The Long-run Effect of Air Pollution on Survival*

Tatyana Deryugina and Julian Reif
University of Illinois and NBER

April 2024

Abstract

Many environmental hazards have long-run health effects, but quasi-experimental studies typically measure outcomes and treatment over short time periods. We develop a new framework for quantifying the effect of air pollution exposure on life expectancy. Using changes in wind direction as an instrument for daily sulfur dioxide levels, we first characterize the dynamic mortality effects of short-run exposure. We then incorporate these estimates into a demographic model to quantify the lifelong effects of a permanent reduction in air pollution exposure. Ninety percent of the survival benefits accrue after the first fifty years of life.

JEL Classification: I18, Q52, Q53

Keywords: Air pollution, chronic exposure, mortality, survival

*Deryugina: deryugin@illinois.edu. Reif: jreif@illinois.edu. Lexi Chen, Prakrati Thakur, and Xinhui Sun provided excellent research assistance. We are grateful to Josh Gottlieb, Peter Hull, Adriana Lleras-Muney, Nolan Miller, David Molitor, Nick Muller, Nicole Riemer, Hannes Schwandt, Matthew West, Nikos Zirogiannis, and seminar participants at Arizona State University, Baltic Economic Association, Cornell University, DePaul University, Exeter University, GSSI School of Advanced Studies, Iowa State University, ITAM, IZA Workshop on Environment and Labor Markets, Jinan University, Johns Hopkins University, London School of Economics, MHEC, Nova SBE, Oxford University, Santa Clara University, Southern Economic Association meetings, Southern Illinois University, Stanford University, Tinbergen Institute, Toulouse School of Economics, University of Birmingham, UC Berkeley, University of Chicago, University of Duisburg-Essen, University of Geneva, UIC, University of Indiana, University of Mannheim, University of Navarra, University of Pennsylvania, and Wake Forest University for helpful feedback. This research was supported by the National Institutes of Health under Award Numbers R01AG053350 and R01AG073365.

1 Introduction

Environmental hazards such as air pollution, extreme temperatures, and water pollution are important causes of human morbidity and mortality. For example, The Lancet Commission on Pollution and Health estimates that air pollution caused 6.5 million premature deaths in 2015, amounting to about 12 percent of all deaths worldwide (Landrigan et al., 2018). Such assessments are generally based on observational studies, which are prone to estimation bias (Dominici, Greenstone and Sunstein, 2014). Although quasi-experimental studies can address this bias, they typically measure health outcomes and treatment exposure over short time periods that span less than one year, thereby overlooking effects that may emerge years or decades later. This limitation is challenging to overcome because quasi-experimental variation in chronic (long-run) treatment exposure is difficult to find, data that track individuals over long time periods are rare, and endogenous responses such as migration complicate the interpretation of estimates. However, optimal health and environmental policies require an accurate estimate of the lifelong health consequences of permanent changes in exposure.

We propose a novel approach to overcome this limitation when the outcome is mortality. The approach combines well-identified short-run estimates with an individual-level model of health production that fits human survival curves well and accommodates a wide variety of mortality dynamics (Lleras-Muney and Moreau, 2022). We show how to calibrate the model using empirical mortality estimates and then perform counterfactuals that can fully characterize the short- and long-run survival effects of both acute and chronic changes in exposure. We validate the calibrated model by comparing its short-run predictions to empirical estimates of outcomes measured using time windows or age groups that differ from those used for calibration, and by comparing its three-year projections to corresponding estimates from Anderson (2020).

We use our approach to estimate the short- and long-run effects of exposure to air pollution on US population mortality. Our study focuses on sulfur dioxide (SO_2), a major source of fine particulate matter and the predominant pollutant measured in the decades following

the 1970 Clean Air Act. We assemble a new dataset that combines the universe of publicly available daily-level death records (1972–1988) with data on air pollution and weather. We first investigate the causal effect of acute (1-day) air pollution exposure on county-level mortality by instrumenting for observed changes in SO_2 with changes in local wind direction. We estimate that a 1-unit (≈10 percent) increase in SO_2 raises 1-day mortality by 0.08 deaths per million (0.33 percent). Secondary analyses that incorporate data on other pollutants indicate that our estimates reflect the causal effects of exposure to both SO_2 as well as secondary particulate matter derived from SO_2.

The cumulative mortality effect of this 1-day exposure more than doubles when we extend the outcome window from one day to one month (28 days), demonstrating that air pollution continues to have lethal effects long after exposure has ended. While statistical power limits our ability to estimate precise effects beyond a one-month window, we can show that cumulative mortality remains elevated for at least three months following exposure. These results are consistent with findings from the medical literature suggesting that air pollution causes "accelerated aging" by, for example, hardening arteries and increasing the risk of heart disease (Rajagopalan and Landrigan, 2021).

We document striking differences in mortality dynamics by cause of death. Our 1-day mortality estimate is driven roughly equally by deaths related to three groups of causes: cardiovascular disease, cancer, and "other diseases," a residual category that includes chronic lower respiratory illness and diabetes. Lengthening the outcome window to encompass deaths over the subsequent month causes the estimates for cardiovascular and other diseases to increase by a factor of 3–4, while the estimate for cancer falls by 70 percent and becomes statistically insignificant. This fall in the magnitude of the cancer-related mortality estimate over time implies that the cancer-related deaths occurred predominantly among frail individuals who had short (less than one month) counterfactual life expectancies—a phenomenon called "mortality displacement." Altogether, we conclude that acute exposure to air pollution produces two distinct mortality patterns: mortality displacement in a subpopulation of

frail individuals, where the cumulative mortality effect quickly dissipates, and accelerated aging in a subpopulation of healthier individuals, where the cumulative effect grows over time. On net, the accelerated aging effect dominates.

Medical and epidemiological research suggests that long-run exposure to air pollution also leads to poor health outcomes that only emerge many years later, and are thus very challenging to estimate directly using quasi-experimental methods. In the second part of our paper, we quantify the effect of lifetime exposure on long-run survival by adapting a dynamic production model of health to our daily mortality setting (Lleras-Muney and Moreau, 2022). While there are many ways to model survival, this model is particularly well-suited to our needs because it can accommodate both the accelerated aging and mortality displacement patterns present in our setting.

Our approach maps our short-run empirical estimates to the model's parameters, allowing us to form long-run projections that comply with well-documented age patterns of human mortality. We use age-specific 1-day cancer mortality estimates to calibrate the effect of pollution exposure on the model parameter governing mortality displacement, and use the corresponding non-cancer mortality estimates to calibrate the effect of exposure on the model's aging parameter. We then use this calibrated model to quantify the short- and long-run effects of both acute and chronic exposure.

We validate the model in several ways. Because it is calibrated using only 1-day mortality estimates, we can measure the model's accuracy by comparing its mortality predictions in the month following acute exposure to our corresponding empirical estimates. We also assess the plausibility of a key modeling assumption—that the effect of exposure on model parameters depends only on current exposure—by testing the model's ability to predict mortality for age groups not used in the calibration. For example, we calibrate a model using IV mortality estimates for ages 70 and over and then assess its ability to predict the pollution mortality effects for 65–69-year-olds, a younger age group that has a different lifetime exposure history. We also compare the estimated mortality effects of short-term (up

to 28-day) chronic pollution exposure to the corresponding model predictions.

Overall, we find that the overwhelming majority of model predictions lie inside the 95% confidence intervals of the IV estimates, supporting the calibrated model's validity. These predictions depend meaningfully on our estimate of the fraction of deaths due to mortality displacement: alternatively assuming that acute exposure produces either 0% or 100% mortality displacement yields predictions that lie far outside the 95% confidence intervals of the IV estimates. These results demonstrate that, while all-cause mortality estimates are insufficient for drawing reliable conclusions about longer-term survival, researchers can overcome this problem by appropriately incorporating information about cause of death into a survival model.

Finally, we use our model to quantify the projected effect of a permanent, 1-ppb decrease in SO_2 on life expectancy. The model predicts that, all else equal, this reduction would extend life expectancy at birth by 1.2–1.3 years (90% bootstrap CI, 0.4–2.5), which substantially exceeds the estimate of 0.17 years obtained when extrapolating our one-month IV estimates to the whole life cycle. Although this decrease in chronic exposure begins at birth, ninety percent of the improvements in life expectancy occur after age 50 and over three-quarters occur after age 65. This result suggests that most of the survival benefits of the drastic reductions in US air pollution levels over the past fifty years have yet to emerge for cohorts born after the passage of the 1970 Clean Air Act. Because the model holds behavior fixed, its projections can be interpreted as the total benefits associated with pollution reduction, uncontaminated by longer-run behavioral responses such as migration (Graff Zivin and Neidell, 2012; Currie et al., 2014; Carleton et al., 2022).

The main contribution of our study is the development and application of a new framework for estimating the long-run mortality effects of chronic exposure to environmental hazards. The standard approach estimates the short-run mortality effects of acute exposure and then quantifies long-run mortality effects using population life tables (e.g., Deschênes and Greenstone, 2011), although recent work has improved the accuracy of this method by

incorporating individual-level predictions of counterfactual life expectancy (Deryugina et al., 2019). However, that approach remains prone to bias if unobserved characteristics are correlated with both life expectancy and the probability of dying from exposure, and it is unable to quantify the effects of chronic exposure. By contrast, our study uses a structural model to infer long-run mortality effects from short-run quasi-experimental estimates, providing a novel example of a "best of both worlds" approach that combines structural and experimental methods (Todd and Wolpin, 2023). Because the model leverages well-documented features of human life-cycle mortality patterns, its projections are more reliable than naive extrapolations that fail to incorporate this information.

Our approach complements largely elusive efforts to directly estimate the long-run mortality effects of pollution. Direct estimation requires identifying quasi-experimental variation in long-run exposure and accounting for behavioral responses that alter exposure, such as migration. Ignoring these responses leads to an underestimate of the health costs of air pollution (Graff Zivin and Neidell, 2012; Currie et al., 2014). As a result, only a few credible quasi-experimental studies estimate mortality effects over multiple years (Chen et al., 2013; Ebenstein et al., 2017; Anderson, 2020; Barreca, Neidell and Sanders, 2021; Andersen et al., 2023). Even among these studies, the possibility of migratory responses cannot be easily ruled out.[1] By contrast, our approach employs quasi-experimental variation at the daily level, which alleviates concerns about many forms of avoidance behavior.

Our short-run analysis also makes important contributions to the literature on the health effects of acute exposure to air pollution. Our paper is the largest quasi-experimental study of acute pollution exposure on mortality, encompassing 18 million deaths and enabling us to decompose mortality effects across the entire age distribution with precision. To our knowledge we are the first to show that pollution exposure causes two distinct mortality patterns: mortality displacement among frail individuals, where the cumulative mortality effect quickly

[1] For example, Barreca, Neidell and Sanders (2021) utilize repeated cross-section mortality data at the county-year level. They rule out large changes in total population following a drop in air pollution, but differential migration remains a threat to validity.

falls to zero, and accelerated aging among healthier individuals, where the mortality effect grows with time. These findings underscore that great care must be taken when inferring longer-run mortality effects from short-run estimates. Although lengthening the outcome window to multiple years helps address this challenge, studies with short outcome windows of one year or less make up the vast majority of papers employing quasi-experimental variation in air pollution.[2] Finally, ours is one of the first quasi-experimental studies to estimate the mortality effects of acute exposure to SO_2, a pollutant that has declined substantially in the US over the past fifty years, but which remains elevated in lower-income countries such as China and India.

The rest of the paper is organized as follows. Section 2 provides background on air pollution and describes our data. Section 3 describes our short-run empirical strategy. Section 4 presents estimates of the short-run mortality effects of acute exposure to SO_2 and other air pollutants. Section 5 presents the dynamic model of health production, calibrates it, and quantifies the long-run survival effects of chronic exposure. Section 6 concludes.

2 Background and data

2.1 Air pollution

Sulfur dioxide (SO_2) is a major air pollutant produced primarily by the combustion of coal and oil. In the US, the main source of SO_2 emissions has historically been coal-burning power plants. These emissions fell significantly during our sample period and in more recent years (Figure A.1a), as a result of transitions to low-sulfur coal, increased use of pollution control equipment, and greater reliance on alternative energy sources such as natural gas.

[2]See, for example, Currie and Neidell (2005); Knittel, Miller and Sanders (2016); Schlenker and Walker (2016); Deschênes, Greenstone and Shapiro (2017); Deryugina et al. (2019); Hollingsworth, Konisky and Zirogiannis (2021); Hollingsworth and Rudik (2021); and Heo, Ito and Kotamarthi (2023). A handful of papers study the effect of early-life air pollution exposure on later-life outcomes (e.g., Isen, Rossin-Slater and Walker, 2017; Voorheis, 2017; Colmer and Voorheis, 2020), including one study that considers mortality before age 55 (Arenberg and Neller, 2023). These studies do not consider chronic exposure, however.

SO_2 harms human health through two main channels. First, human clinical trials have shown that direct exposure to SO_2 impairs respiratory function, especially in people with asthma (Agency for Toxic Substances and Disease Registry, 1998). Animal experiments have also demonstrated that SO_2 inhalation can cause brain damage (Sang et al., 2010; Yao et al., 2015) and contribute to cardiac and mitochondrial dysfunction (Qin et al., 2016). Second, SO_2 transforms naturally into sulfate (SO_4^{2-}) at a rate of several percent per hour (Luria et al., 2001). Sulfates are a major component of fine particulate matter ($PM_{2.5}$), a catch-all term for particles whose diameter is 2.5 micrometers (μm) or less. $PM_{2.5}$ is thought to be particularly harmful to health because of its ability to cross the blood-alveolar and blood-brain barriers. Prior quasi-experimental research has found causal links between short-run exposure to $PM_{2.5}$ and a number of health-related outcomes, such as short-run healthcare spending, hospitalizations, and mortality (e.g., Barwick et al., 2018; Deryugina et al., 2019; Heo, Ito and Kotamarthi, 2023).

While research on the exact pathophysiological mechanisms underlying these health effects continues, medical studies have documented significant associations between air pollution and hypertension, diabetes, coronary artery calcification, and the progression of chronic kidney disease, all of which are risk factors for cardiovascular disease (Rajagopalan and Landrigan, 2021). Air pollution has also been linked to the initiation, promotion, and progression phases of lung cancer (Turner et al., 2020; Hill et al., 2023). Once initiated, lung cancer typically grows for over 10 years before it is diagnosed (Nadler and Zurbenko, 2014). Thus, while the quasi-experimental literature has consistently found significant adverse effects of short-run exposure on short-run health, long-run health effects are probably larger.

We measure air pollution using the EPA's Air Quality System database, which provides hourly data at the pollution-monitor level for criteria pollutants regulated by the EPA. The extent of spatial and temporal coverage varies by pollutant (see Appendix Section A.1 for additional detail). Although our analysis focuses on SO_2, we also examine four other air pollutants that have been monitored since the 1970s or the 1980s: nitrogen dioxide (NO_2),

total suspended particulates (TSP), ozone (O_3), and carbon monoxide (CO). TSP comprises all particulates with diameters less than 100 μm, thus including $PM_{2.5}$. Because $PM_{2.5}$ was not consistently monitored until the late 1990s, we cannot include it directly in our analysis.

Panel A of Table 1 shows county-level summary statistics for daily ambient pollution concentrations during our 1972–1988 sample period. The average SO_2 concentration is 9.0 parts per billion, with a standard deviation of 12.6. The average levels of NO_2 and O_3 are higher, at about 21 and 26 parts per billion, respectively. The most prevalent pollutant is CO, with an average concentration of 1.64 parts per million (1,640 parts per billion). We are more than twice as likely to observe SO_2 levels than any of the other four pollutants.

2.2 Mortality

We obtain daily death counts from the National Vital Statistics. These data are based on death certificate records and include information on the cause of death and the county of occurrence. We focus our analysis on the years 1972–1988, the only timeframe for which we have access to publicly available information on the exact date of death.[3] We calculate death rates by dividing death counts by annual population estimates provided by the Surveillance, Epidemiology, and End Results (SEER) Program.

Figure A.2 reports annual death rates by age group and cause of death during our sample time period. The infant mortality rate steadily declines over time, and by 1988 nearly equals the average death rate for ages 45–64. Panel B of Table 1 summarizes daily mortality rates for various subgroups over this time period. The all-age daily death rate is about 25 per million. The rate is higher for infants (33 deaths per million), and much higher for those over age 85 (443 deaths per million).

We classify causes of death into four main categories: cardiovascular, cancer, external, and other. Panel B of Table 1 reports that cardiovascular disease is the leading cause of death during our time period, accounting for nearly half of all deaths in our sample (12 daily deaths

[3]The exact date of death is unavailable prior to 1972, and is available after 1988 only in the Research Data Center of the National Center of Health Statistics.

per million). Cancer deaths make up slightly more than twenty percent of overall mortality (5 deaths per million). External causes of deaths are responsible for about eight percent of all deaths and include car accidents, poisonings, suicides, and other causes not originating in the body. We group the remaining twenty percent of deaths, which includes deaths from respiratory illness, into the "other" category. We report estimates for subcategories of cardiovascular and other disease deaths in secondary analyses.

2.3 Wind and weather

Our empirical strategy is motivated by the well-known fact that wind currents carry air pollution over long distances. For example, regional contributions to sulfate substantially exceed local contributions in many US cities (Environmental Protection Agency, 2004). In theory, one could use predictions from an atmospheric model of wind transport as an instrument for changes in local pollution levels. In practice, doing so at the daily level is computationally intensive and requires comprehensive data on emissions, which are largely unavailable during our study period. Instead, we follow Deryugina et al. (2019) and instrument for changes in SO_2 using changes in wind direction.[4] Our key identifying assumption is that, conditional on other climatic variables and comprehensive fixed effects, wind direction affects mortality only through its effects on air pollution.

We obtain wind speed and wind direction data from a 6-hour reanalysis dataset published by the Japan Meteorological Agency (JMA). The data consist of vector pairs, one for the east-west wind direction (u-component) and one for the north-south wind direction (v-component), reported in 6-hour intervals on a grid with a resolution of 1.25 degrees (≈86 miles).[5] We first interpolate between the grid points to calculate the 6-hour u- and v-

[4] Deryugina et al. (2019) estimate the effects of $PM_{2.5}$ on 3-day mortality and medical spending among the elderly Medicare population. Our empirical analysis, which includes all ages and a longer sample period, focuses on mortality dynamics by age and cause of death. These dynamics help characterize the incidence of short-run pollution exposure and are crucial to forming the long-run projections presented in Section 5.

[5] These data are available starting in 1958 from `http://rda.ucar.edu/datasets/ds628.0/`. The JMA's relatively coarse grid is an appealing feature: as we explain in Section 3.2, regional wind patterns are a better source of variation than local differences in wind direction, which may produce undesirable measurement

components at the centroid of each county. We then average the u- and v-components within a county-day to match the frequency of our mortality data. Finally, we use trigonometry to convert the average u- and v-components into daily wind direction and wind speed.

We obtain temperature and precipitation data from Schlenker and Roberts (2009), who construct a gridded weather dataset at the daily level by combining monthly data from the PRISM Climate Group with daily data from weather stations. The final dataset spans the years 1972–1988 and includes total daily precipitation and daily maximum and daily minimum temperatures for each point on a 2.5-by-2.5 mile grid covering the contiguous US.[6] We aggregate these data to the county-day level by averaging the daily measures across all grid points located in a particular county.

3 Empirical strategy

3.1 Estimating equations

Our first objective is to estimate the causal effect of acute (1-day) exposure to SO_2 on short-run mortality. We model this relationship using the following regression:

$$Y_{cd}^{k} = \beta^{k}\text{SO2}_{cd} + X_{cd}^{k\,\prime}\delta + \alpha_{cm} + \alpha_{my} + \varepsilon_{cd} \tag{1}$$

where Y_{cd}^{k} is the cumulative mortality rate in county c in the k days following exposure on day d (including same-day mortality). The parameter of interest is β^{k}, the coefficient on daily SO_2 levels. Similar to Deryugina et al. (2019), the controls X_{cd}^{k} include contemporaneous and $k-1$ leads of our weather variables (described below) in order to ensure that β^{k} does not capture the effects of weather conditions during the outcome window.

Equation (1) includes fixed effects for county-by-calendar-month (α_{cm}) and calendar-

error in county-level exposure.

[6] See http://www.prism.oregonstate.edu/ for the original PRISM dataset and http://www.wolfram-schlenker.com/ for the daily data.

month-by-year (α_{my}), hereafter referred to as "county-by-month" and "month-by-year". The county-by-month fixed effects control for geographic and geography-specific seasonal differences in mortality, air pollution, and wind patterns. The month-by-year fixed effects control for common time-varying shocks, such as those induced by environmental policy changes during our study period. Standard errors are clustered by county, and the regression is weighted by the relevant county-year population.

Our main specification controls for daily maximum temperature, precipitation, and wind speed. We control for maximum temperature using a set of indicators based on each county's temperature distribution. The literature on temperature and mortality generally finds that it is extreme temperatures that matter for the mortality rate (e.g., Barreca et al., 2016) and that the average climate of a county determines which temperatures are considered extreme (e.g., Heutel, Miller and Molitor, 2021). We therefore generate nine indicators for maximum temperatures falling into intervals whose minimum is defined by the following percentile cutoffs: 0, 1, 5, 10, 25, 75, 90, 95, and 99. We determine the maximum temperature range corresponding to each interval using each county's in-sample temperature distribution.

Guided by the principle of controlling for extremes, we control for daily precipitation by including four indicators for whether precipitation is below the 75th percentile for that county (which in most counties corresponds to very little or no precipitation), between the 75th and 95th percentiles, between the 95th and 99th percentiles, or above the 99th percentile. We control for daily average wind speed with six indicators whose minimum is defined by the following county-specific percentile cutoffs: 0, 25, 75, 90, 95, and 99. Finally, we also control for the set of all possible interactions of these atmospheric controls, yielding 201 different temperature-precipitation-wind-speed combinations.[7]

OLS estimates of Equation (1) are susceptible to bias because SO_2 exposure is not random and is measured with error, as monitor-level fluctuations are a noisy measure of population exposure. We therefore instrument for daily SO_2 using contemporaneous wind direction in

[7]The number of possible combinations is larger than this value, but we do not observe all possible combinations in our data.

the county, allowing the effect of wind direction on SO_2 to vary by geographic group g:

$$\text{SO2}_{cd} = \sum_{g=1}^{50} f^g(\theta_{cd}) + X_{cd}^{k\,\prime}\delta + \alpha_{cm} + \alpha_{my} + \varepsilon_{cd} \tag{2}$$

where:

$$f^g(\theta_{cd}) = \gamma_g^1 1[G_c = g] \times \sin(\theta_{cd}) + \gamma_g^2 1[G_c = g] \times \sin(\theta_{cd}/2)$$

The indicator function $1[G_c = g]$ is equal to 1 if county c is a member of group g and 0 otherwise. The variable θ_{cd} is the local wind direction, measured in radians. The excluded instruments are the 100 regressors formed by the interaction of group indicators, $1[G_c = g]$, with measures of contemporaneous wind direction, $\sin(\theta_{cd})$ and $\sin(\theta_{cd}/2)$. Appendix Section A.2 shows that our results are robust to alternative ways of parameterizing $f^g(\theta_{cd})$.

The wind direction instruments θ_{cd} in Equation (2) vary at the county level. We allow the instruments' effects to vary across different geographic groups—as captured by the parameters γ_g^1 and γ_g^2—because we expect, for example, a westerly wind blowing clean ocean air into California to have a different effect on pollution levels than a westerly wind blowing air pollutants into a community located due east of Chicago. We construct these groups using a k-means clustering algorithm that classifies all SO_2 pollution monitors into 50 spatial groups based on monitor location.[8] To account for autocorrelation in wind direction, which could cause our estimate of β^k to additionally reflect the effects of past or future SO_2 exposure, we include two leads and two lags of wind direction interacted with geographic indicators in the controls X_{cd}^k. Thus, contemporaneous wind direction is an excluded instrument, while past and future wind directions are included instruments.

[8]The inputs into the k-means clustering algorithm are monitor latitude and longitude and the desired number of groups (50). If monitors in the same county are assigned to different groups, we assign the larger integer group number to the county, which is effectively random assignment. The locations of in-sample SO_2 monitors and geographic groups are displayed in Figure A.3. On average, each geographic group encompasses 57 SO_2 monitors and 12 counties.

3.2 Identifying variation

Ambient air pollution comes from both local and distant sources. For example, a city's air pollution levels are influenced by nearby vehicle exhaust as well as by emissions from fires and power plants located hundreds of miles away. The source location matters because local emissions disperse unevenly across the immediate vicinity whereas emissions from a distant source disperse uniformly across that same vicinity. For this reason, using variation from local pollution sources to estimate the county-level effects of pollution exposure can lead to measurement error bias.[9] Thus, our first-stage Equation (2) allows the effect of county-level wind direction on air pollution levels to vary across geographic groups (regions) but not within a group. This restriction avoids using local variation in air pollution produced by sources located inside the geographic group. Local variation will have differing effects on the exposure levels of nearby people, depending on their location relative to the source, resulting in undesirable measurement error. Instead, our first stage favors variation in air pollution that affects the entire geographic group in a similar manner.

Figure 1 illustrates our first-stage variation, using the Greater Philadelphia and Southern California geographic groups as examples. The black dots on the maps show the locations of the SO_2 monitors in these two areas. The plots on the right show the group-specific relationships between daily average wind direction and SO_2.[10] Each specification controls for maximum temperature, precipitation, wind speed, and county-by-month and month-by-year fixed effects in the same way we control for them in our main specification. To show the wind-SO_2 relationship flexibly, we group wind direction into 36 10-degree bins and estimate the first-stage equation separately for each geographic group. The plots also show the relationship based on the sine function specification, which for consistency we also estimate separately for each group. For the main estimates we report later in our paper, we

[9] Consider a county with one air pollution monitor that is placed due west of a major power plant. Suppose that this power plant is located in the center of the county and is the only source of pollution. The monitor will register high pollution readings when the wind blows from the east, and low levels when it blows from the west. However, wind direction in this example has no effect on average pollution levels in the county.

[10] Figure A.4 shows corresponding plots for all 50 geographic groups.

estimate our first and second stages jointly using two-stage least squares.

Figure 1 reveals a strong first-stage relationship between wind direction and SO_2 levels. In the Greater Philadelphia area, pollution levels are highest when the wind blows from the west-southwest direction, and lowest when the wind blows from the east-southeast direction, where the Atlantic Ocean lies. This pattern differs in the Southern California area, where SO_2 levels are highest when the wind blows from the east, a densely populated area, and lowest when the wind blows from the south-southwest direction, where the Pacific Ocean lies. Figure 1 shows that a change in wind direction can change SO_2 levels by 3–4 ppb, equal to 30–40 percent of the national mean during this time period (Table 1).

