see Appendix Figure 2).

Estimating counterfactual life expectancy

We model counterfactual life expectancies for Medicare beneficiaries by estimating a semi-parametric Cox proportional hazards model.[2] This model assumes that the hazard rate of death for individual i can be factored into two separate functions:

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

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:

[2] We have also estimated fully parametric models that assume survival rates are governed by either the Gompertz or Weibull distributions. Those results are very similar.

$$\ln L(\beta) = \sum_{i=1}^{N} \delta_i \left[x_i'\beta - \ln \sum_{j \in R(t_i)} \exp[x_j'\beta] \right] \tag{A2}$$

where the indicator variable 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 and consists of all individuals who are alive at that time. Thus, individuals whose deaths we do not observe (censored observations) affect the partial likelihood function only through the terms indexed by j in equation (A2).

Once $\hat{\beta}$ has been obtained by maximizing the log partial likelihood, we nonparametrically estimate the baseline hazard function following Breslow (1972):

$$\hat{h}_0(t_i) = \frac{d_{t_i}}{\sum_{j \in R(t_i)} \exp[x_j'\hat{\beta}]} \tag{A3}$$

The numerator, d_{t_i}, is the number of deaths that occur at t_i. The corresponding baseline survival function is calculated as

$$\hat{S}_0(t_i) = \exp[-\hat{H}_0(t_i)]$$

where $\hat{H}_0(t_i)$ is the cumulative hazard function, calculated as $\hat{H}_0(t_i) = \sum_{\tau=1}^{t_i} \hat{h}_0(\tau)$. The individual-specific survival function, which allows us to calculate life expectancy, can then be estimated as:

$$\hat{S}(t_i|x_i, \hat{\beta}) = \hat{S}_0(t_i)^{\exp[x_i'\hat{\beta}]}$$

In practice, the nonparametric estimate of the baseline hazard function is limited to the ten years of Medicare data we have available for this survival analysis. We extrapolate the baseline hazard function to future years by assuming it follows a log-linear form. As shown in Appendix Figure 3, this appears to be a very reasonable assumption.

We estimate the Cox proportional hazards model (A2) using data from the 2002 cohort of Medicare beneficiaries, which we observe from January 1st, 2002. We observe all deaths that occur among this cohort on or before December 31, 2011. During this 10-year time period, 50 percent of our sample dies; the

remaining deaths are censored.[3] To ensure that we have 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. For computational ease, we further limit the analysis to a random 5 percent sample of these beneficiaries. The final estimation sample consists of 1,211,585 individuals.

The life-years lost analysis presented in the main text varies the set of individual characteristics included in the vector x_i in order to understand whether they affect the results. As described in the text, we use an increasing number of characteristics as we move across columns in Table 4. Column (2) includes no characteristics; column (3) includes age and sex, and column (4) includes age, sex, and indicators for 27 different chronic conditions. As we describe in detail below, column (5) utilizes a machine learning algorithm to optimally incorporate information from the large set of additional variables available in the Medicare dataset.

The Medicare dataset includes thousands of additional variables, any of which could be significant predictors of survival. Including all of them in a standard survival model 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. This may cause bias due to overfitting (Harrell et al. 1996). Moreover, computational limitations prevent us from including thousands of regressors when performing conventional maximum likelihood estimation using standard numerical procedures.

Recent advances in machine learning techniques overcome these challenges. These methods are well-suited to settings like ours, where the researcher has access to a large number of variables and cares only about predictive accuracy. Column (5) of Table 4 presents estimates of life-years lost when we estimate life expectancy with a Cox proportional hazards model using the Least Absolute Shrinkage and Selection Operator (LASSO) estimator (Tibshirani 1997).[4] Specifically, we maximize a penalized version of objective function (A2):

$$\ln L(\beta) = \left(\sum_{i=1}^{N} \delta_i \left[x_i'\beta - \ln \sum_{j \in R(t_i)} \exp[x_j'\beta] \right] \right) - \lambda \sum_{i=1}^{k} |\beta_i| \tag{A4}$$

[3] Although earlier cohorts are observable for a longer period of time, 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.

[4] We also tried other estimators like ridge regression. The results are similar.

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 (number of elements in β). The optimal penalty parameter λ is selected using 5-fold cross validation.[5] We include the following 1,062 regressors (not including omitted categories) when estimating this model of life expectancy:[6]

1. Age in days as of January 1, 2002
2. Indicator variables for sex and for 7 different races
3. Indicator variables for the presence of the following 27 different chronic conditions as of December 31, 2001: acute myocardial infarction, Alzheimer's disease, senile dementia, atrial fibrillation, cataracts, chronic kidney disease, chronic obstructive pulmonary disease (COPD), heart failure, diabetes, glaucoma, hip/pelvic fracture, ischemic heart disease, depression, osteoporosis, rheumatoid arthritis, stroke, breast cancer, colorectal cancer, prostate cancer, lung cancer, endometrial cancer, anemia, asthma, hyperlipidemia, benign prostatic hyperplasia, hypertension, and hypothyroidism
    a. Indicator variables for all pairwise interactions of these 27 chronic conditions
4. Indicator variables for the interaction of 27 chronic conditions with 7 race indicators
5. Indicator variables for the interaction of 27 chronic conditions with sex
6. Indicator variables for 12 quantiles (10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 99, 99.9) of the *beneficiary's* total prior year spending (i.e., spending that excludes payments made by Medicare)
    a. Indicator variables for the same 12 quantiles for each of the following 17 different categories of *total* prior year medical spending: hospice, home health care, hospital outpatient, acute inpatient, other inpatient, skilled nursing facility, ambulatory surgery center, Part B drugs, evaluation and management, anesthesia, dialysis, other procedures, imaging, tests, durable medical equipment, other Part B carrier, Part B physician
7. Indicator variables for various quantiles (listed in parentheses) of the total annual number of:
    a. Dialysis events (10, 30, 50, 70, 90)
    b. Home health visits, hospital outpatient emergency room visits (10, 30, 50, 70, 90, 95)
    c. Anesthesia events, hospital outpatient visits, other Part B carrier events, acute inpatient stays, durable medical equipment (10, 30, 50, 70, 90, 99)