We interpret our IV estimate as a weighted average of treatment effects among compliers, where weights are larger for compliers with larger first stages (Angrist, Graddy and Imbens, 2000). Figure A.5 shows the geographic distribution of the strength of the first stage, as measured by the difference in predicted SO_2 levels between the most and least polluting wind directions.[11] The areas with the largest variation (4+ ppb) are located primarily in the Midwest and the Northeast, although there are fairly strong compliers (variation of 2–4 ppb) throughout the country. The weakest compliers have variation of <1 ppb and thus contribute relatively little to our main estimates. Appendix Section A.2 presents an analysis of complier characteristics and discusses the monotonicity assumption.

Our empirical approach permits us to instrument for multiple pollutants simultaneously because the relationship between wind direction and pollution levels within and across regions is pollutant-specific.[12] In a later analysis, we investigate the sensitivity of our main estimate to controlling for the four other pollutants measured during our sample time period: NO_2, CO, O_3, and TSP. Including these additional controls causes our sample size to fall by 90 percent or more because these other pollutants are monitored less frequently than SO_2 during this time period. We therefore do not include them in our primary specification.

[11] For each geographic group g, we calculate $\widehat{\gamma_g^1}\sin(\theta) + \widehat{\gamma_g^2}\sin(\theta/2)$ for $\theta \in [0, 2\pi)$ and take the difference between the maximum and minimum values.

[12] For example, manufacturing plants and landfills often emit fine particulate matter, but little to no SO_2.

4 Short-run empirical results

4.1 Mortality by age and cause

We begin by estimating the effect of daily SO_2 exposure on same-day mortality. Table 2 presents OLS and IV estimates of Equation (1). Column (1) reports that a 1-day, 1-ppb increase in SO_2 is associated with a same-day mortality increase of 0.008 deaths per million, about ten times smaller than the corresponding IV estimate of 0.08 deaths per million reported in Column (2). This downward bias for OLS is common in quasi-experimental studies of air pollution, and is often hypothesized to be at least partly due to measurement error in pollution exposure (Deryugina et al., 2019; Alexander and Schwandt, 2022). Our first-stage F-statistic exceeds 500, indicating a strong instrument.[13]

Our key identifying assumption is that changes in wind direction are unrelated to mortality except through their effects on pollution levels. This assumption would be violated if wind direction is correlated with unobserved weather patterns that cause mortality. While impossible to test directly, the plausibility of our identifying assumption can be evaluated by considering the sensitivity of our estimates to different ways of controlling for weather conditions, which we do in Columns (3)–(5) of Table 2. Column (3) reports the estimate from a specification that additionally controls for bins of minimum temperature—with thresholds defined as for maximum temperature—in our weather interaction indicators. In Column (4), we change the parameterization of the weather controls, controlling for all possible interactions of minimum temperatures (specified as indicators for minimum temperatures falling into 3-degree Celsius bins, with outer bins defined by temperatures below -15 Celsius or above 30 Celsius); maximum temperatures (specified in the same way); ten deciles of wind speed; and ten deciles of precipitation. This flexible definition of weather gives us 28,899 possible weather conditions ($17 \times 17 \times 10 \times 10 - 1$), although only about one-third of these

[13]The F-statistic is computed assuming errors are homoskedastic, which means it can be compared to the well-known critical values published in Stock and Yogo (2005). Heteroskedasticity-robust F-statistics also lie well above conventional thresholds. The weak-instrument test of Olea and Pflueger (2013) is not computationally feasible in our setting.

combinations exists in our data. Finally, Column (5) reports estimates from a specification with no weather controls at all. Our estimate remains stable across all specifications.

Figure 2 shows the IV estimates of the effects of a 1-day, 1-ppb increase in SO_2 levels on mortality up to one month following exposure. The first blue point depicts the estimate from Column (2) of Table 2. If short-term mortality displacement were the predominant driver of this 1-day mortality effect, the cumulative mortality effect would decline over time, potentially all the way to zero. Instead, the estimated effect increases steadily to 0.10 deaths per million one week after exposure and to 0.18 deaths per million one month after exposure. Figure A.6 presents cumulative mortality estimates up to 90 days following exposure. We can strongly reject an estimate of 0 throughout this 90-day follow-up period. While the estimated effect stabilizes after about 35 days, the standard errors grow steadily with the length of the outcome window, limiting our ability to obtain precise estimates beyond a one-month window.

Figure 3 shows how our estimates vary by cause of death as reported on death certificates. The increase in 1-day mortality is split roughly equally among cardiovascular disease, cancer, and other diseases. As we consider longer time horizons, however, the estimated effect on cancer-related mortality falls, implying substantial short-run mortality displacement. One month after exposure, the cancer estimate is small and statistically insignificant. We therefore cannot reject the null hypothesis that all cancer-related deaths would have occurred within one month even absent exposure to SO_2. By contrast, the death rate estimates for cardiovascular and other diseases increase with the time horizon—more than tripling over one month—implying that acute SO_2 exposure continues to have lethal effects even after exposure has ended. For each of these cause-of-death categories, we can strongly reject equality between the 1-day and the 28-day estimates ($p<0.0001$).[14]

[14]Figure 3 also shows a small but statistically significant increase in same-day external deaths (0.005 deaths per million), although this estimate becomes statistically insignificant when measured over longer time horizons. These effects may be due to negative effects of pollution on cognitive function (Crüts et al., 2008; Fonken et al., 2011; Bishop, Ketcham and Kuminoff, 2023). For example, recent work has suggested that pollution exposure increases vehicle fatalities through the cognitive function channel (Burton and Roach, 2023).

An alternative explanation for this disparity in trends by cause of death is composition bias, which could arise if cancer-related deaths occurring weeks after exposure are misattributed to other causes like cardiovascular disease. To investigate this possibility, we estimate the effect of acute exposure on deaths where cancer was listed as the underlying cause of death *or* as a secondary (contributing) cause of death. Those estimates, shown in Figure A.7, are very similar to our main estimates, indicating that misattribution in the underlying cause of death does not explain our finding of short-run mortality displacement.

Figure A.8 shows cause-specific estimates of the mortality effect of SO_2 over time when deaths from cardiovascular and other diseases are disaggregated into 5 and 21 different subcategories, respectively. Three of the five subcategories of cardiovascular disease (heart disease, cerebrovascular disease, and atherosclerosis) show significant same-day and monthly mortality increases. We also find strong and growing effects for deaths from chronic obstructive pulmonary disease (COPD), pneumonia and influenza, and other respiratory diseases, which collectively account for over 99 percent of all deaths from respiratory illness (Table A.13). By contrast, we find no significant monthly effects for conditions not previously linked to air pollution exposure, such as stomach ulcers, Parkinson's disease, appendicitis, and chronic liver diseases.[15]

Figure 4 shows how the 1-day mortality effect varies by age group. Panel (a) shows the absolute magnitude of the effect (in deaths per million), while Panel (b) reports it as a percent of the average 1-day mortality for that age group. We fail to detect significant mortality increases for the two youngest age groups (covering ages 0–19). For older ages, our estimates are statistically significant and range from 0.016 deaths per million for 20–44-year-olds to 2.3 deaths per million for 85+ year-olds (Figure 4a). When expressed in relative terms, estimates vary little with age. The smallest relative effect (0.18 percent of daily mortality) is found among 1–19 year-olds, while the largest relative effect (0.51 percent of

[15]Interestingly, we estimate a significant effect of SO_2 on deaths from meningitis. Although not obviously pollution-related, prior studies have found that air pollution weakens local immunity of the pharynx, leading to increased susceptibility to meningitis (Michele et al., 2006; Jusot et al., 2017; Shears et al., 2020).

daily mortality) is found among 85+ year-olds. However, we cannot reject that the relative effects are equal to each other for most age group pairs.

Age-specific estimates for outcome windows up to one month (28 days) following exposure are reported in Figure A.9 and Table A.4. For ages 65 and over, the monthly estimates are at least twice as large as the 1-day estimates, suggesting that any short-run mortality displacement among this group is more than offset by delayed effects of acute exposure. For ages 20 to 44, however, the monthly estimates are smaller than the 1-day estimates and are statistically insignificant, indicating substantial mortality displacement. Thus, among young and middle-aged adults, acute pollution exposure increases mortality primarily for frail individuals who would have died within one month even absent exposure. By contrast, older adults who are killed by acute air pollution exposure have longer counterfactual lifespans.

In relative terms, our estimates for older age groups are similar to corresponding estimates reported in Deryugina et al. (2019), who investigate the effect of acute exposure to $PM_{2.5}$ on 3-day elderly mortality. To the best of our knowledge, no prior quasi-experimental estimates exist for the effect of air pollution on short-run mortality for ages 1–64. For infant mortality, the two most comparable studies are Currie and Neidell (2005) and Knittel, Miller and Sanders (2016), who estimate the effect of PM_{10} (coarse particulate matter with diameter less than 10 μm) on weekly infant mortality. Currie and Neidell (2005) find null effects, while Knittel, Miller and Sanders (2016) estimate significant effects. Our estimates are weakly consistent with the mixed findings from these two studies: while we find a null (albeit noisy) effect of acute exposure on 1-day and monthly infant mortality, our estimates for 3-day and weekly infant mortality are statistically significant (Table A.4).

Appendix Section A.2 presents a series of robustness checks that largely follow Deryugina et al. (2019). These include using different sets of controls, clustering standard errors at different levels, estimating LIML rather than 2SLS, and performing falsification and placebo tests. The results support the validity of our main estimates.

4.2 Other air pollutants

We have thus far interpreted our mortality estimates as the causal effect of exposure to SO_2. It is possible, however, that other harmful air pollutants are co-transported with SO_2. A second, related issue is that SO_2 transforms rapidly into sulfate (SO_4^{2-}), a major component of $PM_{2.5}$. Because the transformation occurs at a rate of several percent per hour, our estimates may reflect the mortality effects of exposure to SO_2 as well as sulfate recently derived from SO_2. We investigate these possibilities using two different methods. The first method controls directly for other air pollutants, instrumenting separately for each one. The second models the atmospheric transport of SO_2, NO_2, and $PM_{2.5}$, as well as the conversion of SO_2 and NO_2 into $PM_{2.5}$, so that we can gauge the potential contributions of these different phenomena to our estimates.

The air pollutants we observe in our data are produced by a variety of sources in different geographic locations, are carried differently by the wind, and exhibit different atmospheric chemistry patterns. This variation can be captured by our instruments, which vary by location and with wind direction. Unfortunately, including multiple pollutants in our regression significantly reduces our sample size because most non-SO_2 pollutants are sparsely measured during our sample period. We therefore include multiple pollutants in our empirical model only in this secondary analysis, focusing on two subsamples. The first one includes measures of TSP, NO_2, O_3, and CO in addition to SO_2, resulting in a sample size that is less than five percent of the size of our main sample. The second subsample is larger but includes only SO_2 and TSP, our best proxy for $PM_{2.5}$.

Panels A and B of Table 3 show the effect of adding pollutants to our main regression specification in each subsample. For reference, Column (1) in Panel A reports that a 1-ppb increase in SO_2 raises 1-day mortality by 0.084 deaths per million in the first, smallest subsample. The magnitude of this effect decreases by nearly 30 percent in Column (2), which additionally controls for TSP, but does not change in Column (3), which controls for NO_2, CO, and O_3. Column (4) controls for all these pollutants simultaneously and again

shows a meaningful decrease in the magnitude of the coefficient on SO_2, suggesting that some of the estimated effect for SO_2 reported in Column (1) is driven by particulate matter. For completeness, Table A.6 shows estimates for a third subsample that is restricted to observations that have pollution readings for SO_2, NO_2, O_3, and CO (but not necessarily TSP), which yields almost four times the number of observations as Panel A of Table 3. The SO_2 coefficient remains stable across all possible combinations of these other pollutants, further confirming that only the inclusion of TSP affects its magnitude.

Panel B shows the effect of controlling for TSP in a larger subsample that conditions only on observing SO_2 and TSP. Here, the coefficient on SO_2 falls by 55 percent after controlling for TSP. Likewise, the coefficient on TSP falls by 30 percent after controlling for SO_2.[16] Overall, we conclude that about half of the SO_2 mortality effect we estimate is potentially driven by particulate matter that was either co-transported with or formed from SO_2.

Further interpreting the estimates in Table 3 is difficult because it is not possible to distinguish "primary" TSP that is directly emitted by a source and then co-transported with SO_2 from "secondary" TSP that is formed from SO_2. This distinction is policy relevant: reducing SO_2 emissions necessarily reduces secondary TSP, but may have little or no effect on primary TSP. We therefore turn to simulations from the Intervention Model for Air Pollution (InMAP), which models how emissions of SO_2, $PM_{2.5}$ (a component of TSP), and NO_2 are transported across the US.[17] Importantly, InMAP simulates the conversion of SO_2 and NO_2 into particulate matter, allowing us to separately measure primary $PM_{2.5}$ and secondary $PM_{2.5}$ and to quantify the amount of secondary $PM_{2.5}$ formed specifically from SO_2. We use the 1990 National Emissions Inventory (NEI) to determine emissions, the earliest year for which appropriate data are available.[18] We report two sets of results, one based on

[16] We observe a similar pattern for longer outcome windows: the coefficient estimates for SO_2 and TSP fall by 30–50 percent when both pollutants are included in the regression, but always remain statistically significant.

[17] The InMAP model is available for download from https://github.com/spatialmodel/inmap/releases/tag/v1.9.6. The evaluation data used in our simulations are from Tessum et al. (2019).

[18] These data are available from https://gaftp.epa.gov/air/nei/nei_criteria_summaries/1990criteriasummaryfiles/.

emissions from coal-fired power plants, the largest source of SO_2 emissions in the NEI, and a second where we additionally include emissions from all sources located in a county with at least one coal-fired power plant. The second scenario thus includes air pollutants that are co-transported, but not necessarily co-emitted, with SO_2.

We use the InMAP simulations to calculate the equilibrium ratio of ambient $PM_{2.5}$ to ambient SO_2, in $\mu g/m^3$ per ppb.[19] That ratio is approximately 2.3 when considering only coal-fired power plant emissions, indicating that each transported ppb of SO_2 is accompanied by 2.3 $\mu g/m^3$ of $PM_{2.5}$. Over 93 percent of this $PM_{2.5}$ is sulfate, a secondary pollutant derived from SO_2. The second scenario, which includes emissions from all pollution sources located in counties with coal-fired power plants, yields a similar ratio of 2.5, with over 70 percent of $PM_{2.5}$ consisting of sulfate. Thus, even if our mortality estimates reflect the effects of exposure to particulate matter, these simulations suggest that the large majority of that particulate matter originates from SO_2 and would therefore fall following reductions in SO_2 emissions. We can also use these results to compute bounds for our estimates. Under the extreme assumption that our mortality estimates are entirely caused by fine particulate matter rather than SO_2, dividing them by 2.5 would yield the mortality effect of one additional $\mu g/m^3$ of $PM_{2.5}$.

5 Long-run survival

5.1 Framework

Our framework for quantifying the long-run survival effects of chronic exposure to air pollution is based on a dynamic production model of health originally developed by Lleras-Muney and Moreau (2022). This model has several appealing features. Its fit is on par with the best models used by demographers across a wide range of population survival curves

[19] InMAP reports all pollutants in units of $\mu g/m^3$. At standard temperature (15 degrees Celsius), 1 ppb of SO_2 corresponds to about 2.62 $\mu g/m^3$.

(Lleras-Muney and Moreau, 2022). The model can also separately capture the mortality displacement and accelerated aging effects we observe in our empirical analysis, which have very different implications for long-run survival. Finally, calibrating the model requires data on mortality only, unlike other models such as Grossman (1972), which require additional information on variables such as incomes, prices, and healthcare utilization, and on how these change with air pollution exposure.

Let H_{it} denote the health capital of individual $i \in \{1, ..., N\}$ at time $t \in \{0, ..., T\}$. At birth, an individual is endowed with a stock of initial health, H^*_{i0}, which is drawn randomly from a normal distribution. This health stock evolves over the individual's lifetime according to the following formula:

$$H_{it} = H_{i,t-1} - d(t) + I + \varepsilon_{it} \tag{3}$$

where:

$$\begin{aligned} H_{i0} &= H^*_{i0} \sim N(\mu_H, \sigma_H), \\ d(t) &= \delta t^{\alpha}, \\ \varepsilon_{it} &\sim N(0, \sigma_{\varepsilon}) \end{aligned}$$

The health stock depreciates at a rate, $d(t)$, which increases with the age of the individual. It is replenished at a constant rate, I, which captures time-invariant factors such as early-life parental investment or lifetime health habits, and varies with an iid health shock, ε_{it}. Death occurs when the individual's health stock falls below a critical threshold, $\underline{\mathrm{H}}$, and is denoted by the indicator variable D_{it}, where:

$$\begin{aligned} D_{i0} &= 1\left[H_{i0} < \underline{\mathrm{H}}\right], \\ D_{it} &= 1\left[H_{it} < \underline{\mathrm{H}} \,\middle|\, D_{i,t-1} = 0\right], \ t > 0 \end{aligned}$$

The model is fully characterized by seven parameters: $\{\alpha, \delta, I, \mu_H, \sigma_H, \sigma_e, \underline{\mathrm{H}}\}$.[20]

Cohort mortality can be calculated to an arbitrary degree of precision by simulating the model given by Equation (3) for a sufficiently large number of individuals. The mortality rate at time t, M_t, is equal to the number of individuals who died in period t divided by the number of individuals alive at the beginning of period t. Survival at time t can then be calculated from mortality:

$$S_1 = 1 - M_0,$$
$$S_t = S_{t-1}\left(1 - M_{t-1}\right),\ t > 1$$

The health effects of pollution exposure can be modeled as temporary or permanent changes to one or more model parameters. We focus on changes to the death threshold, $\underline{\mathrm{H}}$, which produces mortality displacement, and the depreciation parameters, α and δ, which govern the aging process.[21] Altering these parameters has very different effects on long-run outcomes. Temporarily raising the death threshold leads to increased mortality among frail individuals close to death but has no effect on the health of survivors. When the threshold reverts to its original value, mortality rates decline because there are few surviving individuals with health capital values near this lower threshold in the short run. Consequently, a temporary rise in the death threshold leads to short-run mortality displacement without any longer-lasting effects.

In contrast, a temporary increase in the depreciation rate, $d(t)$, affects the health capital of all individuals—both healthy and ill—leading to persistent changes in mortality. Like

[20] The model can be extended to incorporate external causes of death such as car accidents by including two additional parameters specifying the age of onset and the severity of these external causes (Lleras-Muney and Moreau, 2022). However, this extension is unnecessary for our analysis, which focuses on deaths from biological (internal) causes.

[21] While exposure could in principle change the level of investment, I, this effect would result in high and persistent mortality increases for both old and young. Because the mortality effects we observe for younger individuals are small and disappear within one month (Table A.4), we rule out I as a possibility. Exposure could also alter the variance of the iid health shock, ε_{it}. We rule out this possibility because it implausibly implies that a 1-day increase in pollution exposure would produce health benefits for half of the population.

an increase in the death threshold, this increase in depreciation triggers an immediate rise in mortality among frail individuals. Additionally, because the health capital of the entire population has been reduced, and because future health capital depends on past health capital, mortality rates will remain elevated even after $d(t)$ reverts to its original trajectory.

To incorporate our empirical estimates into the model, we assume that the effect of pollution exposure on model parameters depends only on current exposure. This assumption implies that exposure alters the death threshold and depreciation rate parameters by the same amount for the old and the young, irrespective of their prior exposure history. However, because depreciation is a power function of age ($d(t) = \delta t^{\alpha}$), an adverse change to the depreciation parameters will reduce health capital more for the old than the young. Moreover, older individuals generally have lower pre-existing health capital, which means exposure will increase mortality more at older ages. This result is consistent with the age-specific mortality patterns from our empirical analysis (Figure 4).

This modeling assumption is supported by medical research, which has found that the underlying biological effects of pollution exposure are similar across different segments of the population. For example, ambient air pollution exposure produces a number of latent pathophysiological responses, such as increased oxidative stress and inflammation, in both healthy and unhealthy individuals (Brook et al., 2010). Likewise, randomized trials find that higher $PM_{2.5}$ exposure leads to increased levels of stress hormones, insulin resistance, inflammatory markers, and blood pressure even among healthy young adults (Li et al., 2017).

We provide two additional pieces of empirical support for our assumption that the effect of exposure on model parameters depends only on current exposure. First, the assumption yields the testable implication that calibrated parameter values from one age group can be used to form accurate mortality predictions for any other age group, regardless of differences in their exposure histories. We confirm this implication in the "leave-one-out" exercise presented in Section 5.3. Second, Section 5.4 shows that the model's chronic exposure projections are consistent with one-month exposure estimates from our empirical model and

with a three-year estimate from Anderson (2020).

Our analysis proceeds as follows. First, we calibrate the baseline model parameters using a period life table from 1972, the beginning of our sample period. Second, we downscale this annual model to the daily level and solve for the change in parameter values required to match our IV estimates of the same-day effect of acute pollution exposure. We assess model performance by comparing its short-run (up to one month) predictions to IV estimates not used for the calibration. Finally, we use the model to project the long-run effects of chronic changes in pollution exposure on survival. We explain these steps in more detail below.

5.2 Calibration

5.2.1 Baseline parameters

The dynamic production model of health given by Equation (3) depends on seven parameters: $\{\alpha, \delta, I, \sigma_e, \mu_H, \sigma_H, \underline{\mathrm{H}}\}$. To achieve identification (i.e., to ensure a unique solution), we follow Lleras-Muney and Moreau (2022) and normalize two parameters: $\underline{\mathrm{H}} = 0$ and $\sigma_H = 1$. We calibrate the five remaining parameters using simulated method of moments. Specifically, we use the Nelder-Meader method to solve numerically for the parameter values that minimize the squared distance between the model's predicted age-specific survival and US population survival in 1972, the first year of our sample period.

Lleras-Muney and Moreau (2022) model mortality at the annual level. To incorporate our IV estimates with maximum accuracy, however, we must downscale the model to the daily level. Our daily-level IV estimates average about 0.1 deaths per million for a 1-unit change in SO_2 (Table 2). Thus, the model must include many millions of individuals to accommodate that level of granularity. However, calibrating the baseline parameters requires solving an optimization problem over the entire lifetime, which is computationally infeasible at the daily level with millions of individuals.[22] We therefore proceed in three steps.

[22] If the maximum lifespan is 110 years, then the number of periods is $T = 110 \times 365 = 40,150$ days. Simulating a population of $N = 10,000,000$ individuals thus produces 401.5 billion health capital values (3.2 terabytes, if each value is an 8-byte number). Simulating a single population of this size on a large server

First, we calibrate the baseline model using annual survival data and $N = 1,000,000$ agents. Second, we recalibrate the baseline model using daily survival data and $N = 100,000$ agents, employing (scaled) estimates from the annual calibration as our starting values.[23] Third, we simulate the baseline model for a population of $N = 20,000,000$ individuals, using parameter estimates from the second (daily-level) calibration. This third step provides the high-resolution model estimates of population health capital that we combine with our IV estimates to calibrate the effect of exposure on model parameters. Because the pollution exposure calibration focuses on a single point in time instead of tens of thousands of points over a lifetime, it can be performed even with millions of individuals.

Figure A.10 illustrates the results of our baseline calibration. The solid blue line shows the survival curve for the US 1972 life table. The dashed red line reports the survival curve produced by our calibrated model. There is a small difference between the model's prediction and the observed data in infancy, but otherwise the two curves track each other closely and produce life expectancy estimates that differ by only 0.1 years, indicating a successful model fit. The model's parameter values are reported in Table A.14.

5.2.2 Pollution exposure parameters

Our IV estimates from Section 4 identify the effect of a 1-day, 1-ppb increase in SO_2 exposure on mortality over the following month. To incorporate these estimates into our model, we assume that exposure can raise the value of the death threshold parameter, $\underline{H}$, and can increase depreciation by raising the value of δ or α. A 1-day increase in the death threshold will produce short-run mortality displacement that increases current mortality and reduces near-future mortality by about the same amount, while an increase in depreciation will

requires several hours of runtime, and the Nelder-Mead method requires simulating the population hundreds of times to converge to a solution.

[23] We obtain the starting values by dividing the annual estimates of I and δ by 365, and of σ_e by $\sqrt{365}$. We do not rescale μ_H or α. While these starting values provide a good guess for the solution, the model's (nonlinear) dependence on t causes the guess to be suboptimal, hence the need for recalibration. In the special case where $\alpha = 1$, one can show that if a year is partitioned into P periods, then the optimal value for δ in the partitioned model approaches $1/P$ times the optimal value from the annual model.

increase both current and future mortality.

In line with the patterns in Figure 3, we assume that cancer-related deaths reflect mortality displacement while other deaths reflect accelerated aging (depreciation). Let $\hat{\beta}^k_{a,c}$ denote the IV estimate of the effect of acute exposure on cumulative mortality for age group a from cause of death c in the k days following exposure. Consider a specific daily age, t, that lies inside the age interval spanned by age group a. We first use the high-resolution model estimates described above to solve numerically for a new death threshold, $\widetilde{\underline{\mathrm{H}}}$, such that elevating the baseline threshold to $\widetilde{\underline{\mathrm{H}}}$ for a single day at age t yields a mortality increase equal to the 1-day estimate for cancer-related deaths, $\hat{\beta}^1_{a,cancer}$. We then solve for the value of $\widetilde{\delta}$ or $\widetilde{\alpha}$ that further increases mortality at age t to match the all-cause (total) estimate, $\hat{\beta}^1_{a,all}$.

Because the effect of pollution exposure on underlying model parameters does not depend on age, we can calibrate those parameters using IV estimates for any age group a. To increase precision, we perform the calibration using several age groups, denoting the results for age group a as $\{\widetilde{\underline{\mathrm{H}}}_a, \widetilde{\delta}_a\}$ or $\{\widetilde{\underline{\mathrm{H}}}_a, \widetilde{\alpha}_a\}$, and then use the simple average across ages when projecting long-run survival. Because we need death rate magnitudes to be large enough to match the granularity of our model, we exclude below-65 age groups.[24] The specific IV estimates used for these calibrations are reported in Table A.15.

Using only a single day to calibrate the parameters produces a noisy solution because health capital in Equation (3) also varies with an iid health shock. To improve precision, we solve for the 1-day changes using 50 different days around the approximate midpoint of each age bin and take the average. For example, for the 65–69 age group, we solve for the changes using ages 68y1d, 68y2d, ..., 68y50d.[25] Appendix Section A.3 provides additional details. The final result is a set of parameters for each age group, $\{\widetilde{\underline{\mathrm{H}}}_a, \widetilde{\delta}_a\}$ or $\{\widetilde{\underline{\mathrm{H}}}_a, \widetilde{\alpha}_a\}$, capturing the effect of an increase in SO_2 exposure of 1 ppb. The parameter $\widetilde{\underline{\mathrm{H}}}_a$ captures the mortality

[24] Our model includes $N = 20$ million individuals, so a daily death rate of 1 per million corresponds to at most 20 deaths. The IV estimates for age groups under 65 are all below 0.1 deaths per million, which is too small for reliable calibration when $N = 20$ million.

[25] The optimal strategy would employ all $365 \times 5 = 1825$ days in the 5-year age bin, giving more weight to the ages near the midpoint. However, doing so is computationally burdensome and unnecessary for achieving sufficient accuracy.

displacement effect, and the parameter $\widetilde{\delta}_a$ or $\widetilde{\alpha}_a$ captures the accelerating aging effect.

To account for econometric uncertainty in the IV estimates, we use a resampling-based methodology. We randomly draw an estimate of the effect of acute pollution exposure on 1-day mortality from a normal distribution with a mean and standard deviation set equal to the mean and standard error of $\widehat{\beta}^1_{a,c}$ in Equation (1), and then calibrate the change in model parameters to match the mortality change draw. We repeat this exercise 100 times and report the 5th and 95th percentiles of the resulting distribution of model parameter estimates.

Our age-specific estimates of the effect of exposure on changes in the mortality threshold and the depreciation parameters are shown in Figure A.11. Estimates are similar across ages, with most values falling inside the 90% confidence intervals of other estimates, which is consistent with our assumption that the effects of exposure on model parameters are constant over the life cycle.

5.3 Validation

We perform two sets of validation exercises. The first compares out-of-sample mortality projections from our model to our IV estimates. The second compares three-year projections to estimates from Anderson (2020). We calibrate only the parameters $\underline{\mathrm{H}}$ (mortality displacement) and δ (accelerated aging) in our validation exercises. As we show later, a model calibrated using α instead of δ produces similar long-run projections. Over shorter time horizons, the projections are nearly identical.[26]

5.3.1 One-month projections

Because calibration is based solely on 1-day IV estimates, our first validation test focuses on model predictions for longer-run (>1 day) outcomes. Figure 5 shows the results of this

[26]The marginal effects of δ and α on depreciation, $d(t) = \delta t^{\alpha}$, vary over time (and thus with age), so in theory one could identify which parameter is a better fit for pollution exposure by using variation in mortality by age group. In our setting, however, the differences are small and statistically indistinguishable from each other even when measured over a lifetime.

exercise for the 65–69 age group. The solid blue line reports the IV estimates, and the green dot-dashed line reports "own-age" predictions, which come from a model that was calibrated using $\hat{\beta}^1_{65,cancer}$ and $\hat{\beta}^1_{65,all}$. By construction, the own-age model prediction matches the IV estimate for the same-day effect. To assess model performance, we compare the own-age predictions of longer-run mortality effects to the corresponding IV estimates. All of these predictions lie within the 95% confidence intervals of the IV estimates.

Next, we compute mortality predictions for a given age group using the average of the calibrated parameters from other age groups. For example, instead of using the calibrated parameters $\{\widetilde{\underline{\mathrm{H}}}_{65}, \widetilde{\delta}_{65}\}$ as we did for the own-age predictions, we predict mortality for the 65–69 age group using the average of the calibrated values from the 70–74, 75–79, 80–84, and 85+ age groups: $\{\frac{1}{4}\sum_{a>65}\widetilde{\underline{\mathrm{H}}}_a, \frac{1}{4}\sum_{a>65}\widetilde{\delta}_a\}$. Because these "leave-one-out" predictions are not based on the IV estimates for the 65–69 age group, the 1-day prediction now also serves as a validation test. The thick red dashed line in Figure 5 shows that all of these leave-one-out predictions lie inside the 95% confidence intervals of the IV estimates. A notable aspect of this validation exercise is its ability to produce accurate predictions despite relying on IV estimates that are significantly larger—by nearly an order of magnitude—than those for the 65–69 age group (Figure 4a). As explained in Section 5.1, a fixed change in the depreciation parameter will produce larger mortality effects in older populations. The successful fit of this leave-one-out prediction, which was calibrated using estimates from much older age groups, indicates that the power function specification for depreciation is a reasonable assumption in our setting.

To illustrate that the consistency between our model predictions and IV estimates is a meaningful result and not merely a consequence of large standard errors, we also plot predictions from a model specification that assumes none of the 1-day exposure deaths are due to changes in the death threshold ("no displacement") and one that assumes all of the deaths are due to changes in the death threshold ("all displacement"). In the no-displacement scenario, depicted by the orange dashed line at the top of Figure 5, the mortality predictions

increase much more rapidly than either the own-age or leave-one-out predictions. In the all-displacement scenario, depicted by the black dashed line at the bottom of the figure, the cumulative mortality effect falls rapidly to zero, reflecting complete short-run mortality displacement. The predictions from these two extreme specifications lie well outside the 95% confidence intervals of our IV estimates, demonstrating that our IV estimates are precise enough to rule out a meaningful range of model predictions. The wide gap between the all-displacement and no-displacement predictions underscores the importance of accurately estimating the fraction of deaths that are attributable to mortality displacement rather than accelerated aging.

Figure A.12 shows leave-one-out predictions for all five age groups 65 and over. The model prediction of the 28-day mortality effect lies inside the 95% confidence interval of the corresponding IV estimate for each age group. The close concordance between these predictions and the IV estimates suggests that our dynamic production model of health accurately captures important determinants of mortality, at least for a one-month window following 1-day exposure. The close concordance also shows that one does not need to account for differences in chronic exposure histories when predicting mortality across different age groups, which supports our assumption that the effect of exposure on model parameters depends only on current exposure.

Another way to validate our chronic exposure assumption is to directly compare model predictions of the effects of longer-run (>1 day) exposure to empirical estimates. Although we lack exogenous variation in long-term chronic exposure, we have sufficient statistical power to aggregate our daily variation in SO_2 exposure to the monthly level. First, we isolate exogenous variation in daily SO_2 by using estimates from our first-stage regression, given by Equation (2), to form the prediction $\widehat{\text{SO2}}_{cd} = \sum_{g=1}^{50} \widehat{f}^g(\theta_{cd})$. We then aggregate our sample up to time periods denoted by $\bar{d}$, where $\bar{d}$ has a span ranging from 2 to 28 days, and calculate $\text{SO2}_{c\bar{d}}$ and $\widehat{\text{SO2}}_{c\bar{d}}$, the average of SO2_{cd} and $\widehat{\text{SO2}}_{cd}$ over the time period $\bar{d}$. Finally, we regress $Y_{c\bar{d}}$, the mortality rate in time period $\bar{d}$, on $\text{SO2}_{c\bar{d}}$ and instrument for

it using $\widehat{\text{SO2}}_{c\bar{d}}$, controlling for county-by-month, month-by-year, and time period ($\bar{d}$) fixed effects.[27] We then compare these short-run chronic exposure estimates to the corresponding predictions from our calibrated model.

Figure 6 reports results for the 65–69 age group. Total mortality rises gradually with the length of exposure, reaching about 5 deaths per million after two weeks. This effect is about ten times larger than the two-week effect of acute exposure (Figure 5). In the subsequent two weeks, the effect increases further, reaching about 15 deaths per million at 28 days, almost twenty times larger than the corresponding acute effect. Although we can strongly reject the hypothesis that the mortality effect is constant over the one-month window, estimates for longer windows are noisy: the upper bound of the 95% confidence interval for the 28-day estimate is about 5 times larger than the lower bound. Nevertheless, it is reassuring that our model projections follow a similar trajectory and fall within the 95% confidence intervals of nearly all the IV estimates.

5.3.2 Three-year projections

While our model's one-month projections perform well for both acute and chronic exposure, that does not guarantee accuracy over a lifetime horizon. Regrettably, there are no quasi-experimental estimates available for lifetime exposure; indeed, this gap in knowledge is precisely what motivates our study. As a second-best option, we turn to quasi-experimental estimates from Anderson (2020), who studies the effect of living downwind of Los Angeles highways on three-year mortality among individuals 75 or older. His estimates imply that a 10 percent increase in NO_2 levels reduces life expectancy by 0.05–0.064 years.[28]

To compare these estimates to our model, we construct a counterfactual for our 1972

[27] Because the time period $\bar{d}$ will sometimes span more than one calendar month, the month-by-year fixed effects are not perfectly collinear with the time period fixed effects. In cases where the time period spans more than one month, we use the fixed effect corresponding to the minimum month.

[28] The key independent variable in Anderson (2020) is the percent of time that a census block group located within 600 meters of a highway was downwind of the highway. To help interpret his reduced-form estimates, he separately estimates a first stage using data on NO_2 levels provided by four pollution monitors located near highways.

cohort where we increase SO_2 levels by 1 unit ($\approx$ 10 percent) beginning at age 72 for 10 years—the approximate time that individuals in the Anderson (2020) sample were exposed to higher levels of pollution from living downwind as opposed to upwind of a highway.[29] We then calculate the ensuing three-year mortality increase for ages 82–84 and express it as a change in life expectancy at birth to make it comparable to Anderson (2020). The result is a life expectancy reduction of 0.09 years. The associated 90% bootstrap confidence interval ranges from 0.03 to 0.16 years, which overlaps with Anderson's estimates of 0.05–0.064 years. An appealing feature of our model is that we can also use it to project mortality effects beyond Anderson's three-year window. Doing so implies that these 10 years of chronic exposure reduce life expectancy at birth by 0.17 years (90% bootstrap CI: 0.06 to 0.30), suggesting that even a three-year outcome window is insufficient for accurately estimating the full mortality consequences of pollution exposure.

5.4 Long-run projections

Finally, we use our calibrated model to quantify the effects on life expectancy of a permanent, 1-ppb ($\approx$ 10%) decrease in SO_2 exposure. We compute counterfactual survival for the 1972 cohort using the average of age-specific calibrated parameter values, $\{\frac{1}{5}\sum_{a\geq 65}\widetilde{\underline{\mathrm{H}}}_a, \frac{1}{5}\sum_{a\geq 65}\widetilde{\delta}_a\}$ or $\{\frac{1}{5}\sum_{a\geq 65}\widetilde{\underline{\mathrm{H}}}_a, \frac{1}{5}\sum_{a\geq 65}\widetilde{\alpha}_a\}$, and assume that the change in exposure begins at birth, i.e., that exposure alters the model parameters for all $t \geq 0$. For purposes of comparison, we also quantify the implied life expectancy effect of extrapolating our age-specific monthly IV estimates to the entire life cycle.[30]

Figure 7 illustrates our results. The calibrated baseline life expectancy of the 1972 cohort is 71.32 years. The two model-based projections indicate that a permanent, 1-ppb decrease

[29] Anderson (2020) estimates that 78 percent of the in-sample individuals have lived in the same location for at least 10 years. The average age of the US population over 75 is 82 in the year 2000, the center of the time period studied in Anderson (2020).

[30] These IV estimates are reported in the last row of Table A.3. The extrapolation assumes that daily mortality rates increase at every age by the monthly IV point estimate for the corresponding age group, including point estimates that are statistically insignificant. We assume the estimates in Table A.4 apply to the midpoint of each age bin and interpolate to calculate the rest of the values.

in SO_2 exposure improves life expectancy for this cohort by 1.18–1.32 years (1.65–1.85%). By contrast, extrapolating our short-run (monthly) IV estimates to the life cycle yields a life expectancy improvement of about 0.17 years (0.24%), which is 7–8 times smaller than the model-based estimates and lies outside the model's 90 percent confidence intervals.[31] The IV extrapolation is biased downwards because it is limited by the one-month regression window and therefore cannot account for the latent effects of pollution on people's health capital, which may not increase mortality until years or even decades later.

Although decreased air pollution exposure begins at birth, Figure 7 indicates that survival gains are concentrated among older individuals: over 90 percent of the life expectancy improvement occurs after age 50 and over 75 percent occurs after the age of 65. This pattern arises for two reasons. First, reduced exposure has a smaller effect on the health capital of the young than the old because health capital depreciates as a power function of age. Second, even though younger individuals do accumulate some health capital as a result of reduced exposure, it has little immediate effect on their observed mortality because the health capital levels of most young individuals far exceed the death threshold. However, as they age and their health capital falls towards the threshold, the additional health capital accrued at younger ages is reflected in their improved survival rates.

5.4.1 Discussion

Some of the SO_2 exposure we measure may be accompanied by unobserved exposure to $PM_{2.5}$. If we make the extreme assumption that our IV estimates are driven entirely by $PM_{2.5}$ and that each 1 ppb of SO_2 is accompanied by 2.5 $\mu g/m^3$ of $PM_{2.5}$ (see Section 4.2), then we would conclude that a permanent 1-$\mu g/m^3$ decrease in $PM_{2.5}$ improves life expectancy by 0.47–0.53 years. Even under this extreme assumption, our estimates remain policy relevant

[31] These estimates and their confidence intervals are reported in the first row of Table A.16. The rest of the table reports results for chronic exposure decreases and increases of up to 3 units. The relationship between survival improvements and permanent changes in air pollution is roughly linear, although we caution against generalizing from these results, since they rely heavily on the assumed linear relationship between air pollution and mortality in Equation (1).

because our InMAP simulations indicate that the majority of the $PM_{2.5}$ captured by our wind instrument would come from the conversion of SO_2 into $PM_{2.5}$.

We do not model optimization behavior, which means that our survival projections should be interpreted as holding long-run behavior fixed. Our quasi-experimental estimates reflect short-run behavior that occurs within the time horizon of our outcome window, but not behavior that may arise over longer time horizons. In the context of air pollution, economists distinguish between two types of behaviors: avoidance and mitigation. Avoidance is *ex ante* behavior that reduces pollution exposure, such as migration. Although avoidance behavior complicates the measurement of exposure, it does not affect the application of our model provided that the timing and extent of exposure is correctly specified. For example, our model can be used to predict survival gains when an individual relocates from a high-pollution area to a low-pollution area, assuming the age at which the relocation occurs is known.

Mitigation is *ex post* behavior that reduces the harmful effects of exposure, such as healthcare consumption. Mitigation behavior must be thoughtfully assessed because it can influence the relationship between exposure and model parameters. However, our validation exercises suggest that mitigation behavior has a minimal impact on our projections, at least in the US context over the medium run. For example, our 1-day IV estimates are sufficient for forming accurate 28-day projections (Figure 5), and our three-year projections are consistent with estimates from Anderson (2020). It remains possible that a long-run mitigation response, developing over a lifetime, could affect the accuracy of the model's life expectancy projections, although we note that such long-run responses have not been documented by the literature. In this circumstance, one could interpret model projections as the total cost or benefits of a change in pollution exposure, rather than the total survival effect, provided one is willing to make additional assumptions. In the context of temperature, Carleton et al. (2022) note that an optimizing agent will invest in mitigation up to the point where the marginal benefit equals marginal cost. Applying this insight to pollution implies that the model's projections of the marginal mortality cost of a rise in air pollution will equal

the sum of the mitigation costs and the realized (post-mitigation) mortality costs, i.e., the total marginal cost of the increase in pollution exposure.

6 Conclusion

Accurate estimates of the long-run effect of chronic air pollution exposure on health are vital for making informed policy decisions. We propose a novel two-step approach that combines well-identified short-run estimates of the mortality effect of air pollution with a dynamic production model of health that can be used to form long-run survival projections. Although we focus on air pollution, our method can be applied to other health hazards, provided that the researcher can estimate short-run mortality effects and credibly identify the relevant model parameters affected by the hazard.

To obtain well-identified estimates, we assemble a new dataset that combines daily data on weather, air pollution, and mortality, and then instrument for changes in SO_2 levels using changes in wind direction. We show that the short-run mortality effects of acute SO_2 exposure can be decomposed into two distinct phenomena: mortality displacement, where exposure kills frail individuals with short counterfactual life expectancies, and accelerated aging, where mortality continues to increase after exposure has ceased.

After incorporating our IV estimates into our survival model, we calculate that a permanent, ten percent decrease in air pollution exposure would improve life expectancy by 1.2–1.3 years, holding behavior fixed. These survival benefits are nearly ten times larger than a naive extrapolation of our IV estimates, demonstrating the importance of accounting for latent health changes caused by pollution exposure. While we do not account for the costs of reducing air pollution emissions—a necessary ingredient for performing a full cost-benefit analysis—our estimates imply that value of reducing pollution exposure may be substantially larger than has previously been recognized.

References

Agency for Toxic Substances and Disease Registry (1998). Public health statement for sulfur dioxide. Technical report, U.S. Department of Health and Human Services.

Alexander, D. and H. Schwandt (2022). The impact of car pollution on infant and child health: Evidence from emissions cheating. *The Review of Economic Studies 89*(6), 2872–2910.

Andersen, S., L. Guiso, F. Marciano, P. Sapienza, and L. Zingales (2023). Measuring the Long-Term Impact of Environmental Externalities. Working Paper.

Anderson, M. L. (2020). As the Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality. *Journal of the European Economic Association 18*(4), 1886–1927.

Angrist, J. D., K. Graddy, and G. W. Imbens (2000). The interpretation of instrumental variables estimators in simultaneous equations models with an application to the demand for fish. *The Review of Economic Studies 67*(3), 499–527.

Arenberg, S. and S. Neller (2023). Ashes to ashes: The lifelong consequences of early-life wildfire exposure. Working Paper.

Barreca, A., K. Clay, O. Deschenes, M. Greenstone, and J. S. Shapiro (2016). Adapting to climate change: The remarkable decline in the us temperature-mortality relationship over the twentieth century. *Journal of Political Economy 124*(1), 105–159.

Barreca, A. I., M. Neidell, and N. J. Sanders (2021). Long-run pollution exposure and adult mortality: Evidence from the acid rain program. *Journal of Public Economics 200*, 104440.

Barwick, P. J., S. Li, D. Rao, and N. B. Zahur (2018). The healthcare cost of air pollution: Evidence from the worlds largest payment network. Technical report, National Bureau of Economic Research.

Bishop, K., J. Ketcham, and N. Kuminoff (2023). Hazed and confused: The effect of air pollution on dementia. *The Review of Economic Studies 90*, 2188–2214.

Brook, R. D., S. Rajagopalan, C. A. Pope III, J. R. Brook, A. Bhatnagar, A. V. Diez-Roux, F. Holguin, Y. Hong, R. V. Luepker, M. A. Mittleman, et al. (2010). Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the american heart association. *Circulation 121*(21), 2331–2378.

Burton, A. and T. Roach (2023). Negative externalities of temporary reductions in cognition: Evidence from particulate matter pollution and fatal car crashes. Working Paper.

Carleton, T., A. Jina, M. Delgado, M. Greenstone, T. Houser, S. Hsiang, A. Hultgren, R. E. Kopp, K. E. McCusker, I. Nath, J. Rising, A. Rode, H. K. Seo, A. Viaene, J. Yuan, and A. T. Zhang (2022). Valuing the global mortality consequences of climate change accounting for adaptation costs and benefits. *The Quarterly Journal of Economics 137*(4), 2037–2105.

Chen, Y., A. Ebenstein, M. Greenstone, and H. Li (2013). Evidence on the impact of sustained exposure to air pollution on life expectancy from chinas huai river policy. *Proceedings of the National Academy of Sciences 110*(32), 12936–12941.

Colmer, J. and J. Voorheis (2020). The grandkids aren't alright: the intergenerational effects of prenatal pollution exposure. Working Paper.

Crüts, B., L. van Etten, H. Törnqvist, A. Blomberg, T. Sandström, N. L. Mills, and P. J. Borm (2008). Exposure to diesel exhaust induces changes in eeg in human volunteers. *Particle and Fibre Toxicology 5*(1), 4.

Currie, J. and M. Neidell (2005). Air pollution and infant health: What can we learn from california's recent experience? *Quarterly Journal of Economics 120*(3), 1003–1030.

Currie, J., J. G. Zivin, J. Mullins, and M. Neidell (2014). What do we know about short- and long-term effects of early-life exposure to pollution? *Annual Review of Resource Economics 6*(1), 217–247.

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

Deschênes, O. and M. Greenstone (2011). Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the us. *American Economic Journal: Applied Economics 3*(4), 152–185.

Deschênes, O., M. Greenstone, and J. S. Shapiro (2017). Defensive investments and the demand for air quality: Evidence from the nox budget program. *American Economic Review 107*(10), 2958–89.

Dominici, F., M. Greenstone, and C. R. Sunstein (2014). Particulate matter matters. *Science 344*(6181), 257–259.

Ebenstein, A., M. Fan, M. Greenstone, G. He, and M. Zhou (2017). New evidence on the impact of sustained exposure to air pollution on life expectancy from chinas huai river policy. *Proceedings of the National Academy of Sciences 114*(39), 10384–10389.

Environmental Protection Agency (2004). The particle pollution report: Current understanding of air quality and emissions through 2003. Technical report, U.S. Environmental Protection Agency.

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

Graff Zivin, J. and M. Neidell (2012). The impact of pollution on worker productivity. *American Economic Review 102*(7), 3652–3673.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy 80*(2), 223–255.

Heo, S. W., K. Ito, and R. Kotamarthi (2023). International spillover effects of air pollution: Evidence from mortality and health data. Technical report, National Bureau of Economic Research.

Heutel, G., N. H. Miller, and D. Molitor (2021). Adaptation and the mortality effects of temperature across us climate regions. *Review of Economics and Statistics 103*(4), 740–753.

Hill, W., E. L. Lim, C. E. Weeden, C. Lee, M. Augustine, K. Chen, F.-C. Kuan, F. Marongiu, E. J. Evans Jr, D. A. Moore, et al. (2023). Lung adenocarcinoma promotion by air pollutants. *Nature 616*(7955), 159–167.

Hollingsworth, A. and I. Rudik (2021). The effect of leaded gasoline on elderly mortality: Evidence from regulatory exemptions. *American Economic Journal: Economic Policy 13*(3), 345–73.

Hollingsworth, A. J., D. M. Konisky, and N. Zirogiannis (2021). The health consequences of excess emissions: Evidence from texas. *Journal of Environmental Economics and Management 108*, 102449.

Isen, A., M. Rossin-Slater, and R. Walker (2017). Every breath you take - every dollar you'll make: The long-term consequences of the clean air act of 1970. *Journal of Political Economy 125*(3), 848–902.

Jusot, J.-F., D. R. Neill, E. M. Waters, M. Bangert, M. Collins, L. B. Moreno, K. G. Lawan, M. M. Moussa, E. Dearing, D. B. Everett, et al. (2017). Airborne dust and high temperatures are risk factors for invasive bacterial disease. *Journal of Allergy and Clinical Immunology 139*(3), 977–986.

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

Landrigan, P. J., R. Fuller, N. J. Acosta, O. Adeyi, R. Arnold, A. B. Baldé, R. Bertollini, S. Bose-O'Reilly, J. I. Boufford, P. N. Breysse, et al. (2018). The lancet commission on pollution and health. *The Lancet 391*(10119), 462–512.

Li, H., J. Cai, R. Chen, Z. Zhao, Z. Ying, L. Wang, J. Chen, K. Hao, P. L. Kinney, H. Chen, et al. (2017). Particulate matter exposure and stress hormone levels: a randomized, double-blind, crossover trial of air purification. *Circulation 136*(7), 618–627.

Lleras-Muney, A. and F. Moreau (2022). A unified model of cohort mortality. *Demography 59*(6), 2109–2134.

Luria, M., R. E. Imhoff, R. J. Valente, W. J. Parkhurst, and R. L. Tanner (2001). Rates of conversion of sulfur dioxide to sulfate in a scrubbed power plant plume. *Journal of the Air & Waste Management Association 51*(10), 1408–1413.

Michele, M., M. Alberto, S. Liana, and D. Francesco (2006). Do environmental factors influence the occurrence of acute meningitis in industrialized countries? An epidemic of varying aetiology in Northern Italy. *European Journal of Epidemiology 21*, 465–468.

Nadler, D. L. and I. G. Zurbenko (2014). Estimating cancer latency times using a weibull model. *Advances in Epidemiology 2014*, Article ID 746769, 8 pages.

Olea, J. L. M. and C. Pflueger (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics 31*(3), 358–369.

Qin, G., M. Wu, J. Wang, Z. Xu, J. Xia, and N. Sang (2016). Sulfur dioxide contributes to the cardiac and mitochondrial dysfunction in rats. *Toxicological Sciences 151*(2), 334–346.

Rajagopalan, S. and P. J. Landrigan (2021). Pollution and the heart. *New England Journal of Medicine 385*(20), 1881–1892.

Sang, N., Y. Yun, H. Li, L. Hou, M. Han, and G. Li (2010). So2 inhalation contributes to the development and progression of ischemic stroke in the brain. *Toxicological Sciences 114*(2), 226–236.

Schlenker, W. and M. J. Roberts (2009). Nonlinear temperature effects indicate severe damages to u.s. crop yields under climate change. *Proceedings of the National Academy of Sciences 106*(37), 15594–15598.

Schlenker, W. and W. R. Walker (2016). Airports, air pollution, and contemporaneous health. *Review of Economic Studies 83*(2), 768–809.

Shears, R. K., L. C. Jacques, G. Naylor, L. Miyashita, S. Khandaker, F. Lebre, E. C. Lavelle, J. Grigg, N. French, D. R. Neill, et al. (2020). Exposure to diesel exhaust particles increases susceptibility to invasive pneumococcal disease. *Journal of Allergy and Clinical Immunology 145*(4), 1272–1284.

Stock, J. and M. Yogo (2005). Testing for weak instruments in linear iv regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Chapter 5, pp. 80–108. Cambridge University Press.

Tessum, C., J. Hill, J. Marshall, and D. Paolella (2019, September). Evaluation data for the Intervention Model for Air Pollution (InMAP) version 1.6.1.

Todd, P. E. and K. I. Wolpin (2023). The best of both worlds: combining randomized controlled trials with structural modeling. *Journal of Economic Literature 61*(1), 41–85.