[5] See Simon et al. (2011) for a detailed discussion of how to implement the Cox proportional hazards estimator with a LASSO penalty term.
[6] Variable names correspond to the descriptions given by ResDAC: http://www.resdac.org/cms-data/files/mbsf/data-documentation

d. Part B drug events (10, 50, 70, 90, 99, 99.5)
e. Other procedures events, evaluation and management events, imaging events, hospital outpatient emergency room visits, tests events, Part B physician events (10, 30, 50, 70, 90, 99, 99.5)

8. Fourth-order polynomials in each of 37 different variables that have been merged to the respondent's 5-digit ZIP code of residence. All variables are standardized so that they follow a normal distribution with mean 0 and variance 1. These zipcode-level data are obtained from the 2007-2011 and 2008-2012 American Community Surveys. The variables include data on the following categories (number of variables in parentheses if more than one): travel time to work (2), fraction below the poverty line (3), median household income, aggregate household income, aggregate household social security income, aggregate household retirement income, fraction in labor force, heating fuel sources (3), aggregate number of vehicles, median home value, fraction immigrant, gini index of household income, fraction with less than high school education, median year housing built, fraction on disability (2), fraction with hearing difficulties (2), fractions with vision difficulty (2), fraction with cognitive difficulty (2), fraction with ambulatory difficulty (2), fraction with self-care difficulty (2), fraction with independent-living difficulty (2), fraction with any health coverage, (2)and fraction with private health coverage (2).

The counterfactual life expectancy that forms the basis of the estimate in Column (5) of Table 4 is based on estimating (A4) when including the 1,062 regressors listed above.

The dashed lines in Appendix Figure 4 show the distribution of estimated counterfactual life expectancies for the subsample of Medicare beneficiaries who were used to estimate our survival model. The range of the distribution is wider when the model includes all 1,062 predictors (the dashed black line) than when it includes only age and gender as predictors (the dashed red line). The model based on age and gender corresponds to a typical life table and consists of only 68 $(= (100 - 67 + 1) \times 2)$ values. The maximum and minimum values in this table correspond to life expectancies for a 67-year-old female and a 100-year-old male, respectively. By contrast, the LASSO model generates a much larger set of predictions, some of which lie outside of the range of a basic life table.

The solid lines in Appendix Figure 4 show how the distribution of predicted values changes when it is limited to the subset of beneficiaries who died during the 2002 calendar year. The distribution produced by the model that includes only age and gender—given by the solid red line—shifts to the left because these decedents are older than the average Medicare beneficiary and thus have below-average life expectancies. The distribution for the LASSO model—given by the solid black line—shifts to the left even more. This

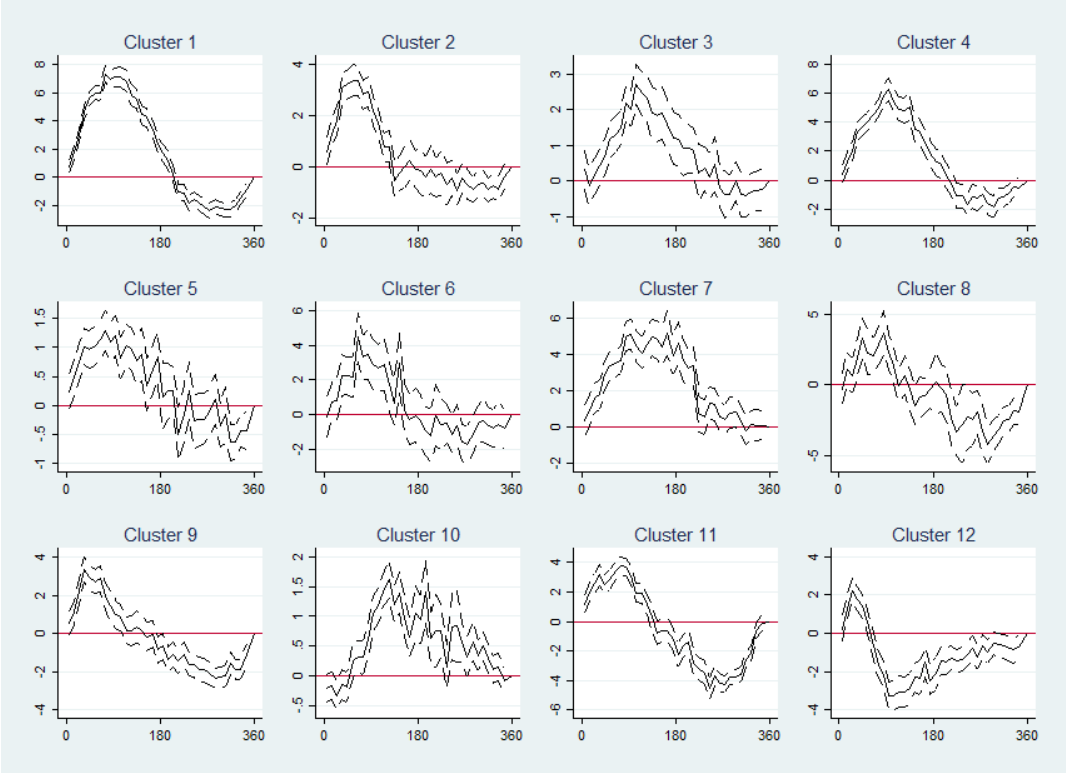
indicates that beneficiaries who died within one year of January 1, 2002 were not only older than the average beneficiary in that year, but also they were less healthy than average, as captured by variables like prior medical spending and prior chronic conditions. Accounting for these additional variables reduces (on average) the prediction of the counterfactual life expectancies for these Medicare beneficiaries. This demonstrates that the Cox LASSO model that incorporates data from many variables generates predictions that are more accurate than a simple Cox model that accounts only for age and gender.