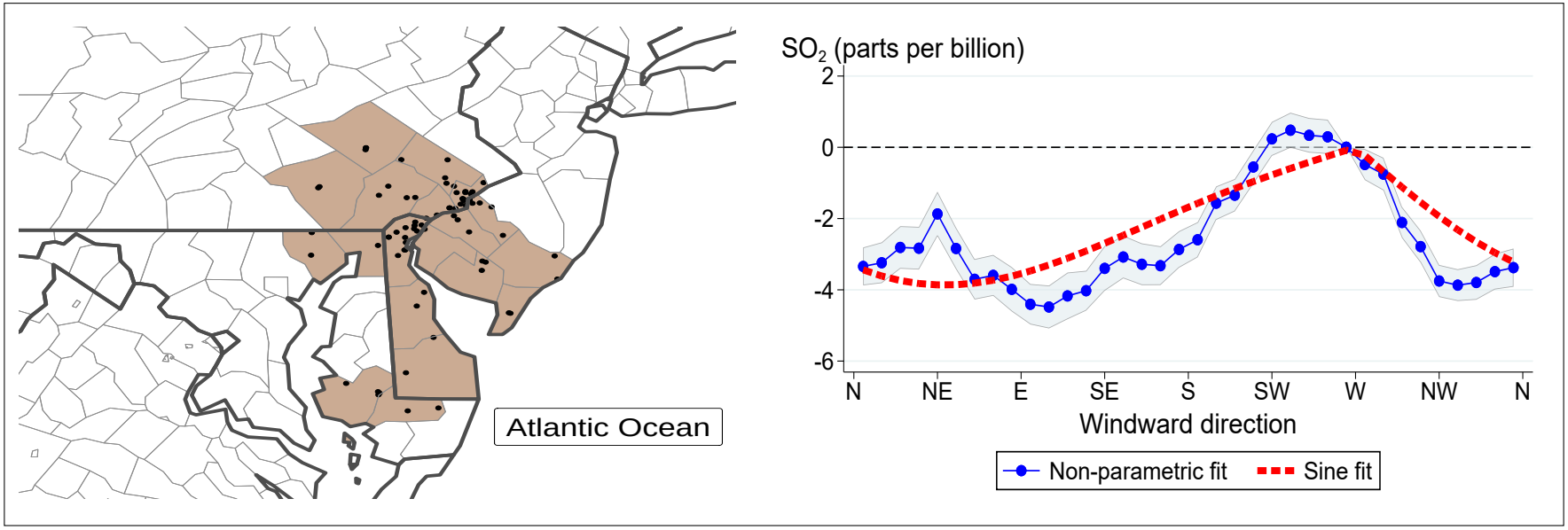
Turner, M. C., Z. J. Andersen, A. Baccarelli, W. R. Diver, S. M. Gapstur, C. A. Pope III, D. Prada, J. Samet, G. Thurston, and A. Cohen (2020). Outdoor air pollution and cancer: An overview of the current evidence and public health recommendations. *CA: a Cancer Journal for Clinicians 70*(6), 460–479.

Voorheis, J. (2017). Air quality, human capital formation and the long-term effects of environmental inequality at birth. US Census Center for Economic Studies, CARRA Working Paper Series.

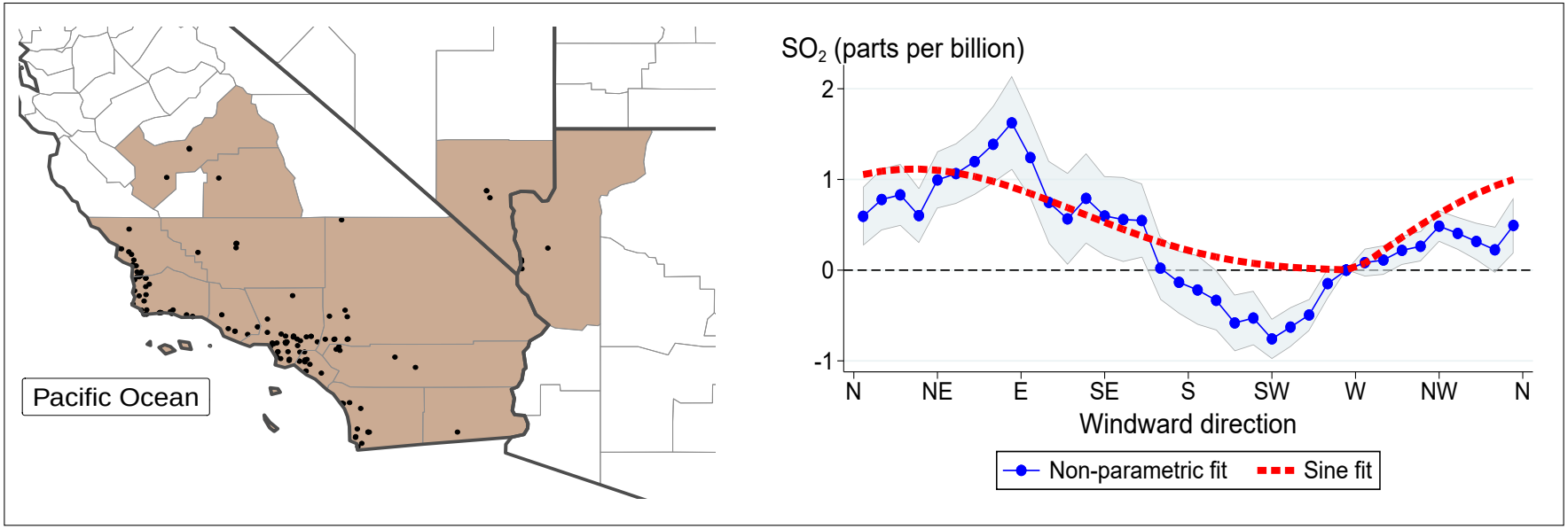
Yao, G., H. Yue, Y. Yun, and N. Sang (2015). Chronic so2 inhalation above environmental standard impairs neuronal behavior and represses glutamate receptor gene expression and memory-related kinase activation via neuroinflammation in rats. *Environmental Research 137*, 85–93.

Figure 1: The relationship between wind direction and SO_2 concentration, Greater Philadelphia and Southern California areas

Greater Philadelphia area

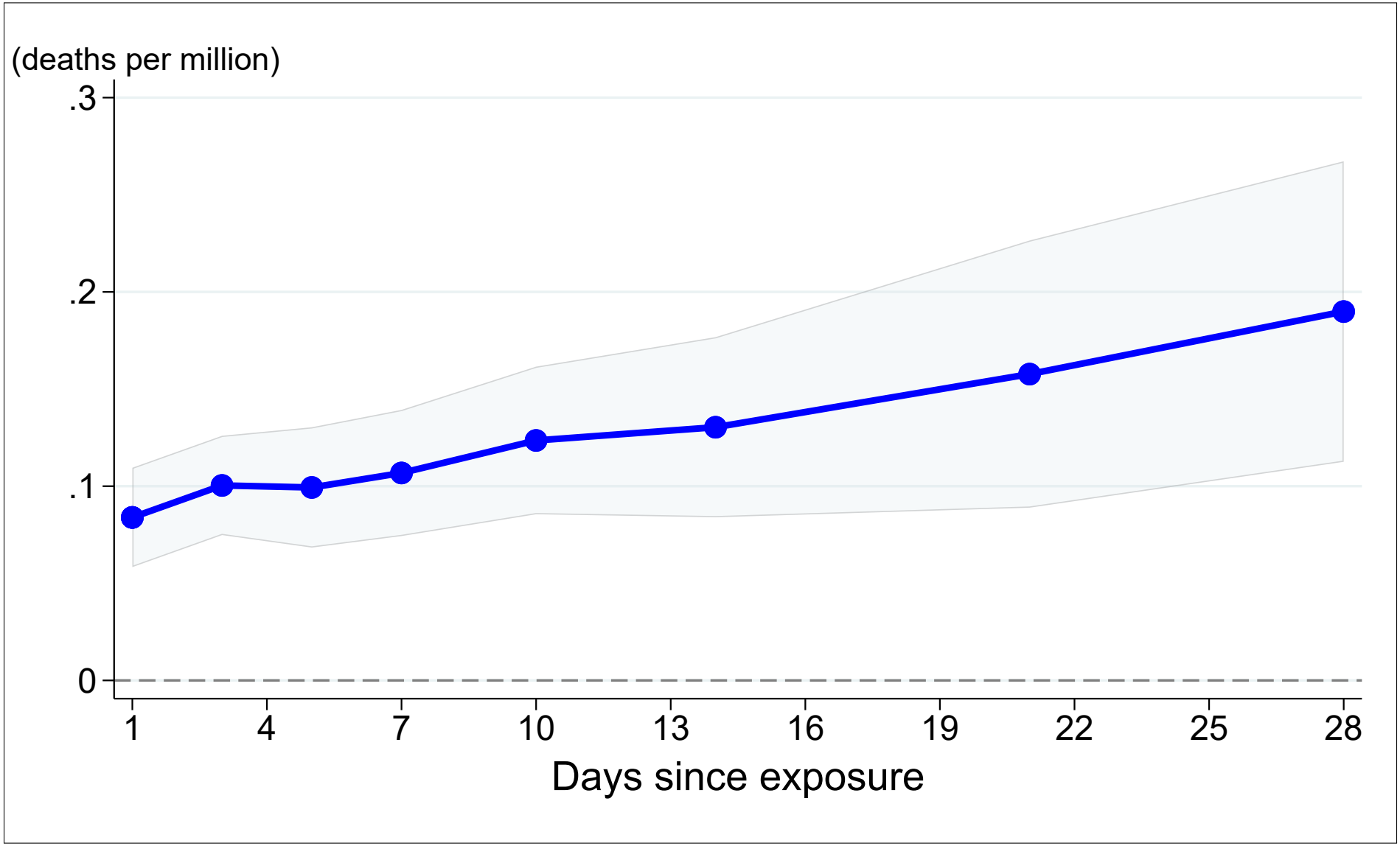


Southern California area



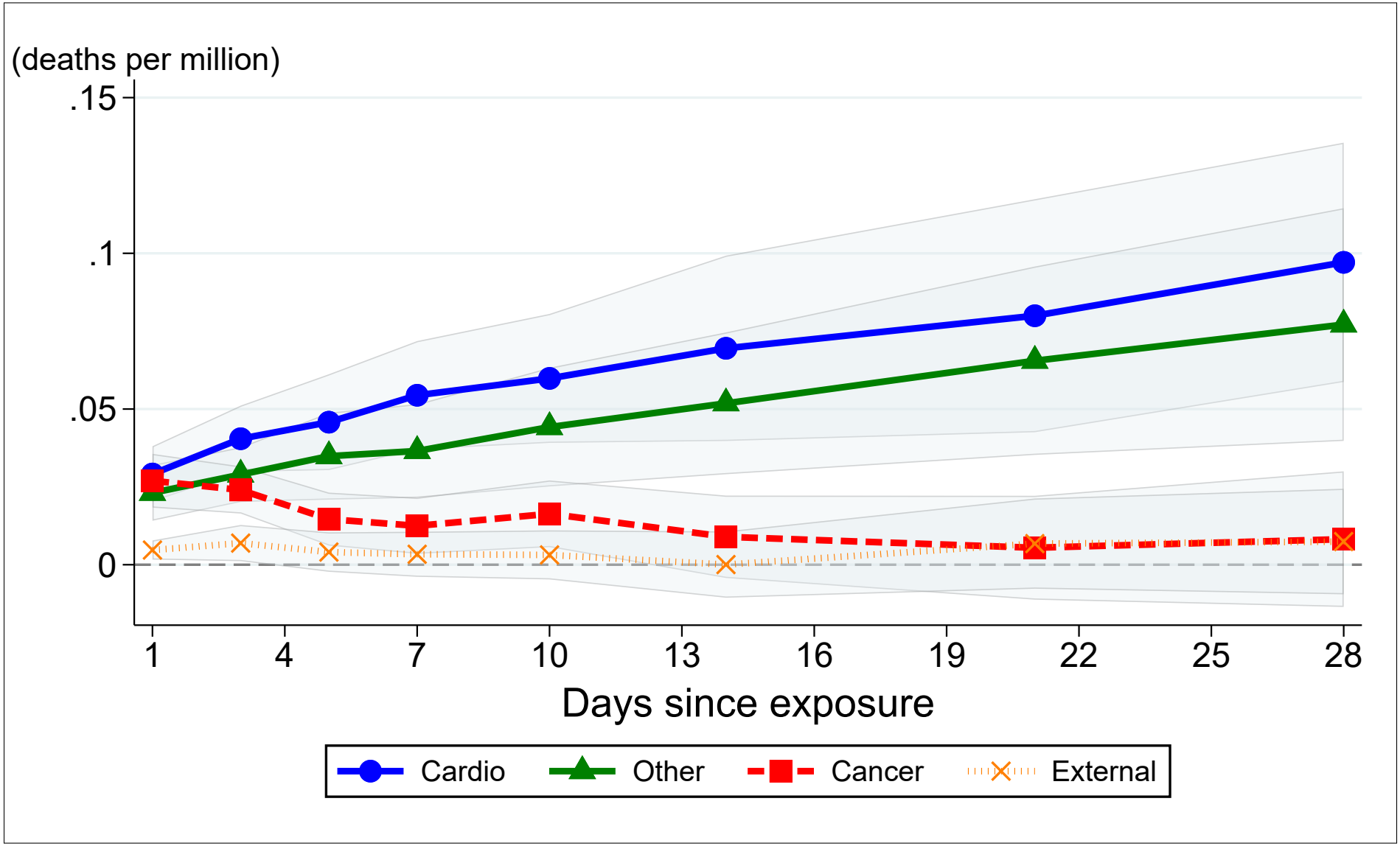
Notes: Sulfur dioxide (SO_2) pollution monitors are depicted as black dots on the two maps. The graphs on the right plot the relationship between SO_2 levels and windward direction in each area. Windward direction describes where the wind is blowing from, with "N" indicating North, "NE" indicating Northeast, etc. The 36 blue points report coefficient estimates from a non-parametric regression of SO_2 on wind direction measured in 10-degree angle bins. The blue shaded area shows the corresponding 95% confidence intervals. The red dashed lines report fitted curves from the parametric sine specification given by $f^g(\theta)$ in regression Equation (2). All regressions include county-by-month and month-by-year fixed effects, as well as flexible weather controls. Standard errors are robust to heteroskedasticity.

Figure 2: IV estimates of effect of acute (1-day) SO_2 exposure on cumulative mortality



Notes: Each point reports an IV estimate from Equation (1) of the effect of acute (1-day), 1-ppb sulfur dioxide (SO_2) exposure on mortality, where mortality is measured as cumulative deaths per million over a time window ranging from 1 to 28 days, as indicated by the x-axis. The shaded area represents 95% confidence intervals. Point estimates are reported in Column (2) of Table A.2. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors are clustered by county.

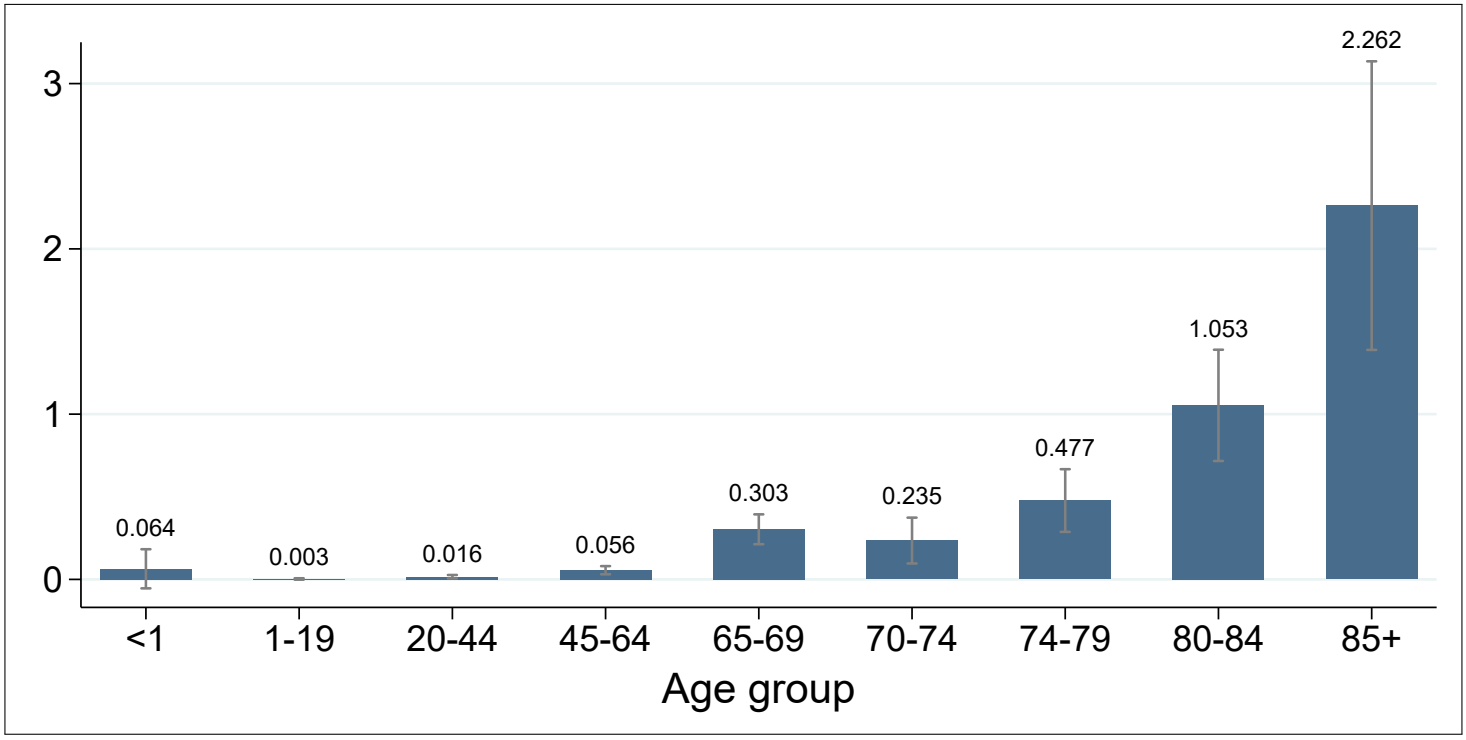
Figure 3: IV estimates of effect of acute (1-day) SO_2 exposure on cumulative mortality, by cause of death



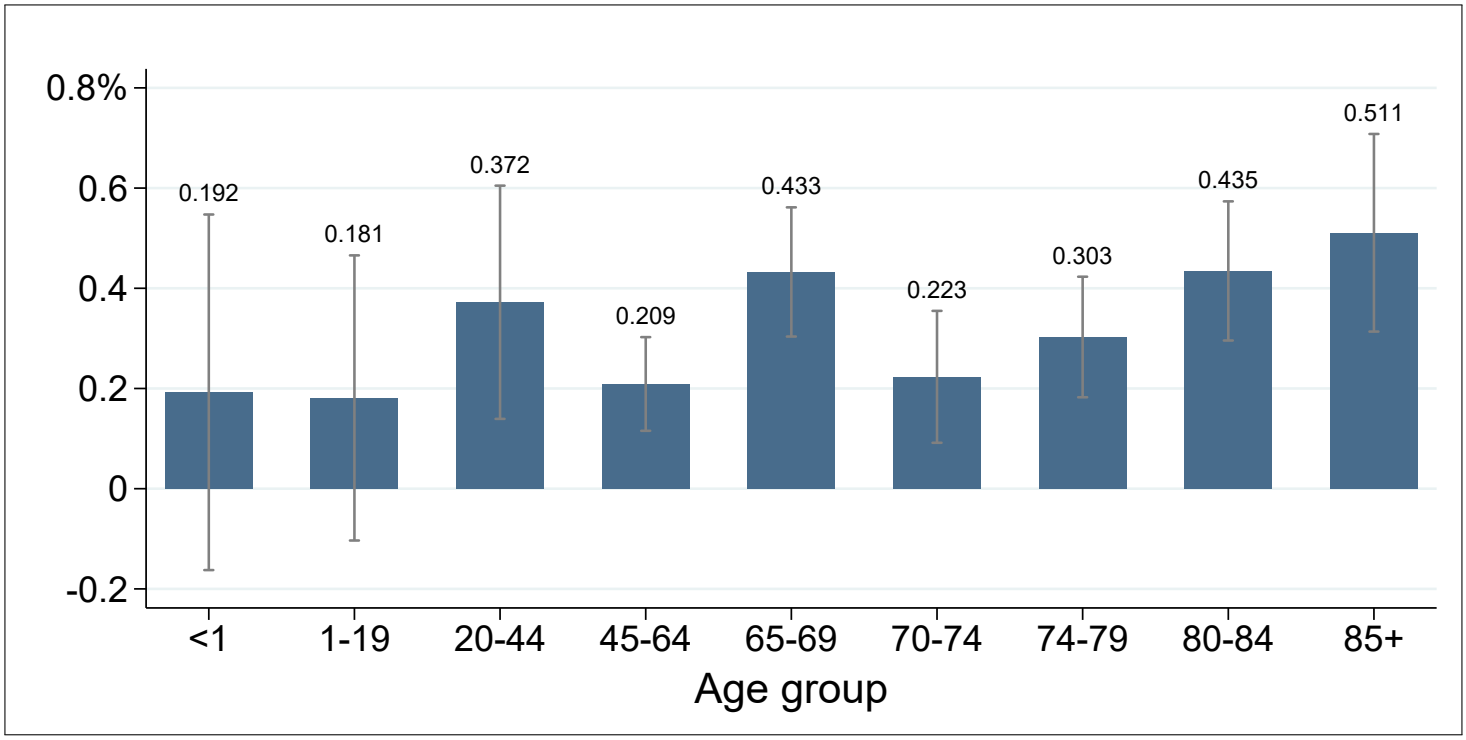
Notes: Each point reports an IV estimate from Equation (1) of the effect of acute (1-day) sulfur dioxide (SO_2) exposure on mortality (deaths per million), for four different causes of death: cardiovascular disease, other diseases, cancer, and external causes of death. Mortality is measured over a time window ranging from 1 to 28 days, as indicated by the x-axis. Shaded areas represent 95% confidence intervals. Point estimates are reported in Table A.5. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors are clustered by county.

Figure 4: IV estimates of effect of acute (1-day) SO_2 exposure on 1-day mortality, by age group

(a) Absolute increase (deaths per million)

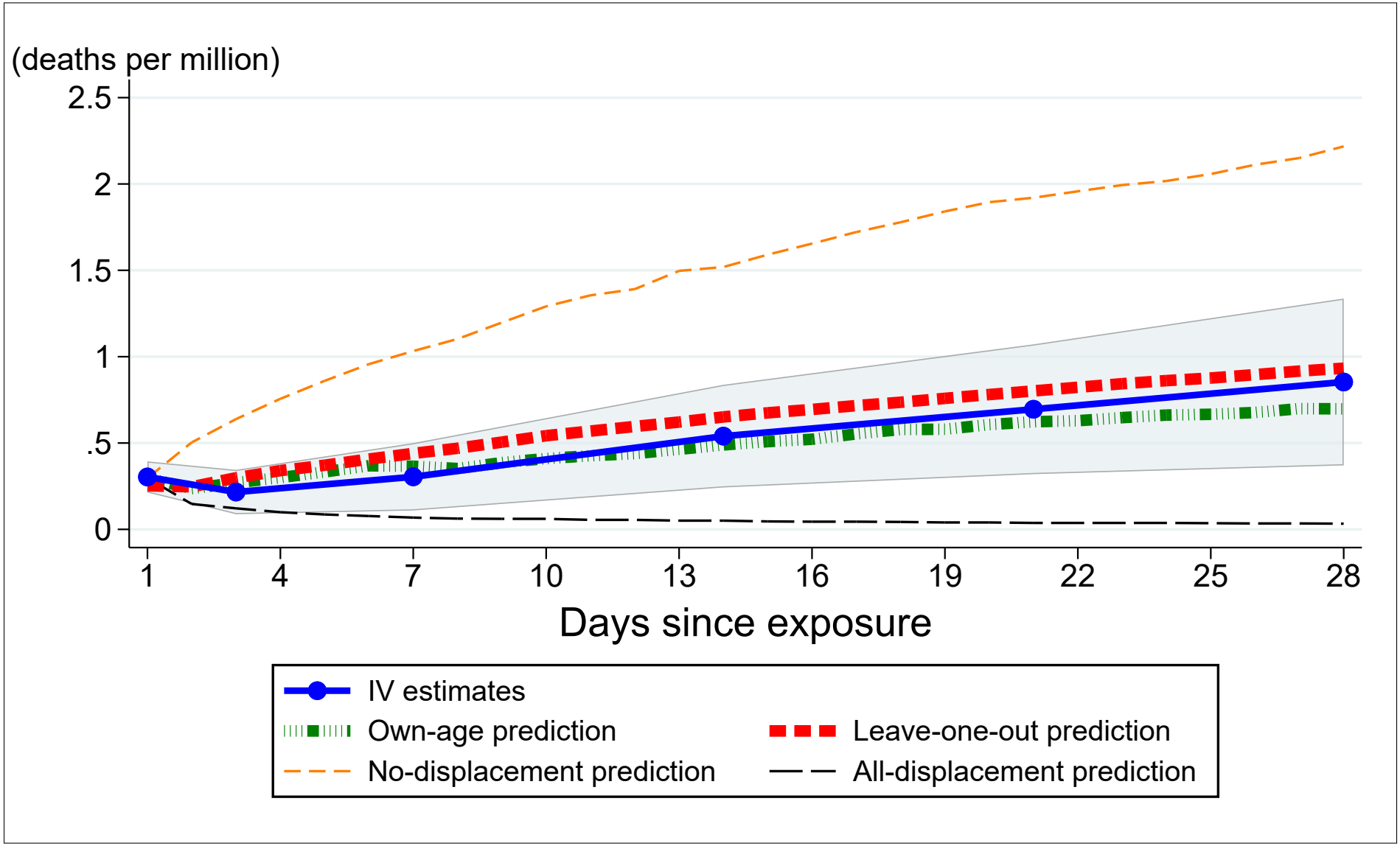


(b) Relative increase (percent of 1-day mortality)



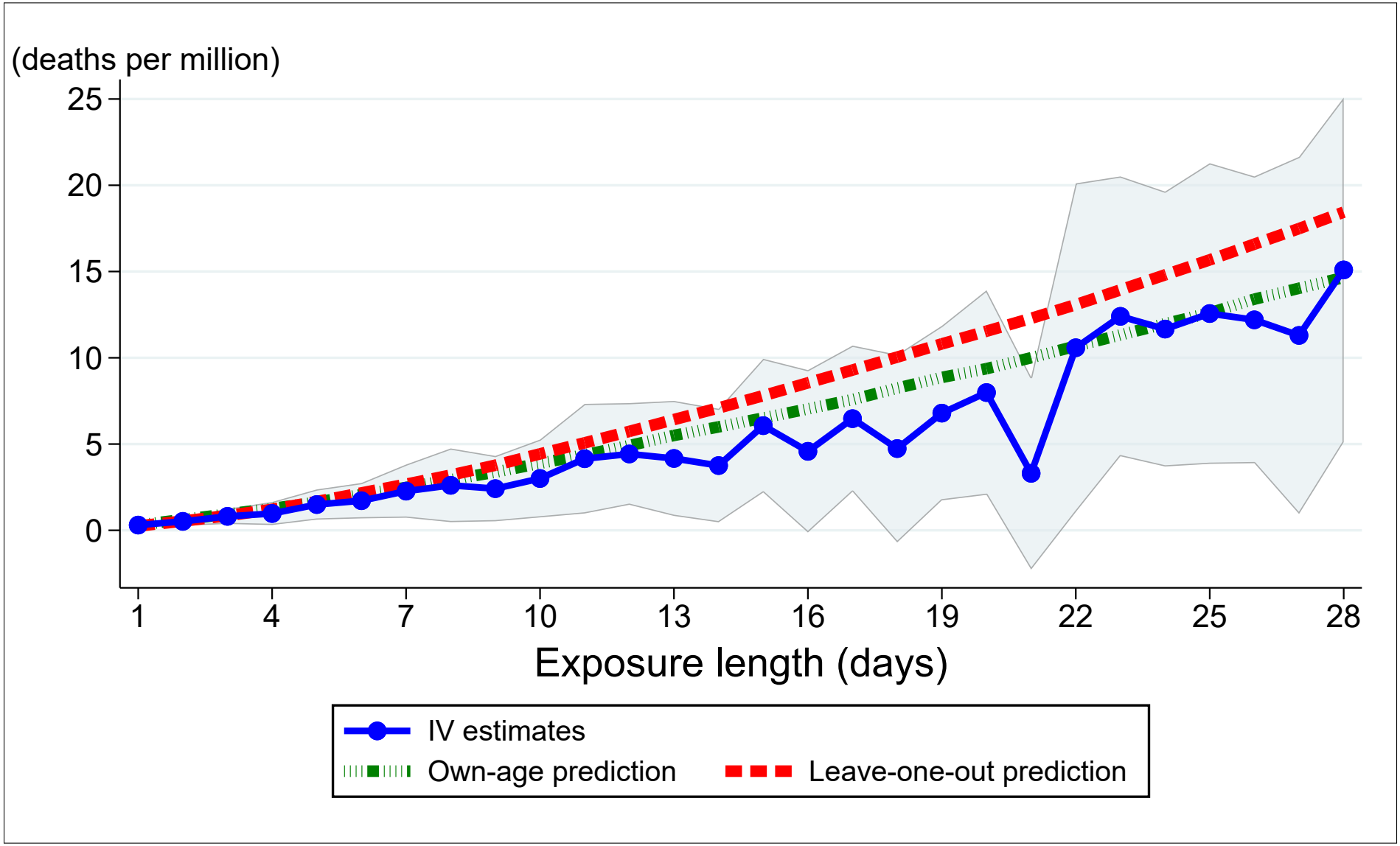
Notes: Each bar represents an IV estimate from Equation (1) of the effect of acute (1-day) sulfur dioxide (SO_2) exposure on 1-day mortality for a particular age group. Error bars represent 95% confidence intervals. Estimates are also reported in Table A.3. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors are clustered by county. IV estimates for 28-day mortality are shown in Figure A.9.

Figure 5: Model predictions of effect of acute (1-day) SO_2 exposure on one-month survival, ages 65–69



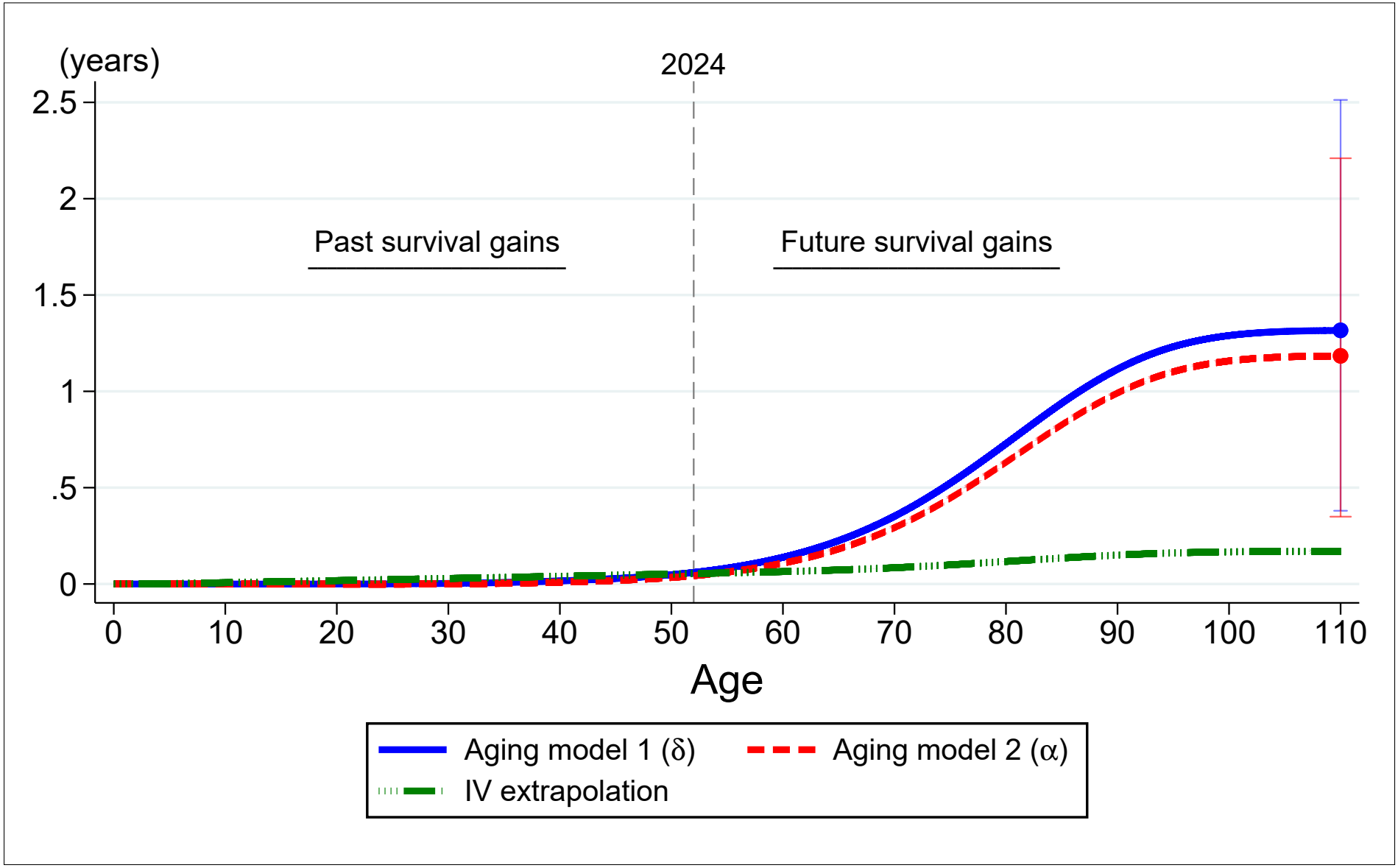
Notes: The solid blue line reports IV estimates from Equation (1) of the effect of acute (1-day) sulfur dioxide (SO_2) exposure on cumulative mortality for the 65–69 age group, with 95% confidence intervals given by the blue shaded area. The thick red and thick green dashed lines report corresponding predictions from the dynamic production model of health described by Equation (3). The "own-age" prediction, depicted by the green dashed line, is calibrated so that its 1-day prediction matches the 1-day IV estimate (first blue point), and it attributes the cancer-related portion of the 1-day IV estimate to the frailest individuals with the lowest levels of health capital. The "leave-one-out" model, depicted by the thick red dashed line, employs the average of the calibrated values for all of the other older age groups, i.e., ages 70–74, 75–79, 80–84, and 85+. The orange dashed line ("no-displacement") at the top if the figure reports model predictions under the extreme assumption that none of the 1-day mortality effect is mortality displacement, while the black dashed line ("all-displacement") at the bottom of the figure reports predictions under the alternative extreme assumption that all of the 1-day effect is mortality displacement. Figure A.12 shows plots for all age groups 65 and over.

Figure 6: Model predictions of effect of short-term chronic SO_2 exposure on survival, ages 65–69



Notes: The solid blue line reports IV estimates of the effect of short-term chronic sulfur dioxide (SO_2) exposure on cumulative mortality for the 65–69 age group, with 95% confidence intervals given by the blue shaded area. These estimates are obtained by isolating exogenous variation in daily SO_2 from our first-stage regression given by Equation (2), aggregating the data up to the level of days denoted by the x-axis, and then regressing the mortality rate on average SO_2 levels, using the average first-stage prediction as an instrument. The other two lines report corresponding predictions from the dynamic production model of health described by Equation (3). The "own-age" prediction, depicted by the green dot-dashed line, is calibrated so that its 1-day prediction matches the 1-day IV estimate (first blue point), and it attributes the cancer-related portion of the 1-day IV estimate to the frailest individuals with the lowest levels of health capital. The "leave-one-out" model, depicted by the thick red dashed line, employs the average of the calibrated values for all of the other older age groups, i.e., ages 70–74, 75–79, 80–84, and 85+.

Figure 7: Projected effect of a permanent 1-unit decrease in SO_2 on survival gains for cohort born in 1972



Notes: This figure shows the cumulative effect of a permanent, 1-unit decrease in sulfur dioxide (SO_2) on survival gains for the cohort of US individuals born in 1972. These projections are produced by the dynamic production model of health (3), which was calibrated using our 1-day IV estimates from Equation (1). The solid blue line ("Aging model 1") shows projections under the assumption that non-cancer-related pollution deaths are governed by changes in the model's depreciation parameter α, and the dashed red line ("Aging model 2") shows projections under the alternative assumption that they are governed by the depreciation parameter δ. The dot-dashed green line ("IV extrapolation") projects changes in life expectancy by extrapolating our age-specific 28-day IV estimates to the whole life-cycle. The cumulative effects at age 110 equal the values reported in the first row of Table A.16. Error bars represent 90% bootstrap confidence intervals.

Table 1: Summary statistics, 1972–1988

	(1)	(2)	(3)
	Mean	Std. Dev.	Observations
A. Pollution outcomes			
SO_2, ppb	8.96	12.63	2,023,369
NO_2, ppb	21.25	15.61	789,506
CO, ppm	1.64	1.37	844,334
O_3, ppb	25.58	13.70	665,512
TSP, μg/m^3	63.14	40.21	627,359
B. One-day mortality rate outcomes			
All-cause mortality, deaths per million	24.69	24.32	2,023,369
Cardiovascular	12.21	16.05	2,023,369
Cancer	5.15	9.16	2,023,369
Other	5.44	10.01	2,023,369
External	1.89	8.00	2,023,369
All-cause mortality by age group, deaths per million			
Age 1 and under	33.37	166.57	2,023,369
Age 1–19	1.51	11.54	2,023,369
Age 20–44	4.41	18.46	2,023,369
Age 45–64	26.60	49.03	2,023,369
Age 65–69	70.13	170.48	2,023,369
Age 70–74	105.32	241.99	2,023,369
Age 75–79	157.60	360.39	2,023,369
Age 80–84	242.30	575.47	2,023,369
Age 85+	442.55	924.22	2,023,369

Notes: Unit of observation is a county-day. Statistics are unweighted. Sample is restricted to observations where both mortality and sulfur dioxide (SO_2) are non-missing. Mortality is calculated as the number of daily deaths per million individuals. Pollution data are from the Environmental Protection Agency, mortality counts are from the National Vital Statistics, and population estimates are from the Surveillance, Epidemiology, and End Results (SEER) Program.

Table 2: OLS and IV estimates of effect of acute SO_2 exposure on 1-day mortality

	OLS	IV			
	(1)	(2)	(3)	(4)	(5)
SO_2, ppb	0.0077*	0.084**	0.083**	0.085**	0.097**
	(0.0031)	(0.013)	(0.013)	(0.012)	(0.014)
First-stage F-statistic		526	523	515	560
Mean outcome	25	25	25	25	25
Sample size	2,023,369	2,023,369	2,023,301	2,022,782	2,023,369
Weather controls					
Baseline weather variables	X	X	X	X	
Minimum temperature variables			X	X	
More granular bins				X	

Notes: Dependent variable is number of deaths per million people on the day of exposure. All regressions include county-by-month and month-by-year fixed effects, as well as two leads and two lags of the instruments. Column (2) reports our baseline specification, which includes county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Column (3) additionally includes controls for minimum temperature. Column (4) further increases the number of weather controls by using more granular bins. Column (5) excludes all weather controls. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table 3: IV estimates of effect of acute SO_2 exposure on 1-day mortality, controlling for other pollutants

	(1)	(2)	(3)	(4)
A. All-pollutant sample				
SO_2, ppb	0.084**	0.060**	0.083**	0.064**
	(0.012)	(0.013)	(0.014)	(0.014)
TSP, μg/m^3		0.012**		0.015**
		(0.0036)		(0.0035)
NO_2, ppb			0.017	0.0024
			(0.020)	(0.017)
O_3, ppb			–0.036	–0.046*
			(0.025)	(0.022)
CO, ppm			–0.21	–0.24
			(0.25)	(0.20)
First-stage F-statistic	81	21	10	10
Mean outcome	27	27	27	27
Sample size	78,946	78,946	78,946	78,946
B. SO_2/TSP sample				
SO_2, ppb	0.079**		0.035*	
	(0.014)		(0.015)	
TSP, μg/m^3		0.027**	0.019**	
		(0.0046)	(0.0045)	
First-stage F-statistic	200	143	50	
Mean outcome	25	25	25	
Sample size	627,290	627,290	627,290	

Notes: Dependent variable is number of deaths per million people on the day of exposure. Regressions are estimated using two-stage least squares where each pollutant is treated as an endogenous regressor. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level. National means for all air pollutants are available in Table 1. Table A.6 presents estimates using a third sample that includes all pollutants except TSP.

Online Appendix

"The Long-run Effect of Air Pollution on Survival"

Tatyana Deryugina, University of Illinois and NBER
Julian Reif, University of Illinois and NBER

A Supplementary information and analysis

A.1 Data

Monitor-level data for sulfur dioxide (SO_2), total suspended particulates (TSP), nitrogen dioxide (NO_2), O_3, and carbon monoxide (CO) for the years 1972–1988 were obtained by email request from the US Environmental Protection Agency (EPA). Each SO_2 observation provides a sample measure, usually recorded over a period of one hour. We dropped SO_2 observations that had a sample duration greater than 24 hours, or that reported an SO_2 measurement above 1000 parts per billion (ppb) or below –2 ppb.[1] We dropped CO, NO_2, and O_3 observations with negative values. All TSP observations had non-negative values. We then aggregated all pollutants to the monitor-day level, weighting by the time duration of each measure. Finally, data were aggregated to the county-day level by averaging over all monitors within a county.

Figure A.1 displays the population-weighted concentrations and the number of counties with at least one operational monitor over time for each pollutant. Except for O_3, the population-weighted means for all pollutants decline substantially during our sample period. CO data are readily available since the mid-1970's and maintain consistent coverage of approximately 225 counties per year, while O_3 data are unavailable prior to 1980. NO_2 coverage is high for most of the 1970s but drops dramatically by the late 1980s, while data on SO_2 are available for a larger number of counties than most other pollutants throughout our sample period. Each year during our sample period, at least 400 counties monitor SO_2 concentrations, and about 50 percent of US individuals live in a county that monitors SO_2. While TSP monitors cover more counties than SO_2 monitors, TSP monitors perform measurements less frequently and are more likely to be located in lower population areas. At the county-day level, we have 62 percent fewer population-weighted observations of TSP than of SO_2.

Figure A.3 shows the locations of the 4,740 active SO_2 pollution monitors during our 1972–1988 sample period. The monitors are present in 1,041 counties.

A.2 Complier analysis and robustness checks

Table A.1 reports an analysis of complier characteristics where we regress county-level measures obtained from the Regional Economic Information System (REIS) dataset on the strength of the first stage. The unit of observation is a county-year, and our regressions include year fixed effects. The strength of the first stage is not significantly associated with population or the percent of residents over age 65. Counties with

[1] According to the AQS Data Coding Manual version 2.38 (February 2, 2010), the maximum allowable sample value for SO_2 is 1000 ppb. The EPA informed us by email that small negative values can arise due to noise and should be included in sample averages to avoid bias. We chose –2 as the bottom cutoff because it appeared to be the minimum allowable sample value.

stronger first stages have a smaller share of residents who are Black, have higher per-capita income, receive higher per-capita transfers, and have a lower employment rate, but the magnitudes of these coefficients are generally small (e.g., a \$90 difference in per-capita income for each 1 ppb difference in the first stage). The largest relative differences arise among the population share that is Black (1.4 percentage points lower Black share per 1 ppb SO_2, relative to a mean of 9.2 percent) and mean SO_2 concentrations (1.4 ppb higher relative to a mean of 7.4 ppb). Along other dimensions, the complier group is not notably different from the US population.

Table A.7 investigates the sensitivity of our estimates to including alternative sets of fixed effects. Column (1) reproduces our baseline estimate, while Columns (2)–(6) present five reasonable alternatives, including variants that control for state-by-calendar-month-by-year (Column 3) or county-by-year (Column 4) fixed effects. These have little effect on the size of our estimate, suggesting that seasonal variation in the climate-mortality relationship and unobserved variables that vary across locations do not bias our estimates. Table A.8 shows that clustering standard errors at larger geographic levels such as state or geographic group has little effect on the size of our standard errors.

Our main specification includes two leads and two lags of our instruments in order to account for autocorrelation in wind direction, which might otherwise cause an upward bias in our estimates. Table A.9 shows that our estimates are insensitive to including more or fewer leads and lags.

We interpret our IV estimate as a weighted average treatment effect where the weights are non-negative, which requires monotonicity of air pollution in the instruments (Angrist, Graddy and Imbens, 2000). In other words, if the instruments increase SO_2 levels in one county, then SO_2 levels in other counties assigned to the same geographic group cannot fall. This assumption would be violated if the relationship between wind direction and SO_2 varies across counties located within the same group. We investigate this possibility with two alternative specifications that allow the effect of wind direction on SO_2 to vary over either a larger or smaller geographic area. Columns (2)–(3) of Table A.10 show that these alternatives produce estimates similar to our main estimate, given in Column (1), suggesting that violations of the monotonicity assumption are not a significant concern in our setting.

Columns (4)–(6) of Table A.10 show that our estimates change little if we employ the following non-parametric specification for $f^g(\theta_{cd})$:

$$f^g(\theta_{cd}) = \sum_{i=1}^{b} \gamma_g^i 1[G_c = g] \times 1\left[\frac{360}{b}(i-1) \leq \theta_{cd} < \frac{360}{b} i\right]$$

where the first indicator function, $1[G_c = g]$, is equal to 1 if county c is a member of group g and 0 otherwise. The second indicator function is equal to 1 if the local wind direction, θ_{cd}, now expressed in degrees rather than radians, lies between $\frac{360}{b}(i-1)$ and $\frac{360}{b}i$ for some fixed value b. For example, when $b = 4$ and $i = 1$, the indicator function will equal 1 if the wind direction lies between 0 and 90 degrees. Columns (4)–(6) report estimates when we set b equal to 9, 6, and 4, respectively.

Our reported F-statistics in Table 2 exceed 500 and our 2SLS estimates differ significantly from OLS estimates, indicating that our estimates do not suffer from weak instrument bias. Nevertheless, we conduct two additional checks to assess the quality of our instruments. First, Table A.11 shows that our 2SLS

estimates are similar to estimates obtained from LIML, which is approximately median unbiased even in the presence of many weak instruments. Second, we conduct a placebo exercise where we generate a set of random wind directions and use those in place of the actual wind direction (Columns 1–3 of Table A.12). The first-stage F-statistics for those placebo exercises never exceed 2.0, which demonstrates that our wind direction instrument picks up meaningful rather than spurious variation in SO_2 levels.

Column (4) of Table A.12 presents the results of a falsification test that regresses 1-day mortality on *future* SO_2 levels, yielding a small and statistically insignificant coefficient. These results provide evidence that our inference methodology does not significantly understate the magnitude of our standard errors and further reinforces the credibility of our identification strategy.

A.3 Model calibration

The dynamic production model of health given by Equation (3) depends on seven parameters: $\{\alpha, \delta, I, \sigma_e, \mu_H, \sigma_H, \underline{\mathrm{H}}\}$. To achieve identification, we normalize two parameters: $\underline{\mathrm{H}}^* = 0$ and $\sigma_H^* = 1$. We calibrate the baseline values for the five remaining parameters using a 1972 period life table, as described in Section 5.2. Let $\{\alpha^*, \delta^*, I^*, \sigma_e^*, \mu_H^*\}$ denote those calibrated values.

We calculate the life expectancy effects of changes in chronic air pollution exposure for two different scenarios. In the first scenario, we assume that pollution exposure only affects the parameters δ and $\underline{\mathrm{H}}$, and denote their post-exposure values as $\{\widetilde{\delta}, \widetilde{\underline{\mathrm{H}}}\}$. In the second scenario, we assume that exposure affects α rather than δ. Below, we describe how we solve for $\{\widetilde{\delta}, \widetilde{\underline{\mathrm{H}}}\}$. It is straightforward to solve for $\widetilde{\alpha}$ instead of $\widetilde{\delta}$ using the methodology described below.

Consider a population of N individuals whose health capital evolves according to Equation (3). Let $\Theta^* = \{\alpha^*, I^*, \mu_H^*, \sigma_H^*, \sigma_e^*\}$ represent the five baseline calibrated parameters unaffected by air pollution exposure. Let S be a random-number seed that fixes the initial stock of health capital, H_{i0}, and the evolution of the iid shock, ε_{it}, for all individuals. Then we can define $M_t\left(\delta^*, \underline{\mathrm{H}}^* \middle| \Theta^*, N, S\right)$ as the deterministic mortality rate at time $t \geq 0$ for a cohort born in 1972, as computed by the model using the baseline parameter values.

Let $\hat{\beta}_{a,c}^k$ denote the IV estimate of the effect of acute exposure on cumulative mortality for age group a from cause of death c in the k days following exposure. Let $[t_a^0, t_a^1]$ define the age interval spanned by age group a, measured in days. Consider a specific age, $t \in [t_a^0, t_a^1]$. To quantify the effect of exposure on model parameters, we first solve for the new mortality threshold, $\widetilde{\underline{\mathrm{H}}}_{at}$, which is defined implicitly by the following equation:

$$\hat{\beta}_{a,cancer}^k = M_t\left(\delta^*, \widetilde{\underline{\mathrm{H}}}_{at} \middle| \Theta^*, N, S\right) - M_t\left(\delta^*, \underline{\mathrm{H}}^* \middle| \Theta^*, N, S\right) \tag{A.1}$$

We then solve for $\widetilde{\delta}$, which is defined implicitly by:

$$\hat{\beta}_{a,all}^k = M_t\left(\widetilde{\delta}_{at}, \widetilde{\underline{\mathrm{H}}}_{at} \middle| \Theta^*, N, S\right) - M_t\left(\delta^*, \widetilde{\underline{\mathrm{H}}}_{at} \middle| \Theta^*, N, S\right) \tag{A.2}$$

Health capital is strictly decreasing in δ for all individuals (see Equation 3), and death occurs when an individual's health stock falls below $\underline{\mathrm{H}}$. Thus, the mortality rate, $M_t(\cdot)$, is monotonically increasing in both

δ and $\underline{\mathrm{H}}$. Consequently, the solutions $\{\widetilde{\delta}_{at}, \widetilde{\underline{\mathrm{H}}}_{at}\}$ to Equations (A.1) and (A.2) exist and are unique.[2] The solutions can be computed using standard root-finding algorithms.

We can solve for $\{\widetilde{\delta}_{at}, \widetilde{\underline{\mathrm{H}}}_{at}\}$ for any $t \in [t_a^0, t_a^1]$. We use the approximate integer midpoint of each age bin, $t = round\,[(t_a^0 + t_a^1)/2]$.[3] However, using only a single day to calibrate the parameters produces a noisy solution because health capital in Equation (3) also varies with an iid health shock. To improve precision, we solve Equations (A.1) and (A.2) using 50 different days around the approximate midpoint of each age bin and take the average. For example, for $a = 65$ (the 65–69 age group), we solve for $\{\widetilde{\delta}_{65,t}, \widetilde{\underline{\mathrm{H}}}_{65,t}\}$ using ages $t =$ 68y1d, $t =$ 68y2d, ..., $t =$ 68y50d.[4] Figure A.13 shows the cumulative mortality effects of acute (1-day) exposure to air pollution for 50 separate daily ages from the 65–69 age group, as predicted by the model using the values of $\{\widetilde{\delta}_{65,t}, \widetilde{\underline{\mathrm{H}}}_{65,t}\}$ for each age t. By construction, the effect shown for the first day is always equal to the 1-day IV estimate for the 65–69 age group (see first row of Column (5) in Table A.4). The subsequent values report longer-run effects of exposure up to 30 days later. Each value reports mortality in the counterfactual relative to mortality in the baseline case (no exposure). A decrease in the cumulative value indicates mortality displacement. The "own-age prediction" in Figure 5 reports the average of these 50 plots.

We compute the age-specific parameter solutions $\{\widetilde{\underline{\mathrm{H}}}_a, \widetilde{\delta}_a\}$ as the averages of these 50 solutions:

$$\widetilde{\underline{\mathrm{H}}}_a = \frac{1}{50} \sum_t \widetilde{\underline{\mathrm{H}}}_{at}$$
$$\widetilde{\delta}_a = \frac{1}{50} \sum_t \widetilde{\delta}_{at}$$

Figure A.11 reports $\{\widetilde{\underline{\mathrm{H}}}_a, \widetilde{\delta}_a\}$ for the five oldest age groups. The estimates are expressed as deviations from the baseline calibrated values (i.e., $\widetilde{\underline{\mathrm{H}}}_a - \underline{\mathrm{H}}^* = \widetilde{\underline{\mathrm{H}}}_a$ and $\widetilde{\delta}_a - \delta^*$). The parameters used to predict the long-run survival effects of air pollution are computed as the average across these five age groups:

$$\widetilde{\underline{\mathrm{H}}} = \frac{1}{5} \sum_a \widetilde{\underline{\mathrm{H}}}_a$$
$$\widetilde{\delta} = \frac{1}{5} \sum_a \widetilde{\delta}_a$$

To account for econometric uncertainty in the IV estimates, we use a resampling-based methodology. We randomly draw an estimate of the effect of acute pollution exposure on 1-day mortality from a normal distribution with a mean and standard deviation set equal to the mean and standard error of $\widehat{\beta}^1_{a,c}$ in Equation (1), and then calibrate the change in model parameters to match the mortality change draw. We repeat this exercise 100 times and report the 5th and 95th percentiles of the resulting distribution of

[2] Alternatively, one could solve first for $\widetilde{\delta}$ and then for $\widetilde{\underline{\mathrm{H}}}$. That would yield estimates that are numerically very close but not identical to the ones resulting from this method. We solve for $\widetilde{\underline{\mathrm{H}}}$ first because it is more efficient: when solving for $\widetilde{\alpha}$, we do not need to separately solve for $\widetilde{\underline{\mathrm{H}}}$ again as it was already computed when solving for $\widetilde{\delta}$.