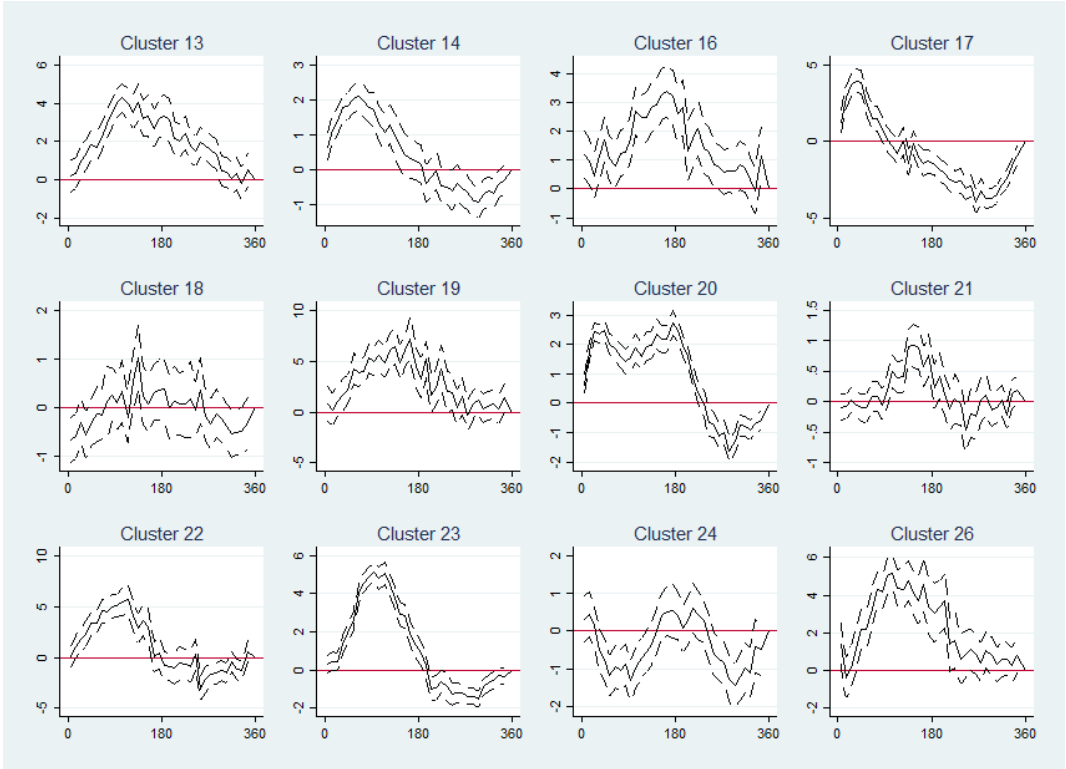
We can perform a similar exercise that incorporates Medicare data from individuals not included in our estimation sample. We first use the results from our estimation to predict life expectancy for Medicare beneficiaries as of January 1 of each calendar year. For each of these years, we then calculate the average life expectancy for all fee-for-service beneficiaries who die during that year ("decedents"). We focus on this group because these decedents form the basis for the life-years lost estimates reported in Table 4.

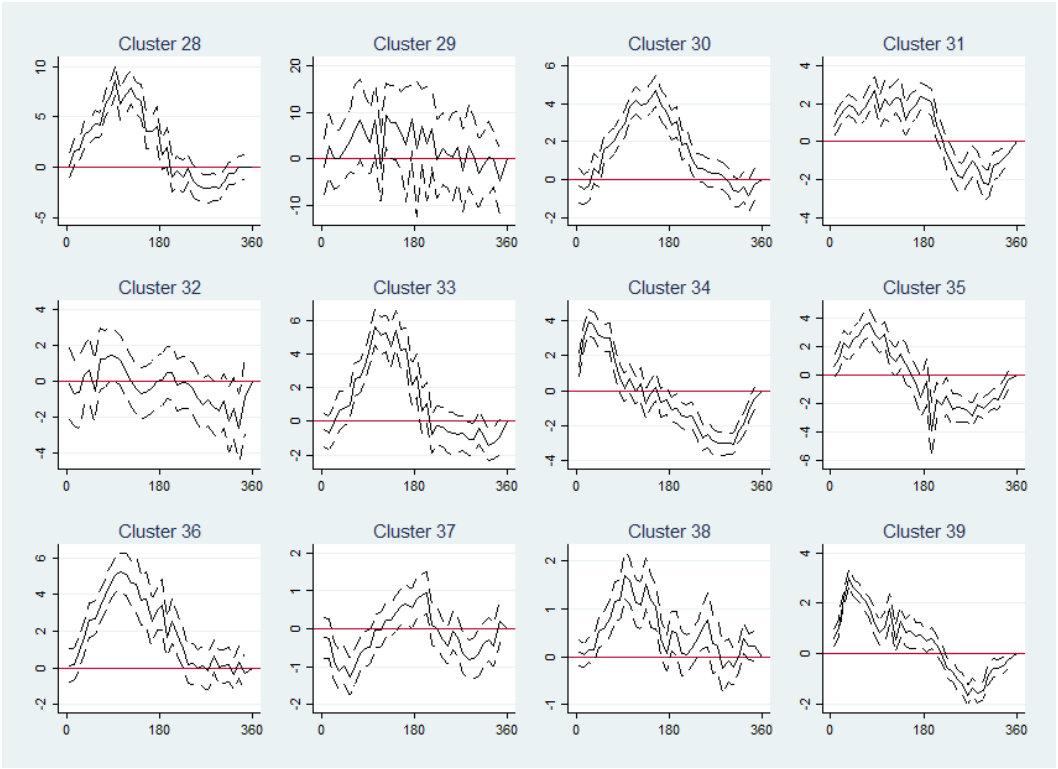
Appendix Figure 5 displays the results of this exercise. The solid green line, which serves as a baseline, displays our estimate of the unconditional life expectancy (11.4 years) for all Medicare beneficiaries. The solid red line displays the average life expectancy among decedents, as predicted by a Cox proportional hazards model that conditions on age and gender. Because the typical decedent is older than the average beneficiary, the predictions from this model are about 2.5 years lower than the baseline. This is clearly a more accurate prediction, since these decedents by definition died ex post within one year. For the sake of comparison, we also include predictions based on a period life table published by the Social Security Administration. Because that life table also conditions on age and sex, its predictions are nearly identical to those of the Cox model. Finally, the solid black line displays estimates based on the LASSO estimation of the Cox proportional hazards model with 1,062 regressors. This reduces the prediction by yet another 2.5 years. The estimates decline slightly over time, which likely reflects the improvement in the recording of chronic conditions in later years.[7]