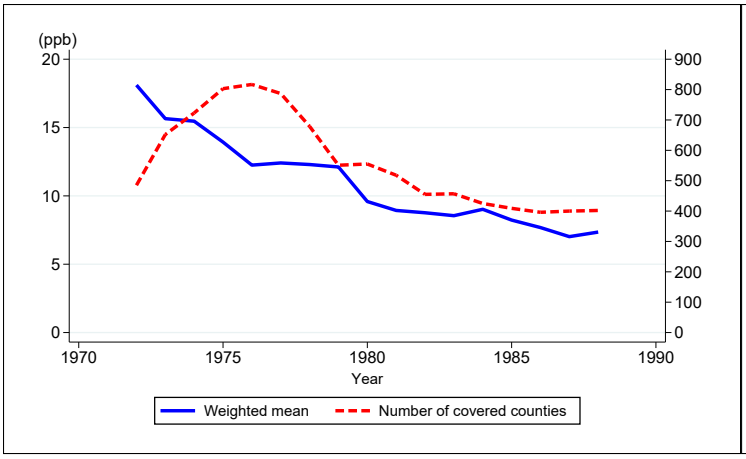
[3] For the 85+ age group, we use a midpoint of 90, which is the average age of death in that group during our sample period.

[4] The optimal strategy would employ all $365 \times 5 = 1825$ days in the 5-year age bin, giving more weight to the ages near the midpoint. However, doing so is computationally burdensome and unnecessary for achieving sufficient accuracy.

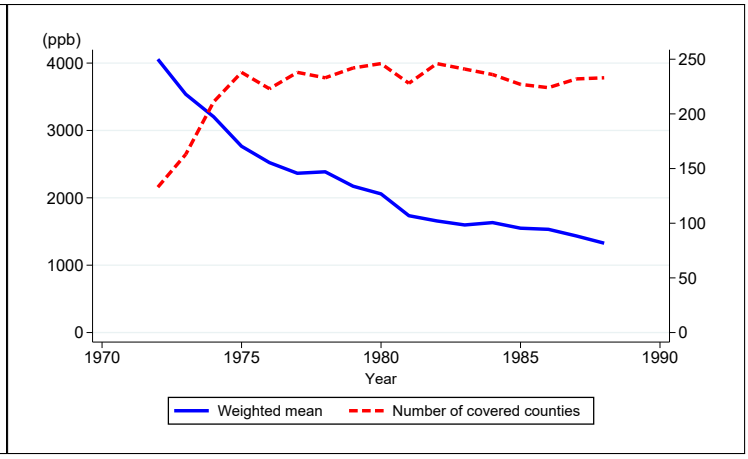
model parameter estimates. To reduce the computational burden, we use 20 different days rather than 50 when computing $\{\underline{\widetilde{\mathrm{H}}}_a, \widetilde{\delta}_a\}$ during the resampling.

Figure A.1: Air pollution means and population coverage levels, 1972–1988

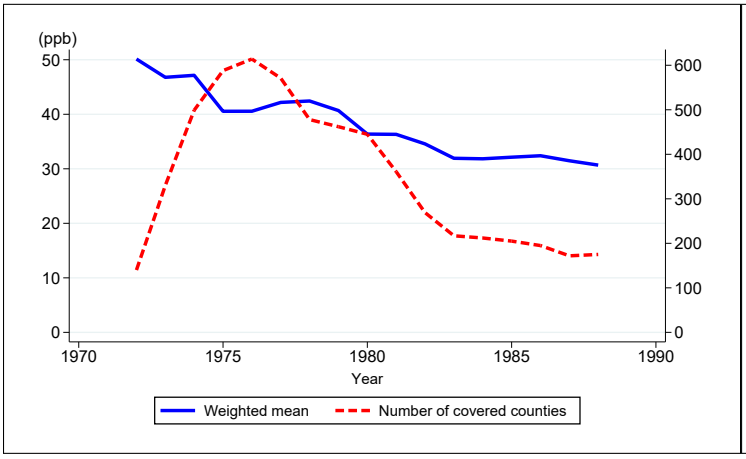
(a) SO_2



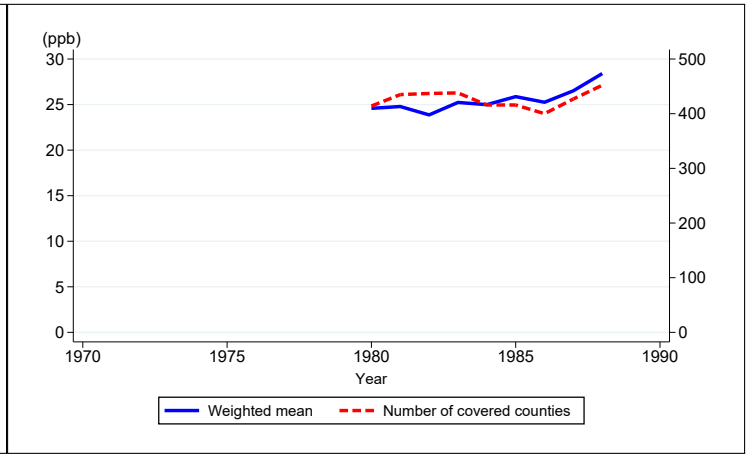
(b) CO



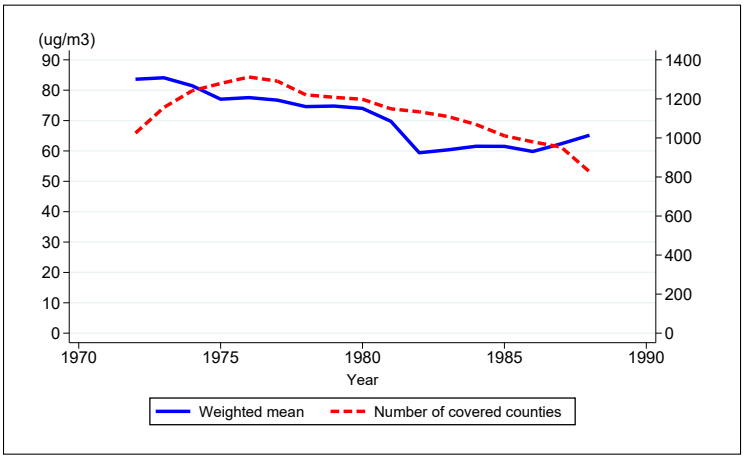
(c) NO_2



(d) O_3



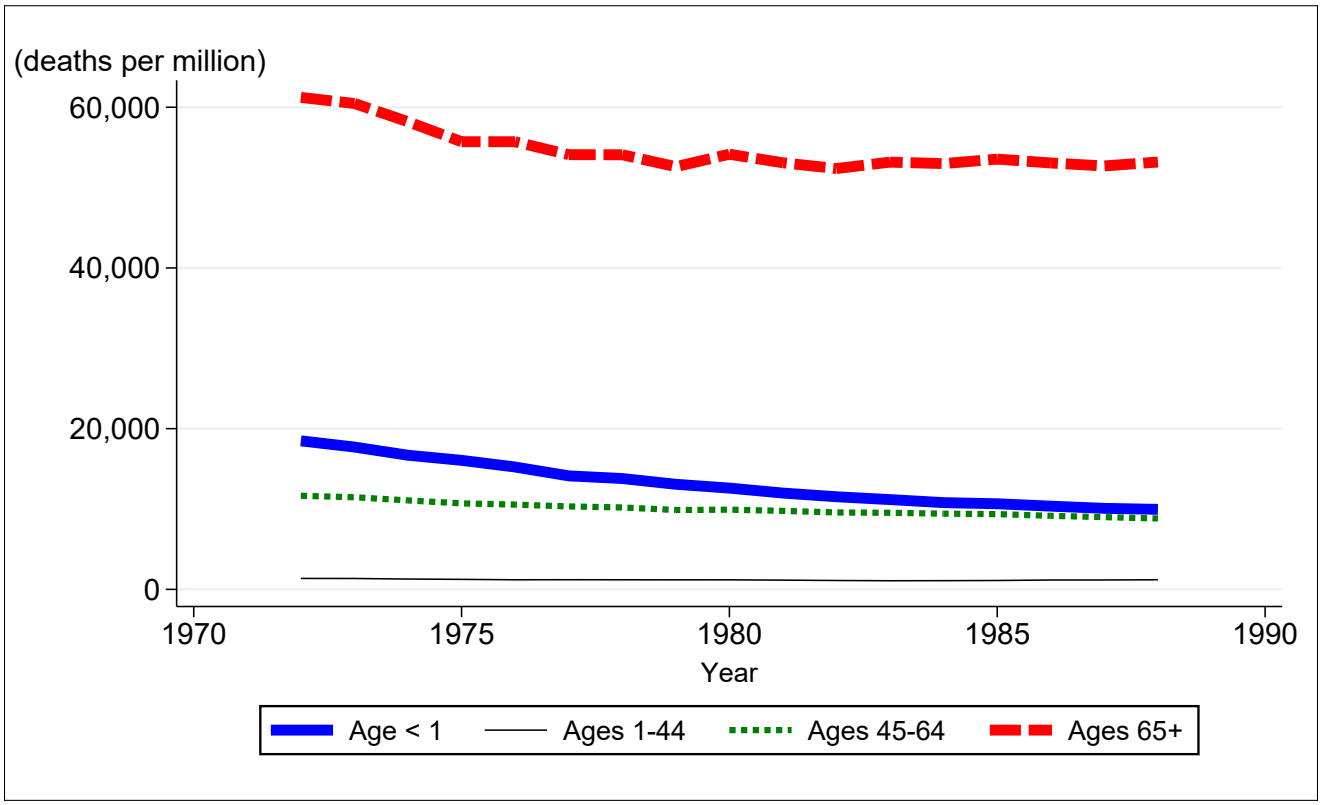
(e) TSP



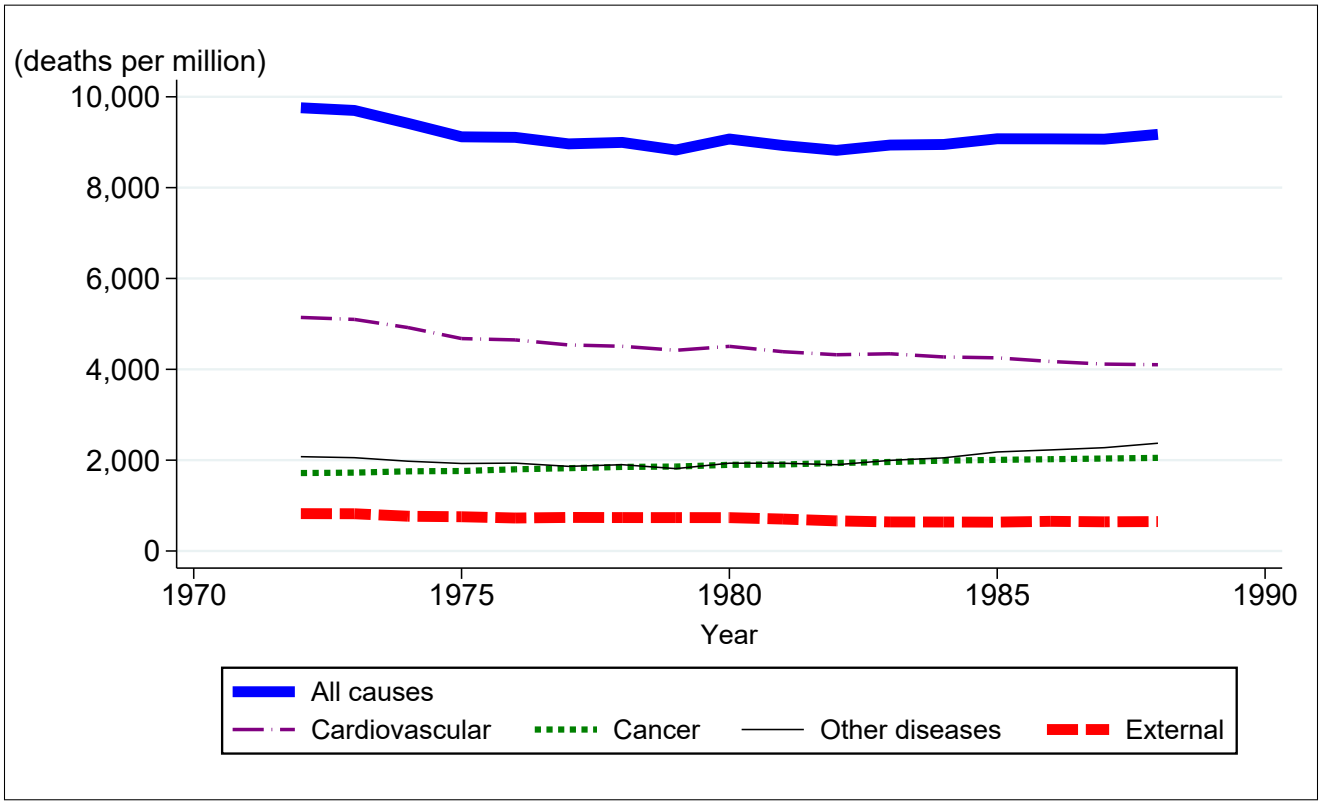
Notes: The solid blue lines report population-weighted pollution levels for all US counties with at least one daily reading for that pollutant. The dashed red line reports the number of counties with at least one operational monitor for the pollutant. Data are obtained from the EPA Air Quality database. SO_2, CO, NO_2, and O_3 are measured in parts per billion (ppb). Total suspended particulates (TSP) is measured in micrograms per cubic meter ($\mu g/m^3$).

Figure A.2: Trends in US mortality rates, 1972–1988

(a) By age group

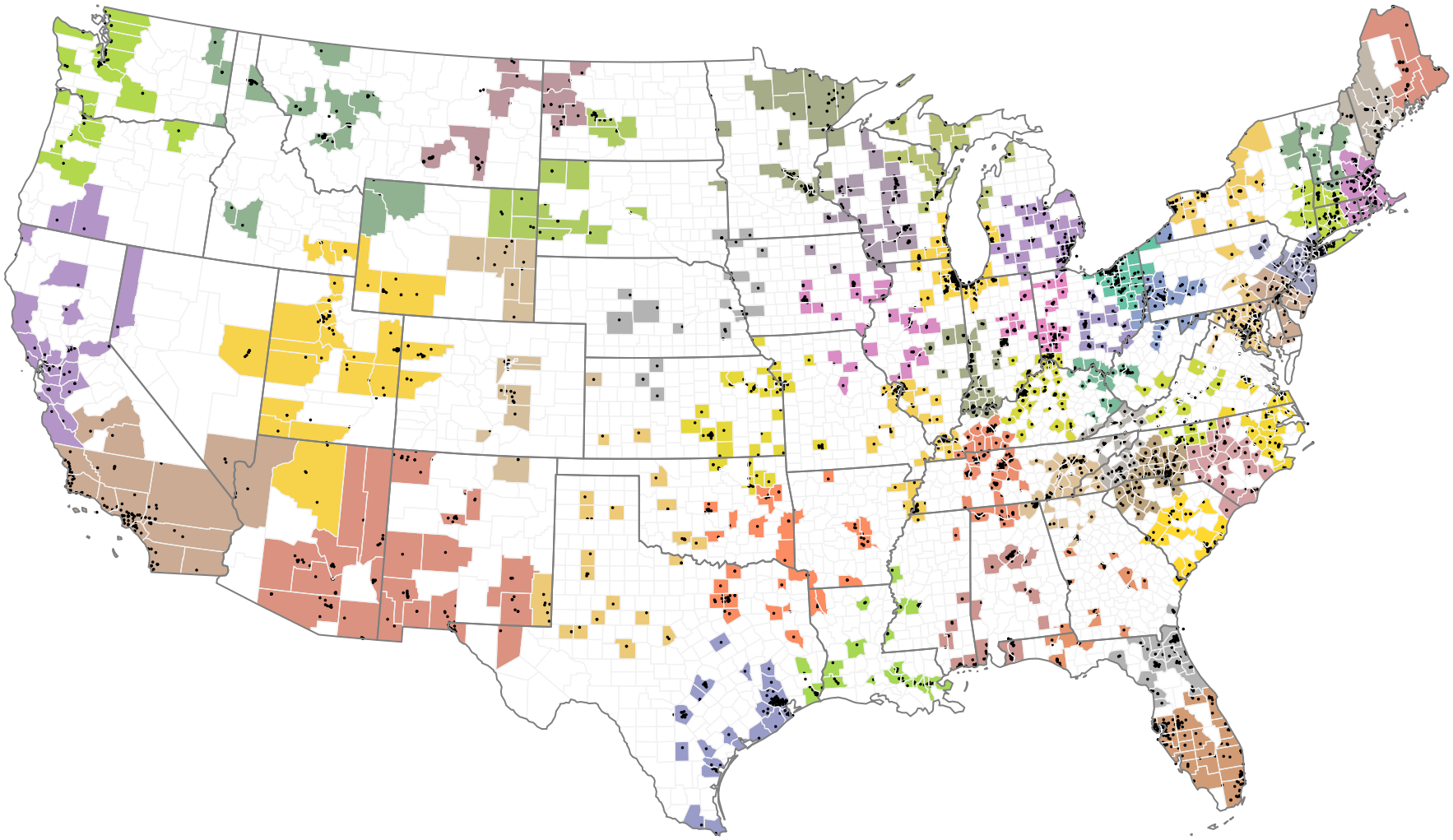


(b) By cause of death



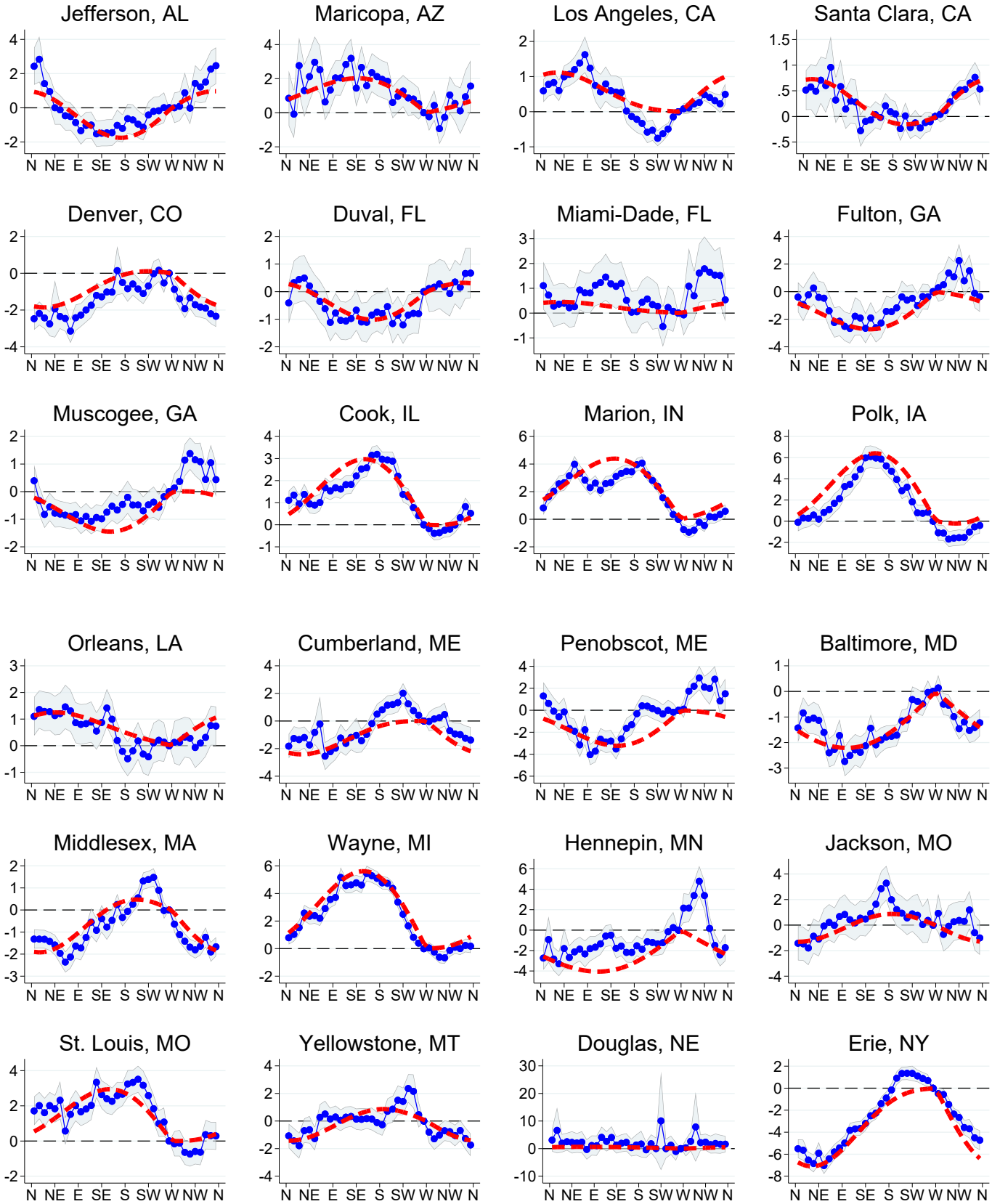
Notes: These two figures report annual mortality rates for the US population. These rates are calculated using mortality data from the National Vital Statistics and population data from SEER. Annual mortality rates are approximately 365 times larger than the daily mortality rates used in the analysis.

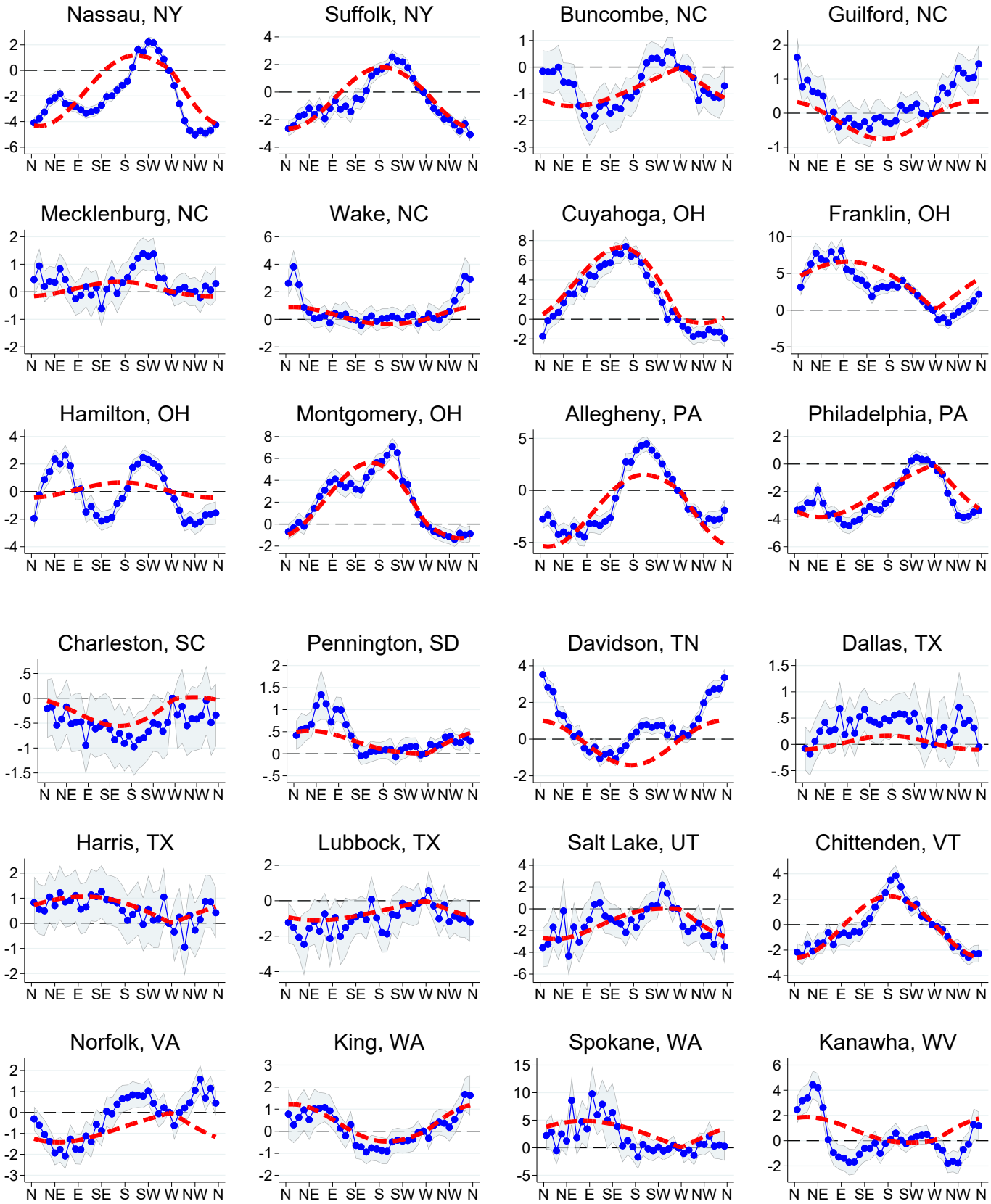
Figure A.3: Locations of the 50 geographic groups and SO_2 monitors

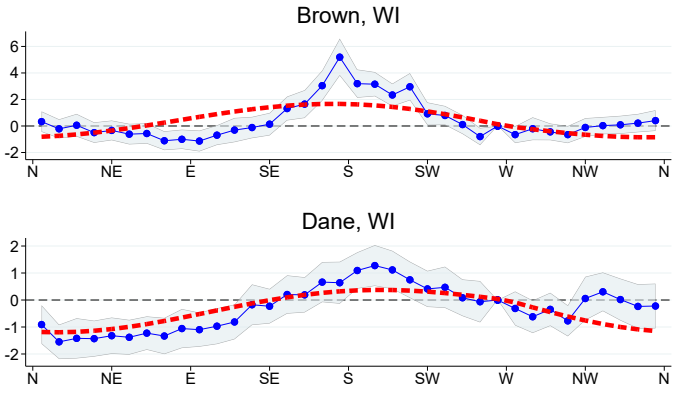


Notes: This map shows the 50 geographic groups included in our main estimation sample. Each group is shaded in a different color. The black dots represent the locations of the SO_2 monitors. Unshaded (white) counties have no SO_2 monitors and are not included in our sample. The Southern California and Greater Philadelphia groups are shown in detail in Figure 1. The first-stage effect of wind direction on air pollution is allowed to vary across geographic groups, as shown in Equation (2).

Figure A.4: The relationship between wind direction and SO_2 concentrations, by geographic group

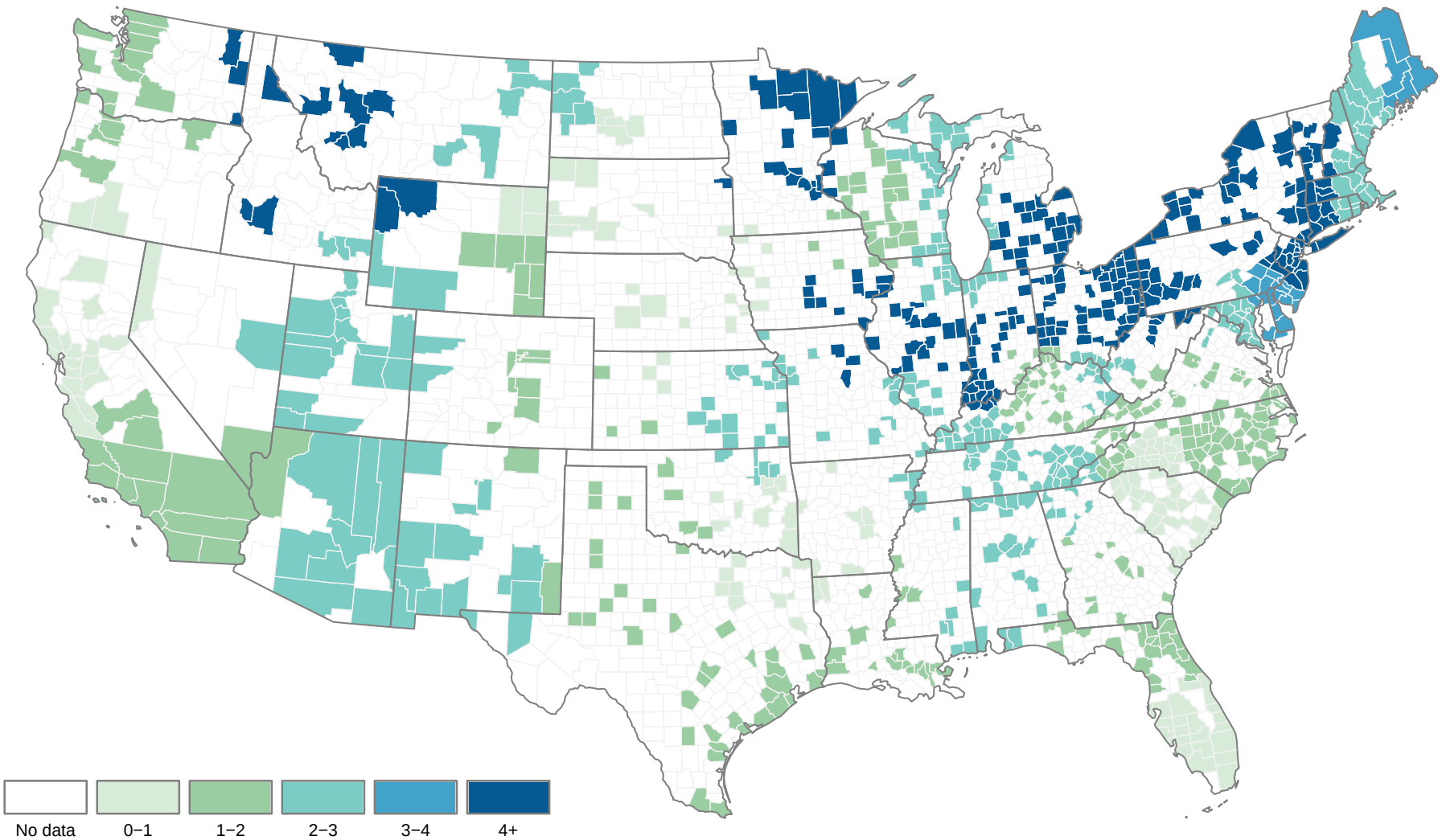



Nassau, NY
Suffolk, NY
Buncombe, NC
Guilford, NC
Mecklenburg, NC
Wake, NC
Cuyahoga, OH
Franklin, OH
Hamilton, OH
Montgomery, OH
Allegheny, PA
Philadelphia, PA
Charleston, SC
Pennington, SD
Davidson, TN
Dallas, TX
Harris, TX
Lubbock, TX
Salt Lake, UT
Chittenden, VT
Norfolk, VA
King, WA
Spokane, WA
Kanawha, WV
N NE E SE S SW W NW N



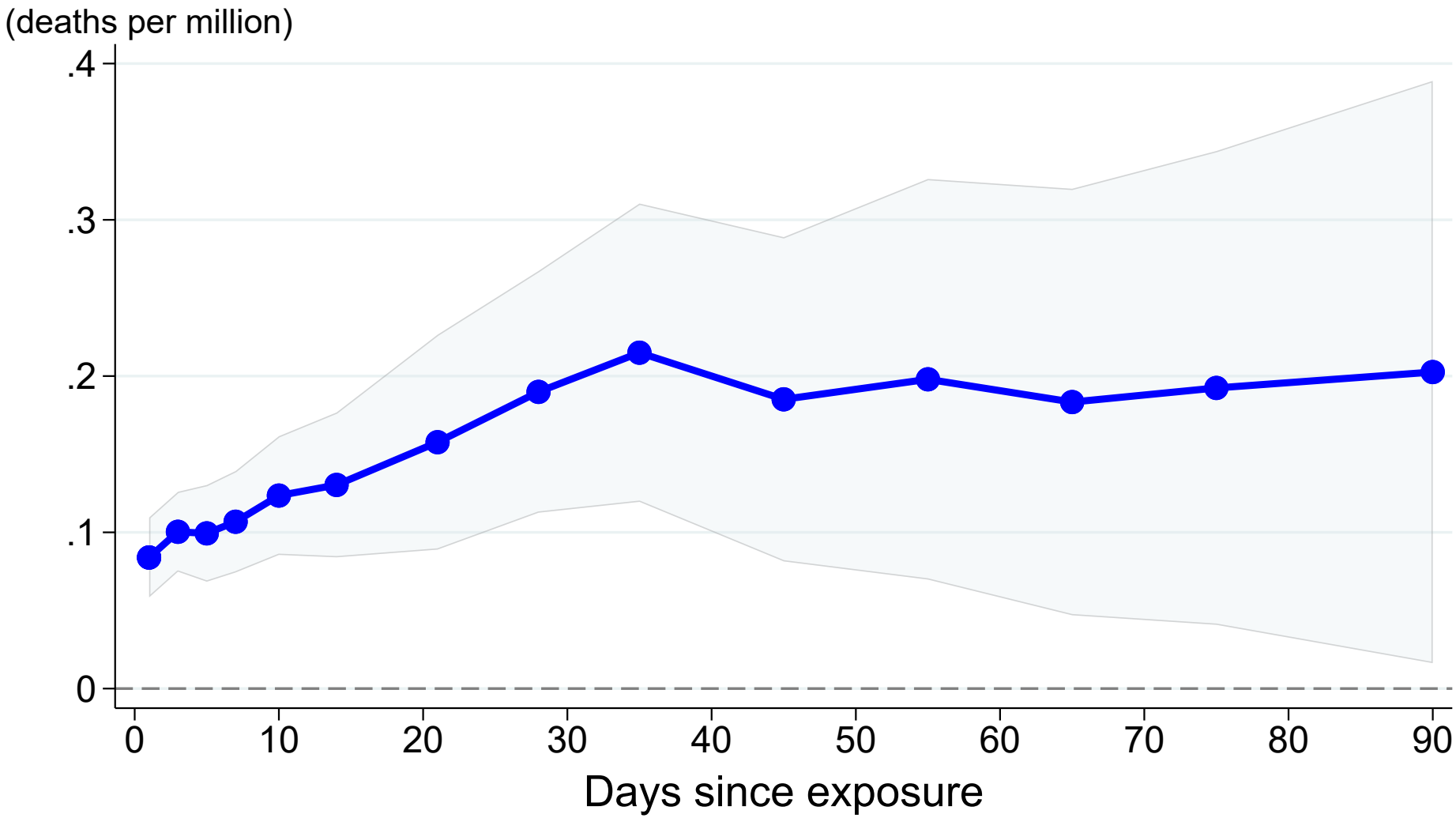
Notes: The graphs plot the relationship between SO_2 levels and windward direction in each area. Windward direction describes where the wind is blowing from, with "N" indicating North, "NE" indicating Northeast, etc. The 36 blue points report coefficient estimates from a non-parametric regression of SO_2 on wind direction measured in 10-degree angle bins. The blue shaded area shows the corresponding 95% confidence intervals. The red dashed lines report fitted curves from the parametric sine specification given by $f^g(\theta_{cd})$ in Equation (2). All regressions include county-by-month and month-by-year fixed effects, and flexible weather controls. Standard errors are robust to heteroskedasticity. The plots for "Philadelphia, PA" and "Los Angeles, CA" are reproduced as "Greater Philadelphia area" and "Southern California area" in Figure 1.

Figure A.5: Strength of the first stage, by geographic group



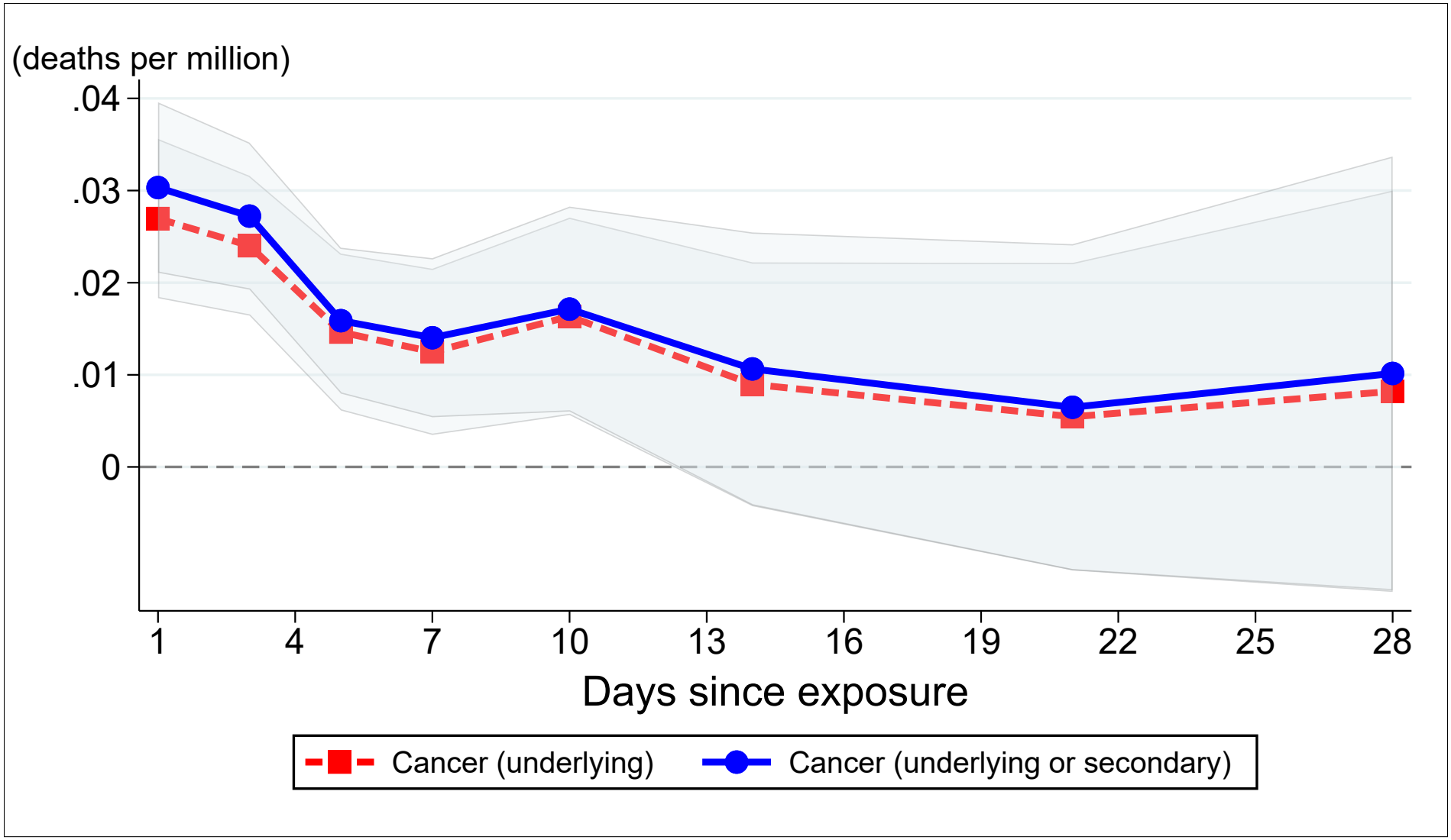
Notes: This map shows the strength of the first stage given by Equation (2), for each of the 50 geographic groups included in our main estimation sample. The strength is measured as the difference in predicted SO_2 levels (parts per billion) between the most and least polluting wind directions. Predictions are calculated using the parametric sine specification $\widehat{f^g}(\theta) = \widehat{\gamma_g^1} \sin(\theta) + \widehat{\gamma_g^2} \sin(\theta/2)$, for $\theta \in [0, 2\pi)$. Unshaded (white) counties have no SO_2 monitors and are not included in our sample.

Figure A.6: IV estimates of effect of acute (1-day) SO_2 exposure on cumulative mortality up to 90 days following exposure



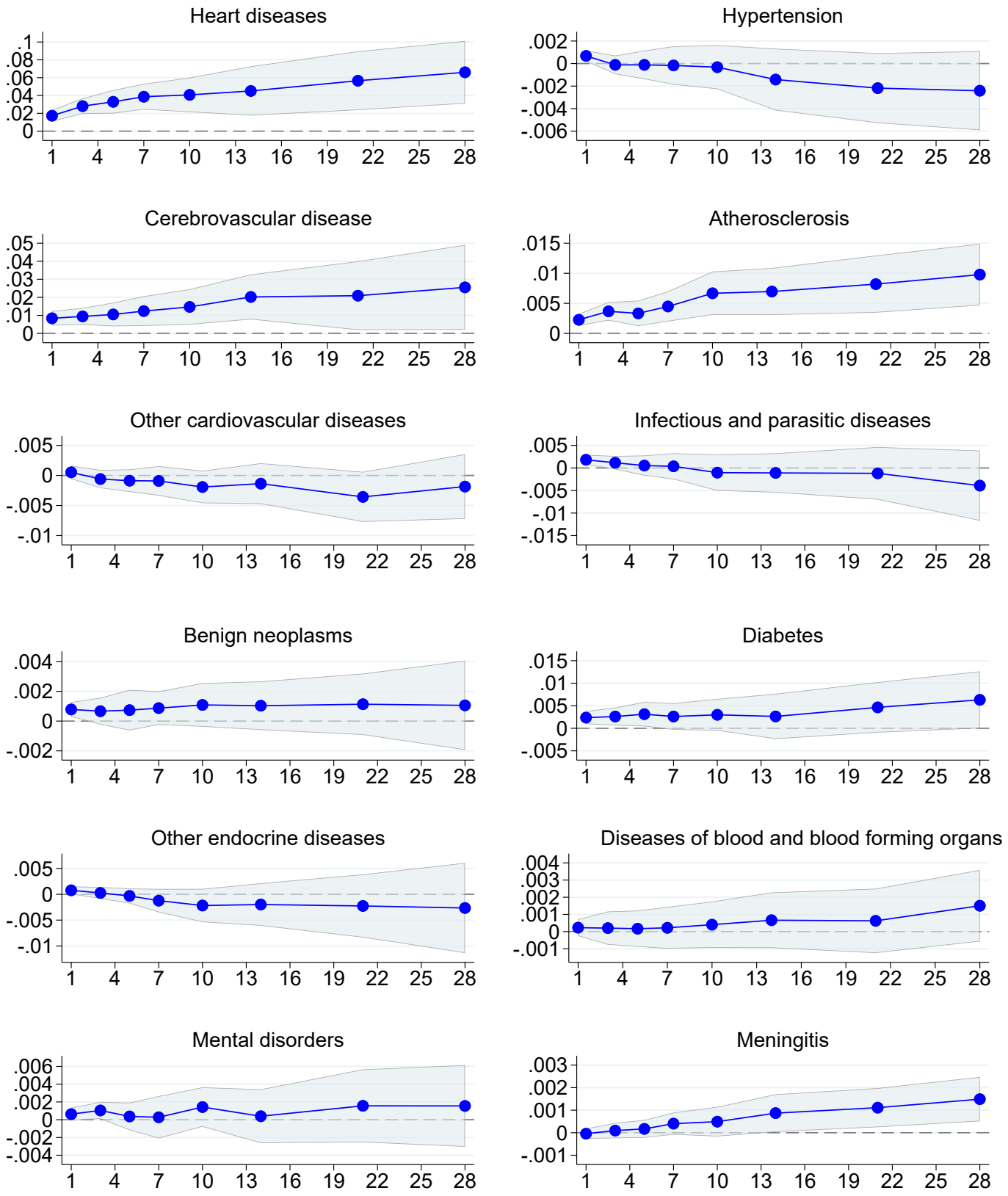
Notes: Each point reports an IV estimate from Equation (1) of the effect of acute (1-day), 1-ppb SO_2 exposure on mortality, where mortality is measured as cumulative deaths per million over a time window ranging from 1 to 28 days, as indicated by the x-axis. The shaded area represents 95% confidence intervals. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors are clustered by county.

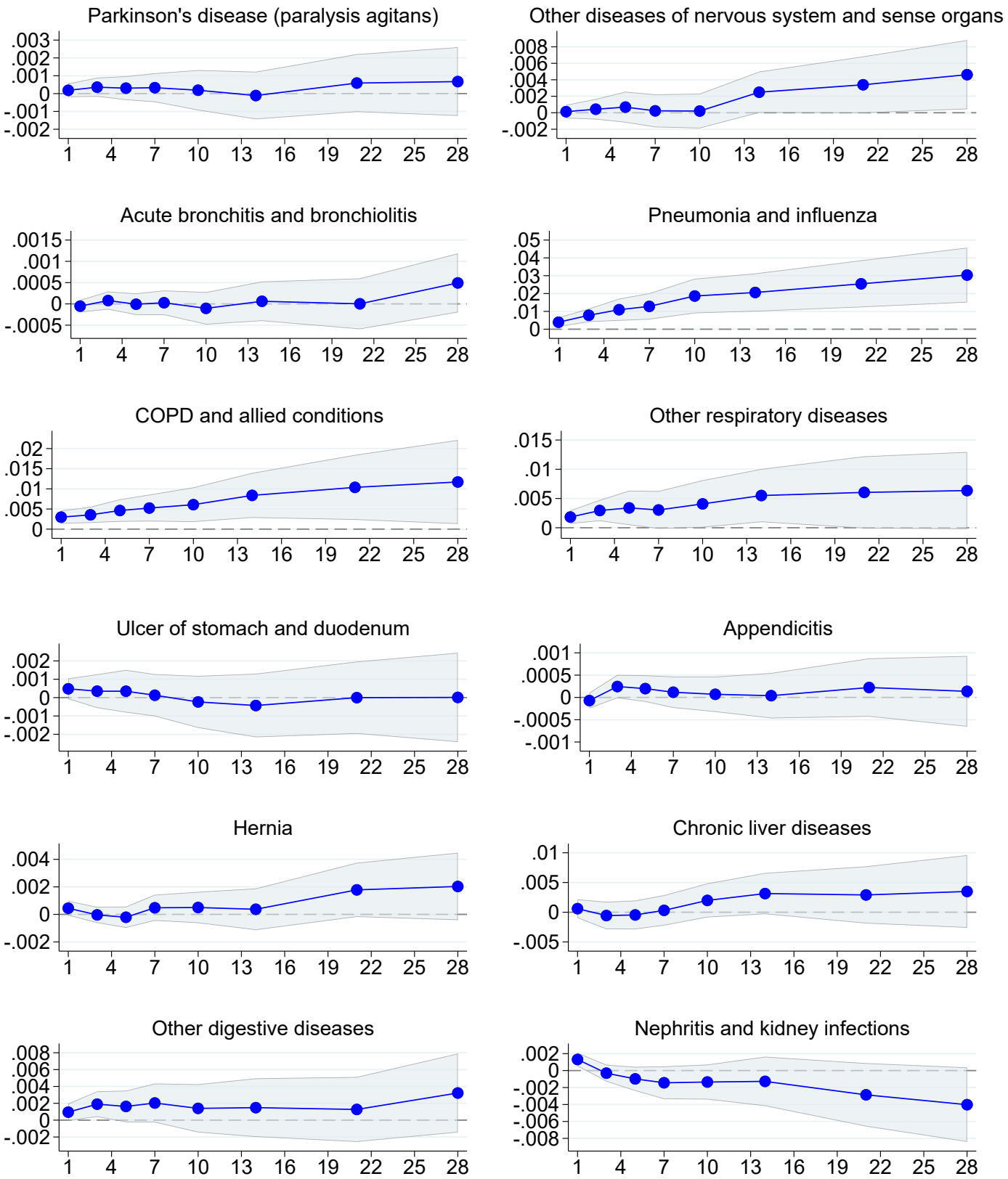
Figure A.7: IV estimates of effect of acute (1-day) SO_2 exposure on cancer-related mortality

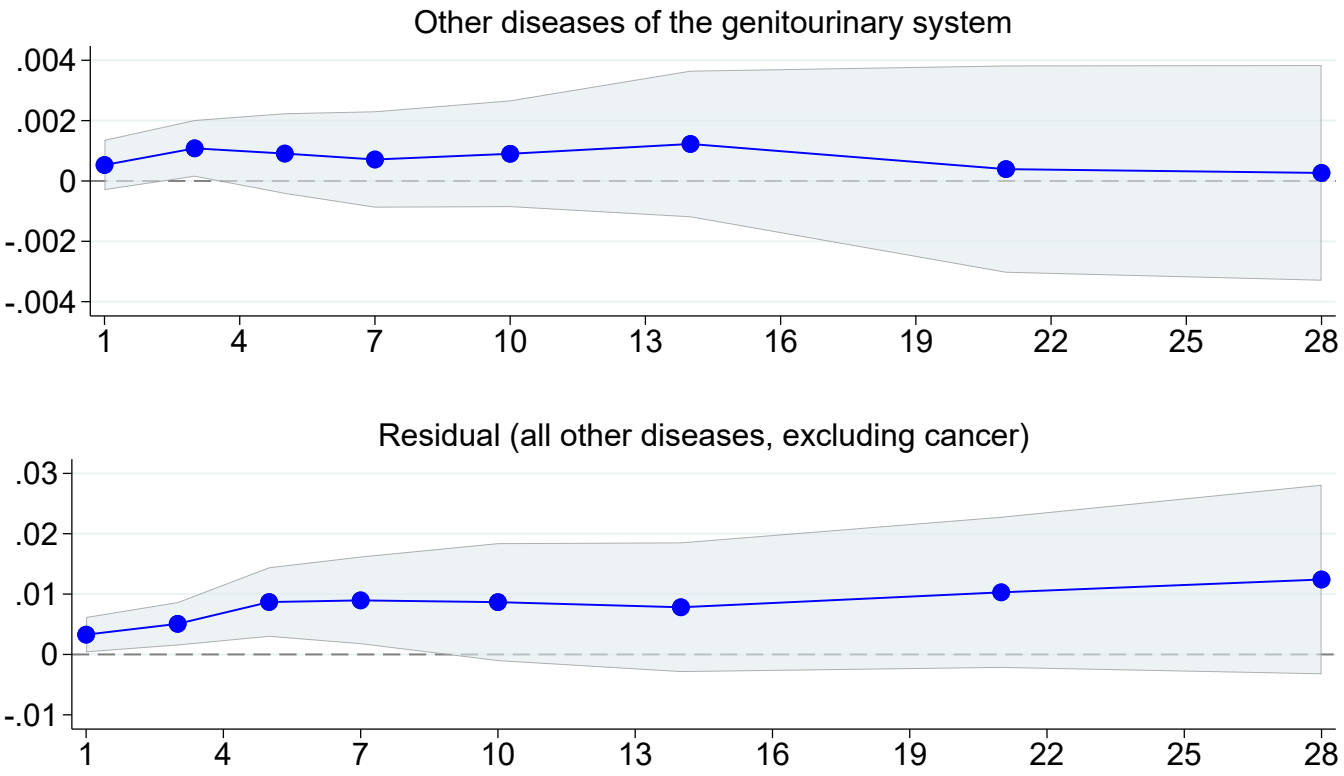


Notes: Each point reports an IV estimate from Equation (1) of the effect of acute (1-day) sulfur dioxide (SO_2) exposure on cancer-related mortality (deaths per million). "Cancer (underlying)", which replicates the cancer estimates shown in Figure 3, includes only deaths where cancer is listed as the underlying cause of death on the death certificate. "Cancer (underlying or secondary)" includes deaths where cancer is listed as the underlying cause of death or as one of the secondary causes. Mortality is measured with a time window ranging from 1 to 28 days, as indicated by the x-axis. Shaded areas represent 95% confidence intervals. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors are clustered by county.