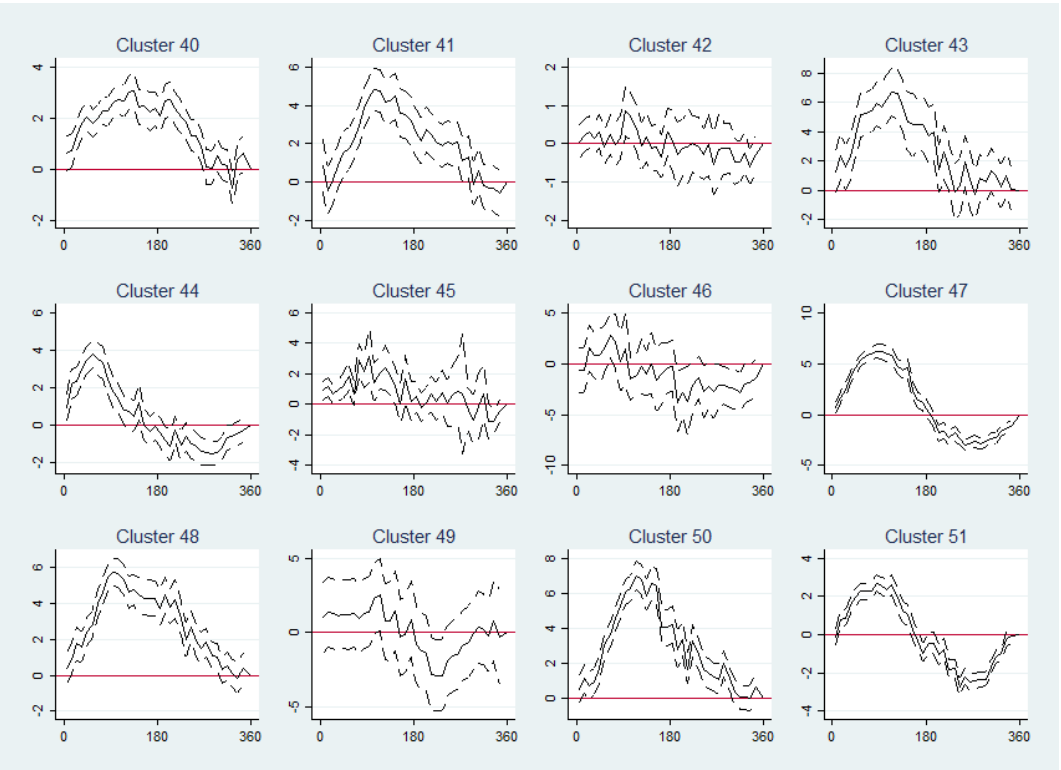
[7] Medicare data on chronic conditions become increasingly incomplete in earlier years beginning in 2006. Because beneficiaries in these earlier years are less likely to have their chronic conditions recorded in the data, their estimated life expectancy is higher than beneficiaries in later years, who are more likely to have chronic conditions.

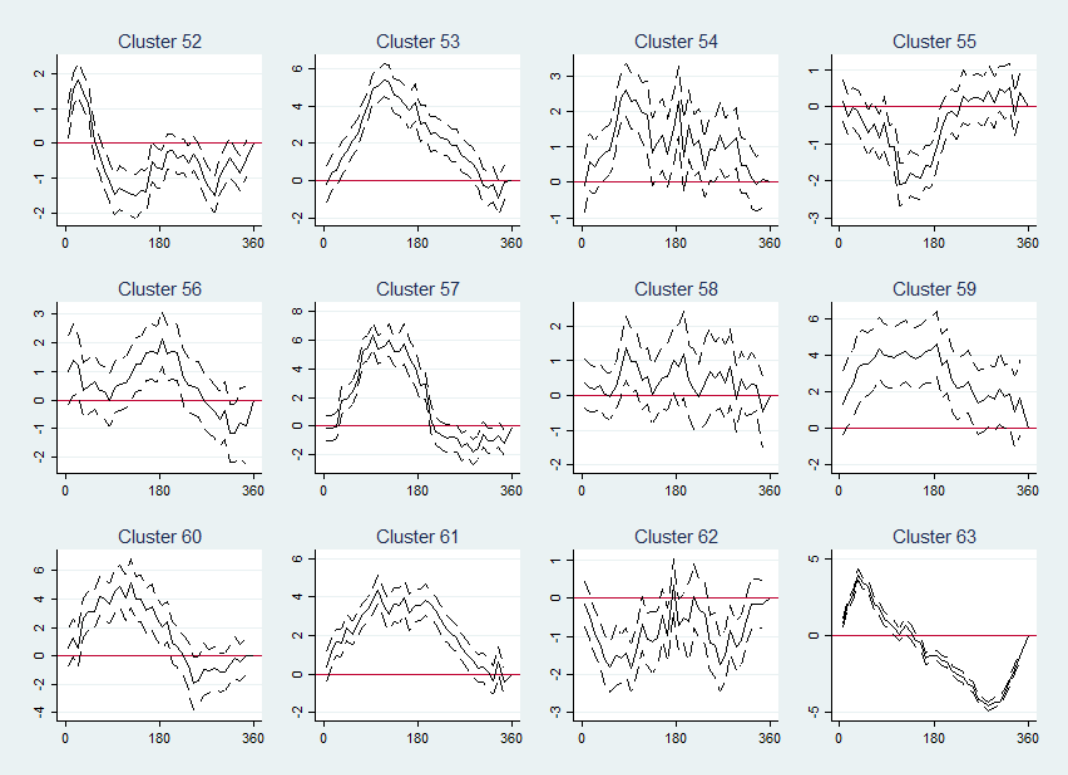
APPENDIX FIGURES

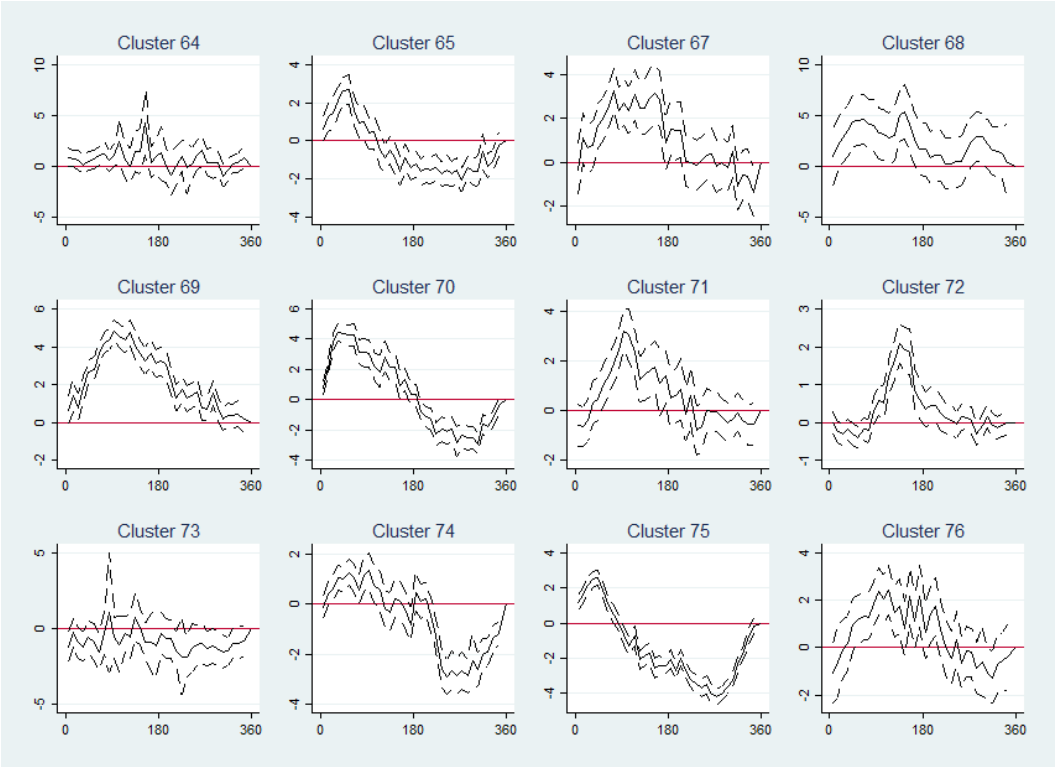


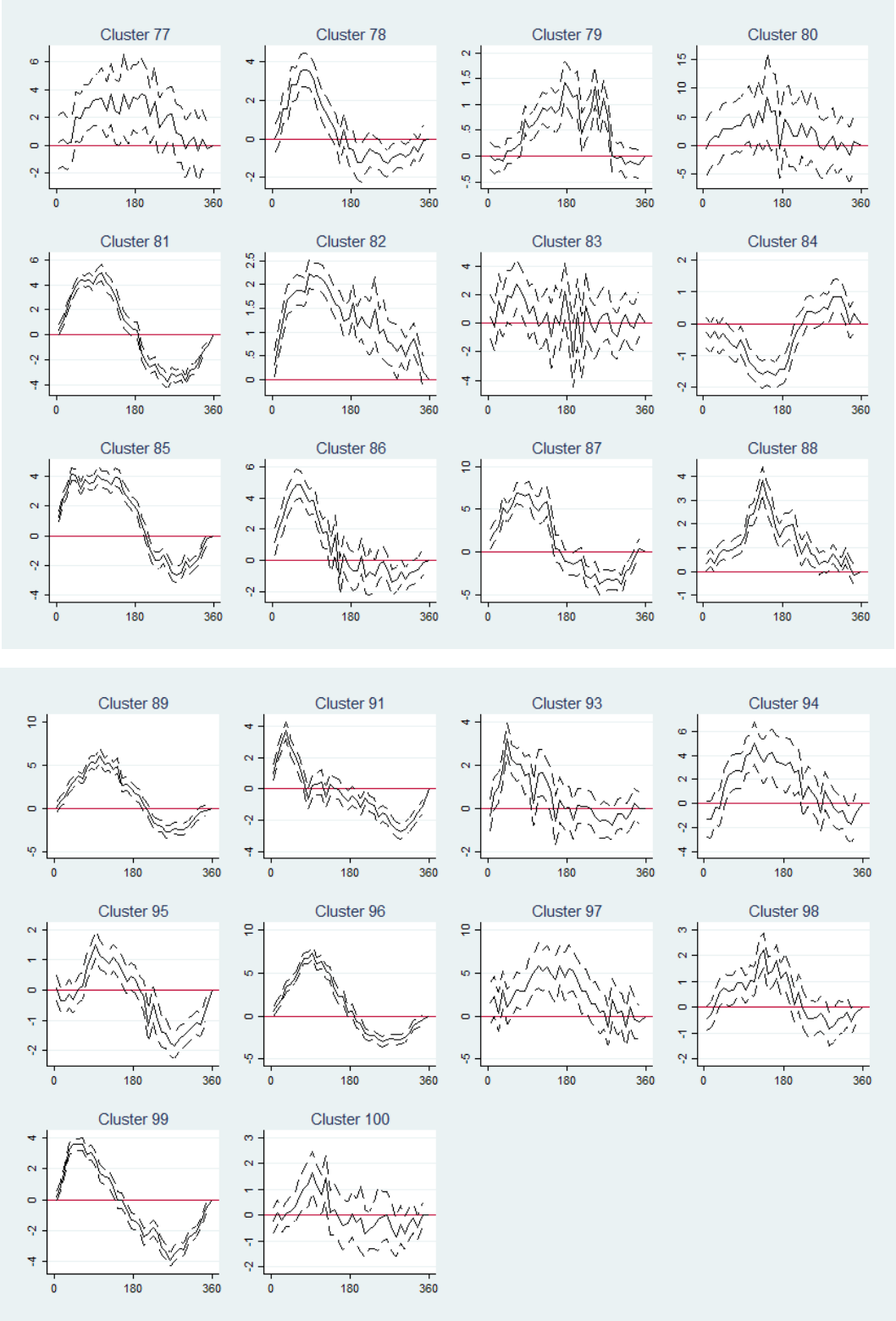



Cluster 28
Cluster 29
Cluster 30
Cluster 31
Cluster 32
Cluster 33
Cluster 34
Cluster 35
Cluster 36
Cluster 37
Cluster 38
Cluster 39

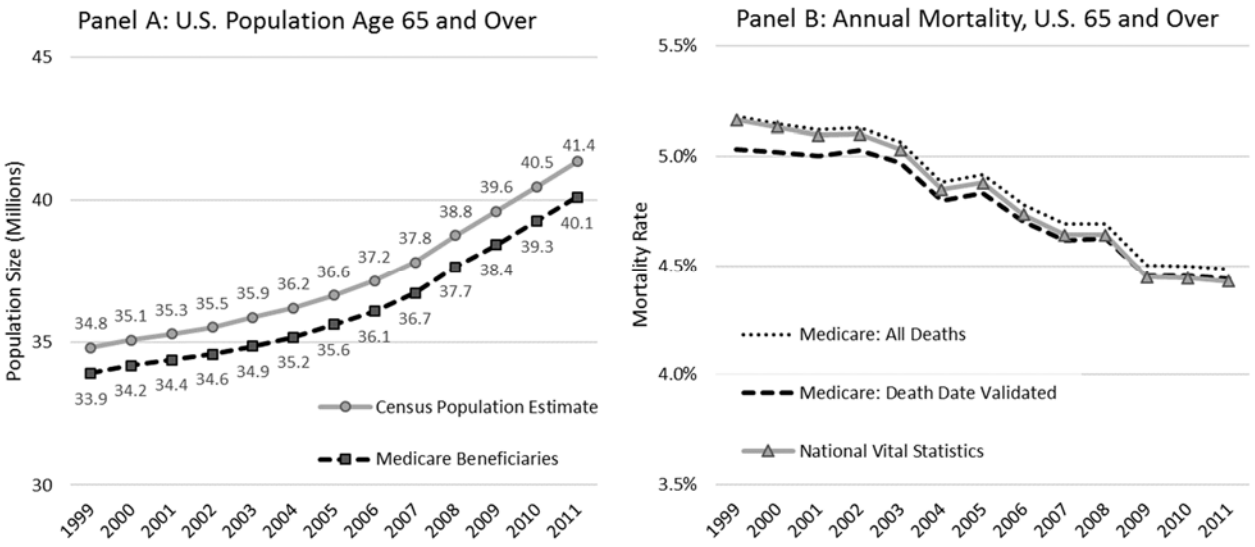

Cluster 40
Cluster 41
Cluster 42
Cluster 43
Cluster 44
Cluster 45
Cluster 46
Cluster 47
Cluster 48
Cluster 49
Cluster 50
Cluster 51


Cluster 52
Cluster 53
Cluster 54
Cluster 55
Cluster 56
Cluster 57
Cluster 58
Cluster 59
Cluster 60
Cluster 61
Cluster 62
Cluster 63


Cluster 64
Cluster 65
Cluster 67
Cluster 68
Cluster 69
Cluster 70
Cluster 71
Cluster 72
Cluster 73
Cluster 74
Cluster 75
Cluster 76



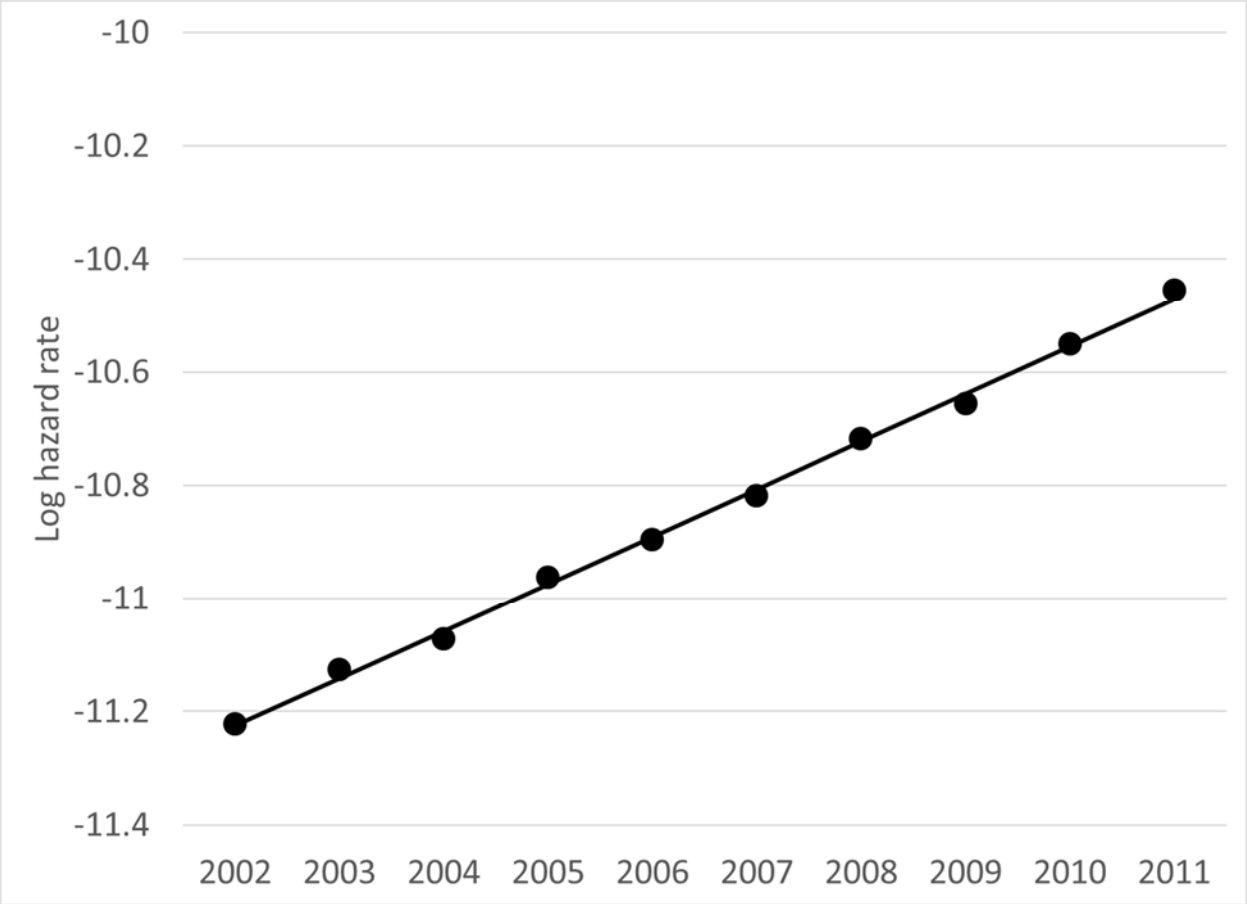
Appendix Figure 1. Relationship between wind direction and PM 2.5 concentrations for each monitor group. Pollution monitor groups with less than 1,000 PM 2.5 readings are not shown.



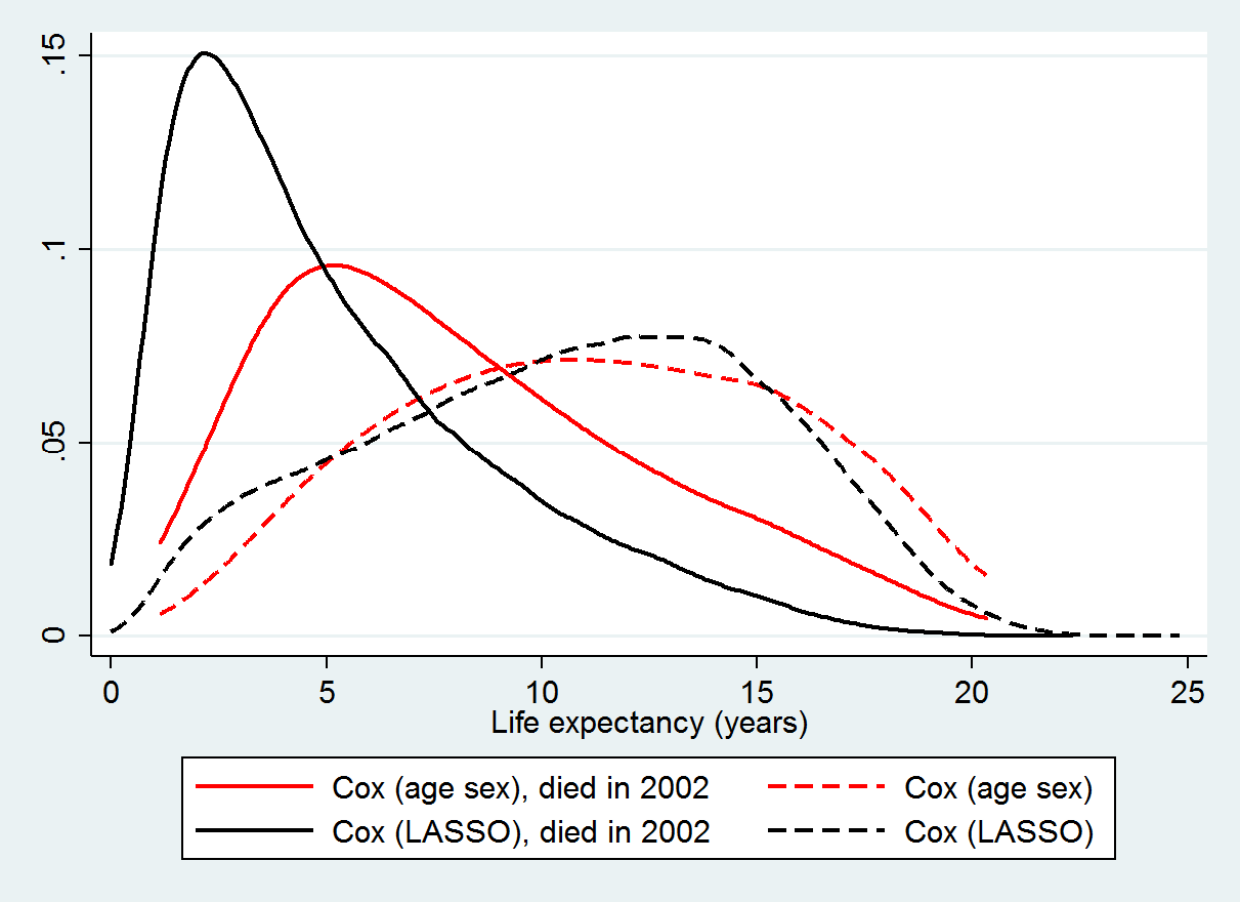
Appendix Figure 2. Population and Mortality Among U.S. Elderly, 1999-2011.

Panel A: Census population estimates are derived from U.S. Census Bureau files. Estimates for 1999-2009 are intercensal estimates of the July 1 resident population age 65 and over; estimates for 2010-2011 are postcensal estimates of the July 1 resident population age 65 and over. Medicare beneficiaries for a given calendar year include all individuals age 65-100 in the corresponding annual Medicare enrollment file, limited to those who turned 65 before July 1 of the year and have a ZIP code of residence located in the 50 states or the District of Columbia.

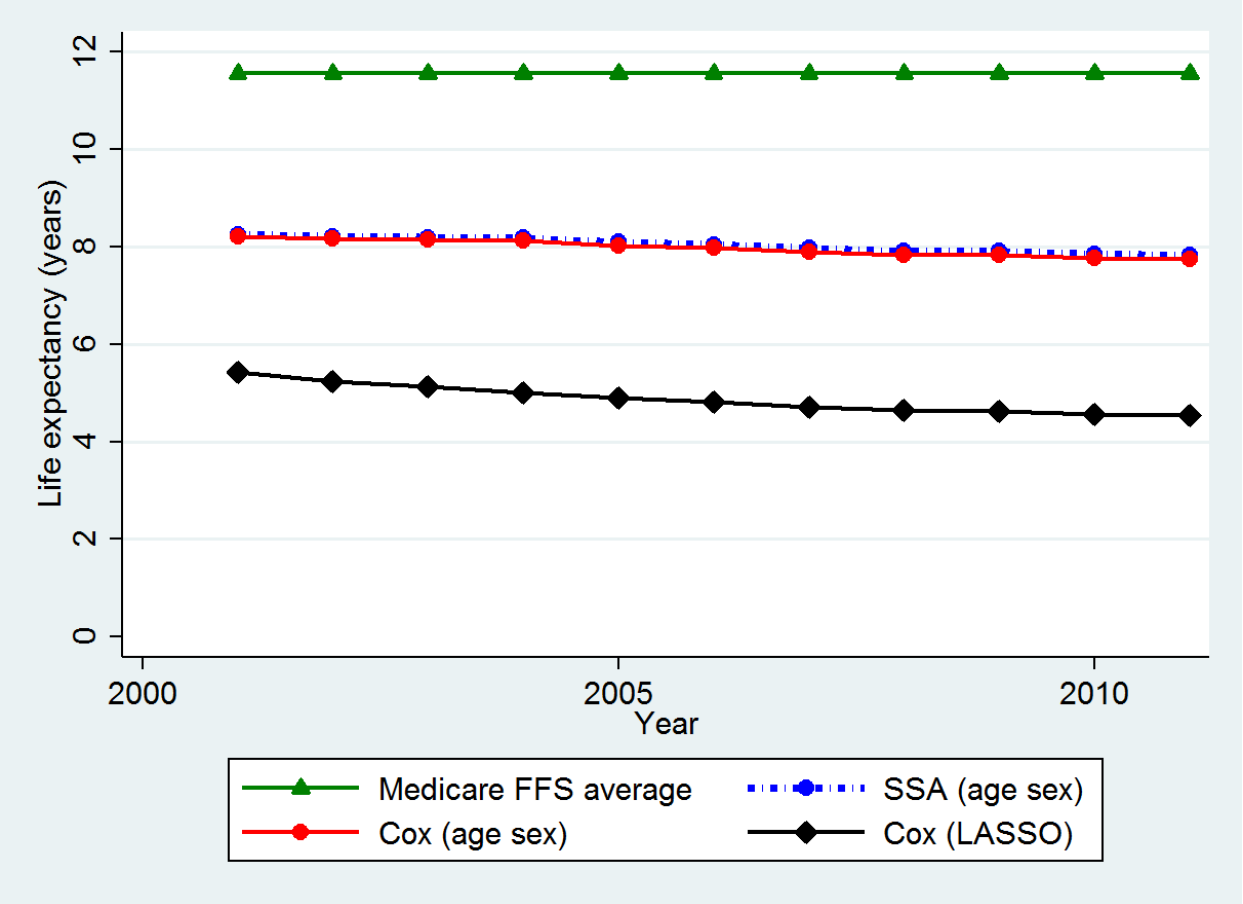
Panel B: National vital statistics mortality data come from the Compressed Mortality File (CMF), which is produced by the National Center for Health Statistics and is based on death certificates filed in the 50 states and the District of Columbia. To obtain vital statistics mortality rates, we divide total CMF deaths among the 65 and over population in a given year by the Census population estimates shown in *Panel A*. The dashed lines report annual mortality rates based on death dates recorded in the Medicare annual enrollment files. The figure reports both the total mortality rate in the Medicare sample, as well as the mortality rate among the analytical sample used in the paper which excludes individuals who have a validated death that year but do not have a validated death *date* flag.



Appendix Figure 3. Estimated baseline hazard rate for the 2002 Medicare cohort. The points represent the log of the baseline hazard rate, as estimated by equation (A2), when we estimate a Cox proportional hazards model with age and gender as controls for the cohort of Medicare beneficiaries who were alive in 2002. Plots of the baseline hazard rate for survival models with different regressors look similar.



Appendix Figure 4. Kernel density plot of life expectancy estimates for Medicare beneficiaries alive on January 1, 2002. The dashed lines display the distributions of life expectancy for all Medicare beneficiaries alive on January 1, 2002. The solid lines limit the distribution to the subset of those beneficiaries who later died during the 2002 calendar year. The red lines display estimates from a Cox proportional hazards model that includes only age and gender as regressors. The black lines display estimates generated by estimating model (A3) using LASSO with $1{,}062$ regressors.



Appendix Figure 5. Average ex ante life expectancy for Medicare fee-for-service beneficiaries who later die within one year, by year. Estimates for "Medicare FFS average" are produced by estimating (A1) with no covariates. Estimates for "Cox (age sex)" are produced by estimating (A1) using only age and gender as predictors. Estimates for "Cox (LASSO)" are produced by estimating (A3) with 1,062 included regressors. Estimates for "SSA (age sex)" are obtained from the 2011 period life table for the Social Security area population (source: https://www.ssa.gov/oact/STATS/table4c6.html, accessed August 7, 2015).