Figure A.8: IV estimates of effect of acute (1-day) SO_2 exposure on mortality, by detailed cause of death



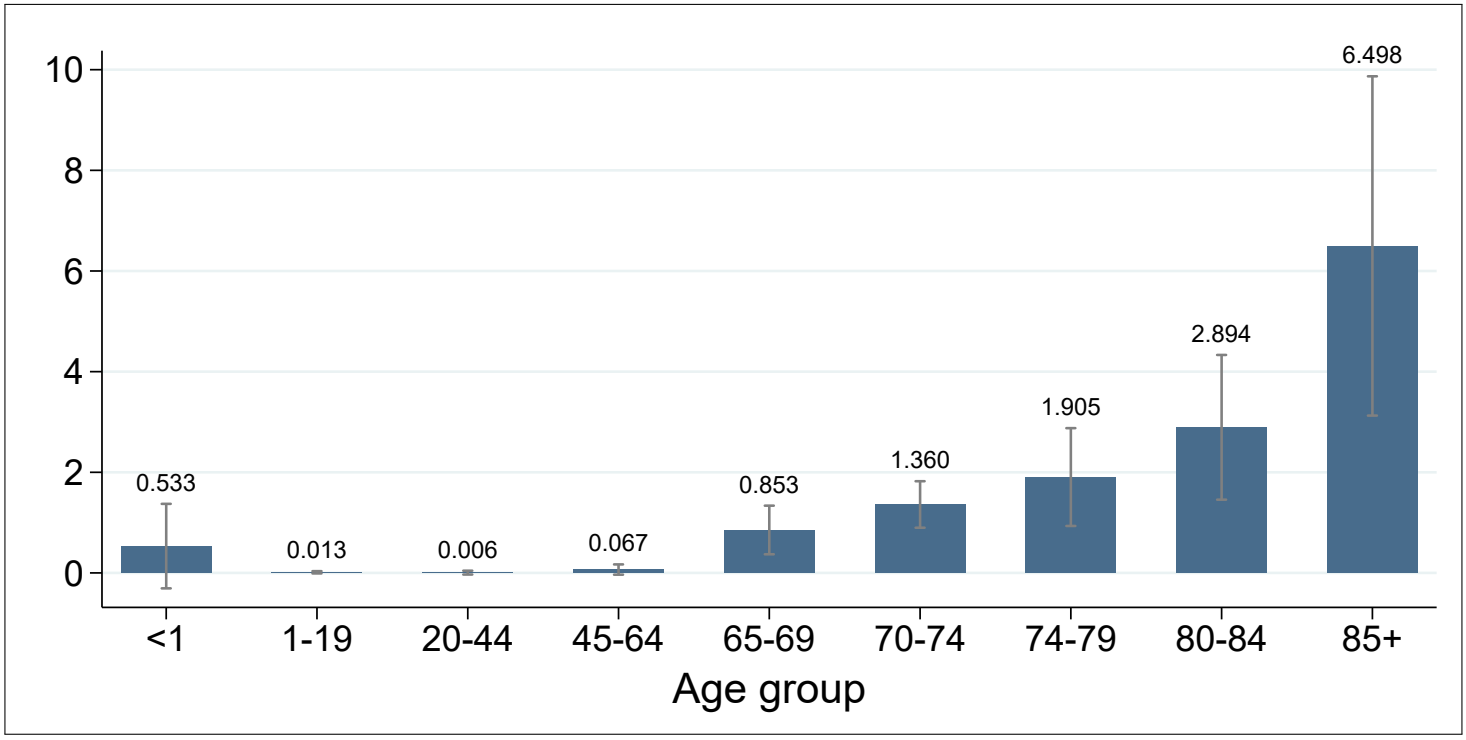

Parkinson's disease (paralysis agitans)
Other diseases of nervous system and sense organs
Acute bronchitis and bronchiolitis
Pneumonia and influenza
COPD and allied conditions
Other respiratory diseases
Ulcer of stomach and duodenum
Appendicitis
Hernia
Chronic liver diseases
Other digestive diseases
Nephritis and kidney infections



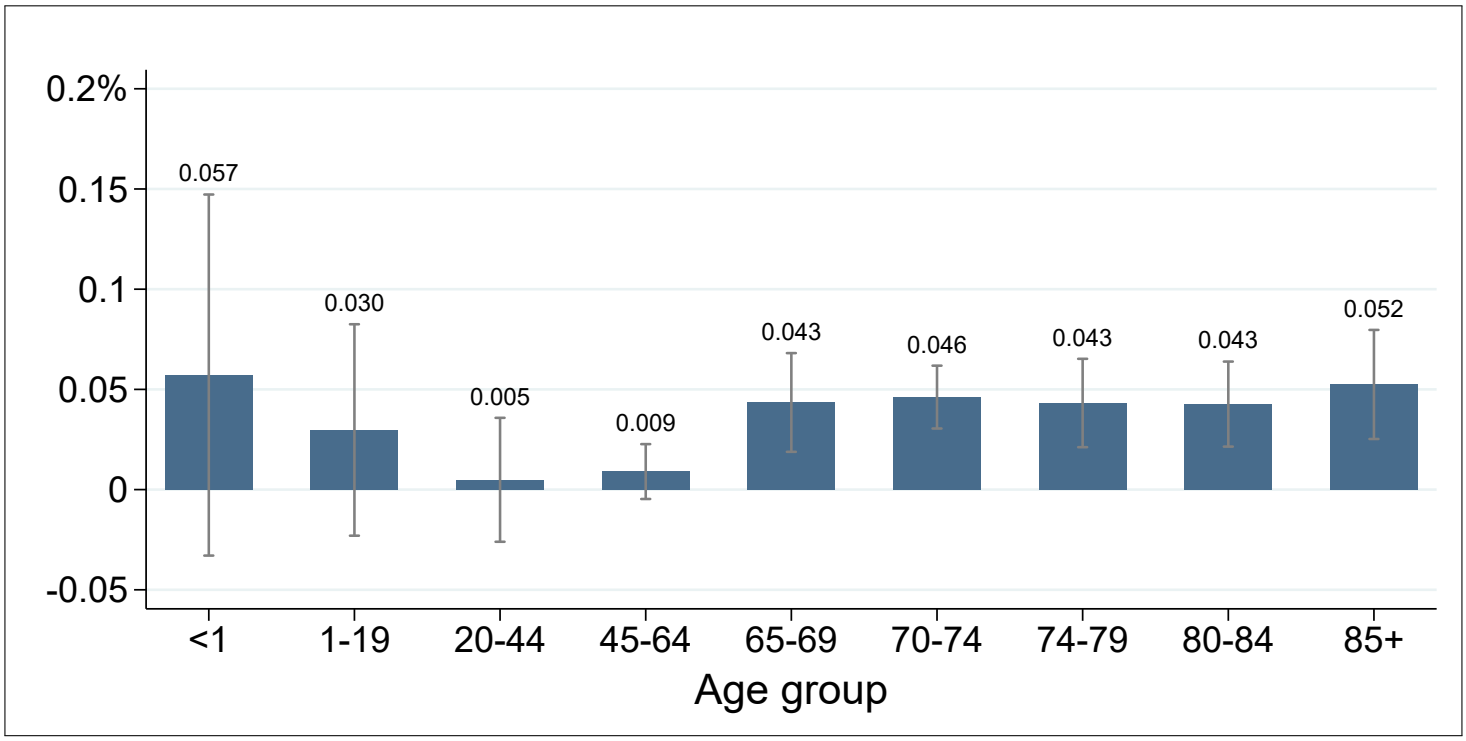
Notes: Each point reports an IV estimate from Equation (1) of the effect of acute (1-day) sulfur dioxide (SO_2) exposure on mortality (deaths per million), for twenty-six different causes of death. Definitions for each cause of death are available in Table A.13. Cumulative mortality is measured over a time window ranging from 1 to 28 days as indicated by the x-axis. Shaded areas represent 95% confidence intervals. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors are clustered by county.

Figure A.9: IV estimates of effect of acute (1-day) SO_2 exposure on 28-day mortality, by age group

(a) Absolute increase (deaths per million)

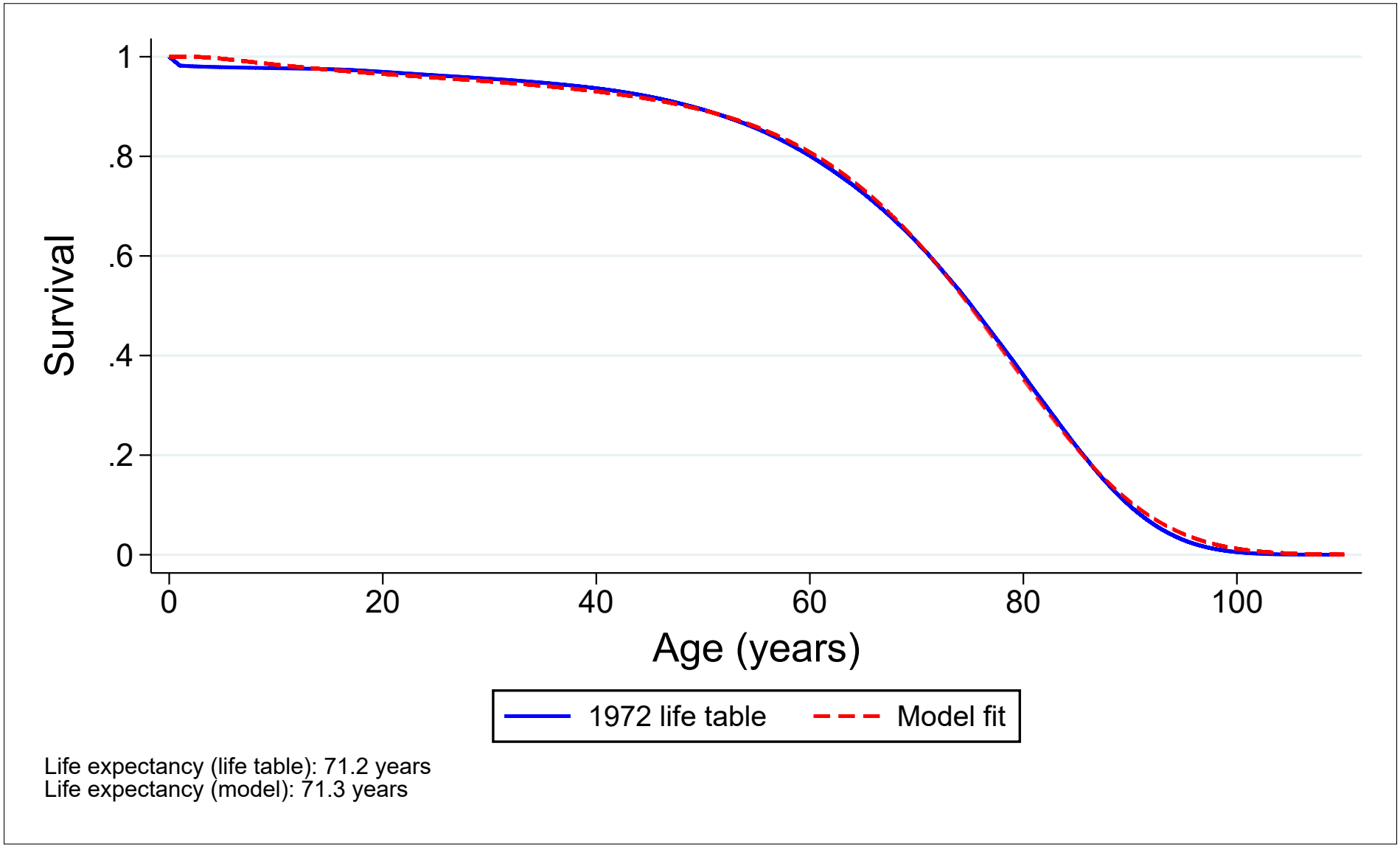


(b) Relative increase (percent of 28-day mortality)



Notes: Each bar represents an IV estimate from Equation (1) of the effect of acute (1-day) sulfur dioxide (SO_2) exposure on 28-day mortality for a particular age group. Error bars represent 95% confidence intervals. Estimates are also reported in Table A.4. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors are clustered by county.

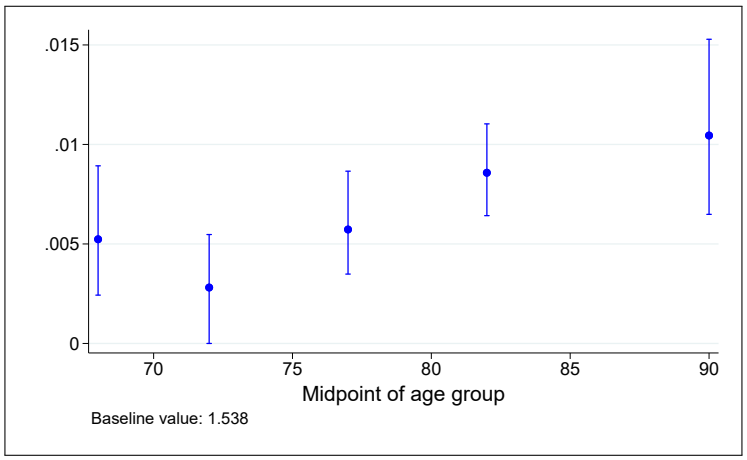
Figure A.10: Baseline calibration of the dynamic production model of health



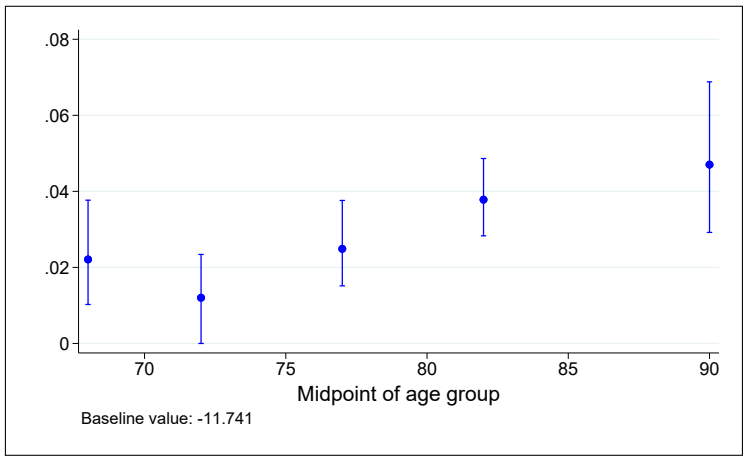
Notes: The solid blue line depicts the survival curve derived from the 1972 period life table for the United States. The dashed red line reports the predicted survival from our dynamic production model of health (3), which was calibrated using these 1972 data. The calibrated model parameters are reported in Column (2) of Table A.14.

Figure A.11: The effect of acute (1-day) SO_2 exposure on model parameters

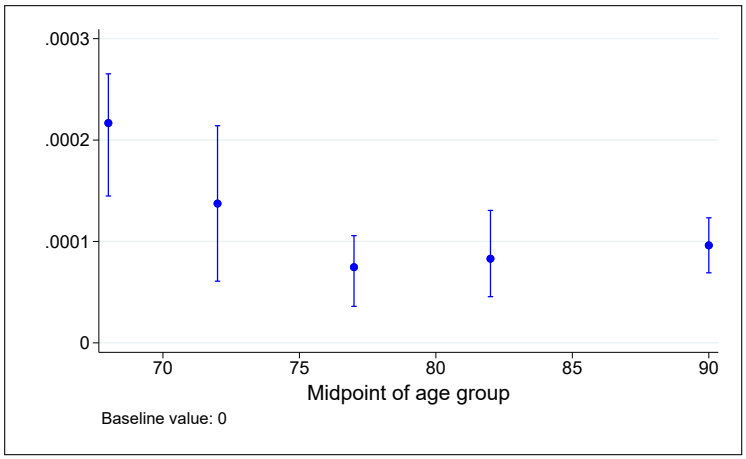
(a) Increase in α ($\widetilde{\alpha}_a - \alpha^*$)



(b) Increase in δ ($\widetilde{\delta}_a - \delta^*$)



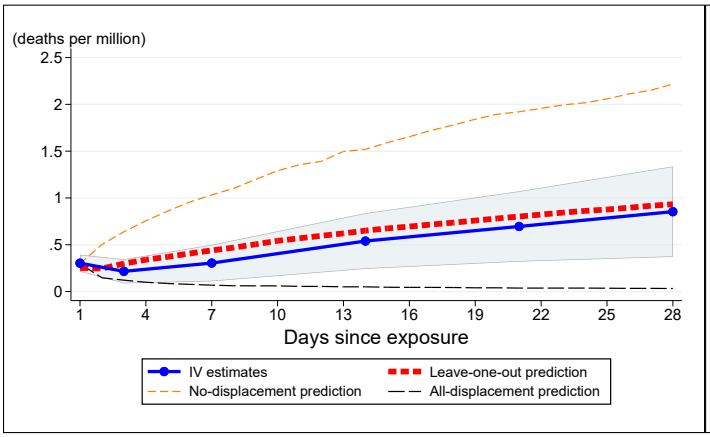
(c) Increase in $\underline{\mathrm{H}}$ ($\widetilde{\underline{\mathrm{H}}}_a - \underline{\mathrm{H}}^*$)



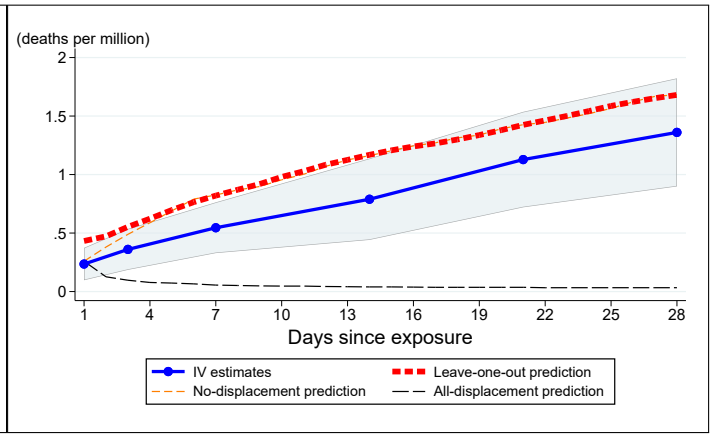
Notes: Panel (a) reports the effect of acute (1-day) pollution exposure on the change in the value of the parameter α from the health model (3), under the assumption that all non-cancer-related deaths operate through changes in α and all other deaths operate through changes in $\underline{\mathrm{H}}$. The baseline value is denoted by α^* and the post-exposure value following calibration using age group a is denoted by $\widetilde{\alpha}_a$. Panel (b) reports estimates under the alternative assumption that all non-cancer-related deaths operate through changes in the aging parameter δ rather than α. Panel (c) reports the effect on the change in the value of the parameter $\underline{\mathrm{H}}$, which governs mortality displacement and is calibrated using cancer-related deaths. The error bars report the 5th and 95th percentiles from a set of 100 bootstrap replications. Baseline values for all model parameters are reported in Table A.14.

Figure A.12: Comparison of model predictions to IV estimates, by age group

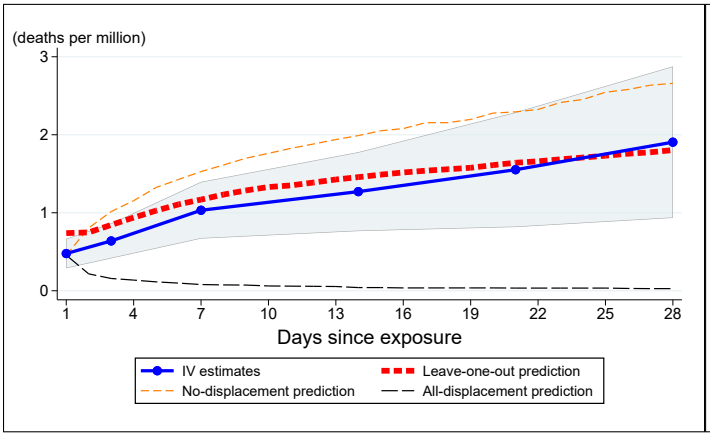
(a) Ages 65–69



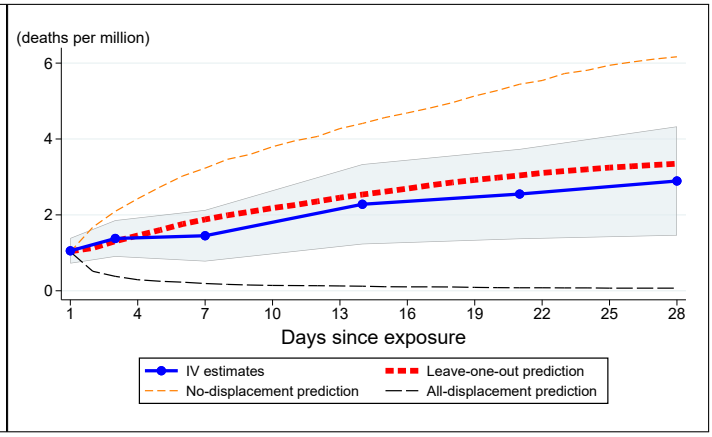
(b) Ages 70–74



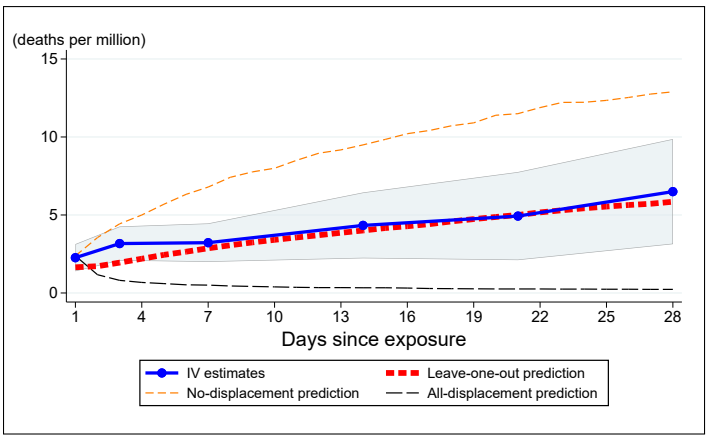
(c) Ages 75–79



(d) Ages 80–84

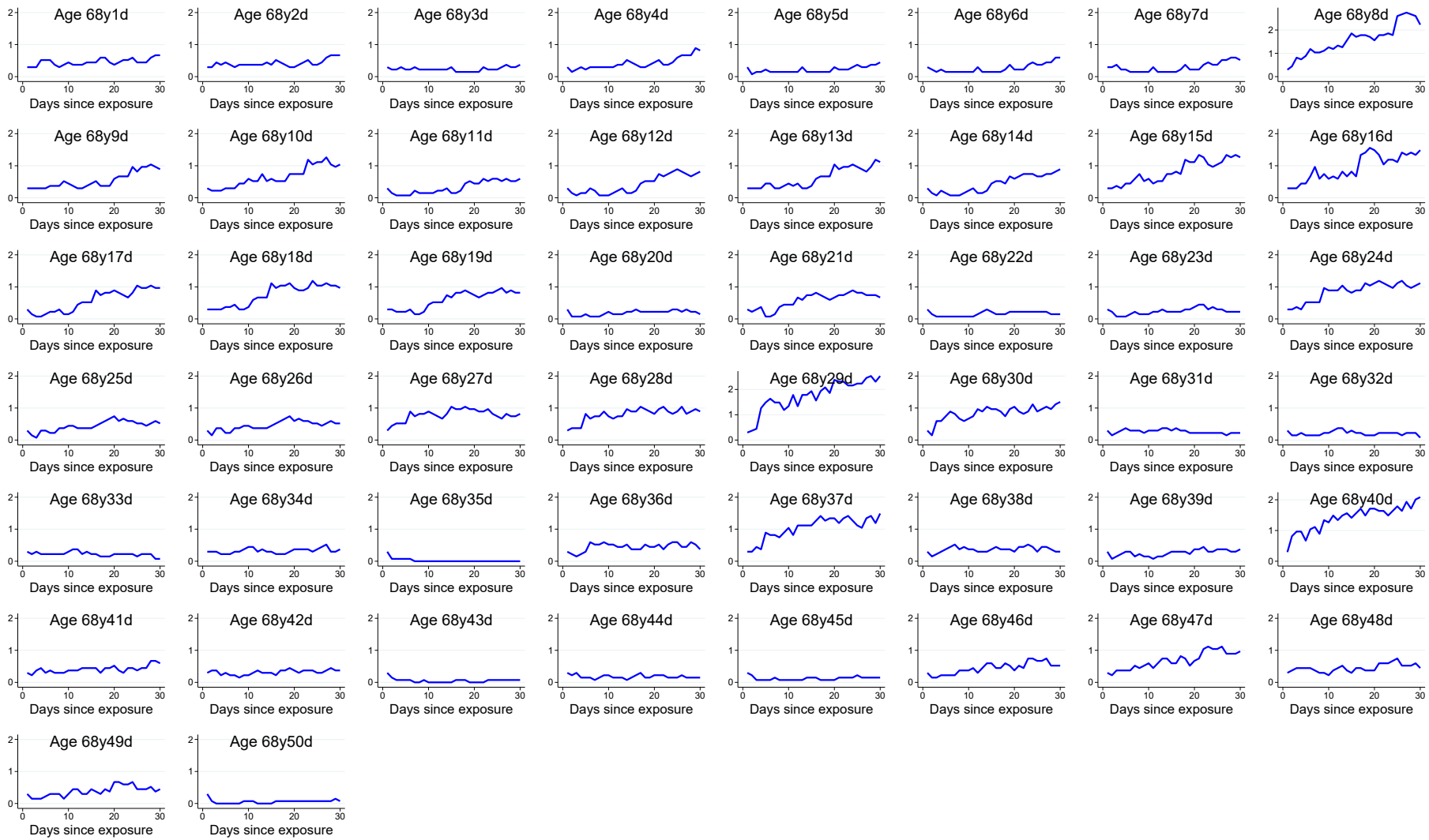


(e) Ages 85+



Notes: The solid blue line reports IV estimates from Equation (1) of the effect of acute (1-day) sulfur dioxide (SO_2) exposure on cumulative mortality, with 95% confidence intervals given by the blue shaded area. The thick red dashed line reports "leave-one-out" predictions from the dynamic production model of health described by Equation (3). The leave-one-out model employs the average of the calibrated values for all of the other age groups, e.g., for ages 65–69 it uses the average of the values for ages 70–74, 75–79, 80–84, and 85+. The orange dashed line ("no-displacement") reports model predictions under the extreme assumption that none of the 1-day mortality effect is mortality displacement, while the black dashed line ("all-displacement") reports predictions under the alternative extreme assumption that all of the 1-day effect is mortality displacement.

Figure A.13: Predicted effects of acute (1-day) SO_2 exposure on cumulative mortality, for selected ages



Notes: Each plot corresponds to a different age (68 years 1 day, 68 years 2 days, etc.) from the dynamic production model of health (3). Plots show the predicted effect of acute exposure on cumulative mortality (deaths per million) up to 30 days following exposure, using the calibration methodology described in Section 5. The first point of each plot (x = 1 day since exposure) is equal to the 1-day IV estimate for the 65–69 age group (Figure 4a). The subsequent values are model predictions. The average of these 50 plots corresponds to the green dot-dashed line ("own-age prediction") shown in Figure 5.

Table A.1: Complier county characteristics for the wind direction instrument

	Population	Percent 65+	Percent Black	Per-capita income (dollars)	Per-capita transfers (dollars)	Employment rate	Mean SO_2 (ppb)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
First-stage strength, ppb	4,138	0.048	–1.4**	84**	32**	–0.56**	1.4**
	(8,988)	(0.054)	(0.15)	(28)	(4.4)	(0.17)	(0.088)
Mean outcome	244,789	11	9.2	8,856	1,115	47	7.4
Sample size	9,427	9,427	9,328	9,286	9,286	9,286	9,427
R-squared	0.0099	0.061	0.069	0.77	0.78	0.047	0.30

Notes: This table reports an analysis of complier county characteristics for the wind direction instrument. Complier characteristics are estimated by regressing the county-level variable reported at the top of each column on the strength of the first stage. First-stage strength is measured as the difference in predicted SO_2 levels between the most and least polluting wind directions. All regressions include year fixed effects. County characteristics are obtained from the Regional Economic Information System (REIS) dataset published by the Bureau of Economic Analysis (BEA). Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table A.2: OLS and IV estimates of effect of acute SO_2 exposure on cumulative mortality, for different outcome windows

	(1)	(2)
Outcome window	OLS	IV
1 day	0.0077*	0.084**
	(0.0031)	(0.013)
3 days	0.021*	0.10**
	(0.0092)	(0.013)
5 days	0.029	0.099**
	(0.015)	(0.016)
7 days	0.031	0.11**
	(0.020)	(0.017)
10 days	0.035	0.12**
	(0.028)	(0.019)
14 days	0.042	0.13**
	(0.039)	(0.024)
21 days	0.049	0.16**
	(0.056)	(0.035)
28 days	0.046	0.19**
	(0.073)	(0.039)

Notes: Dependent variable is cumulative number of deaths per million people in the days following acute (1-day) exposure. Each estimate comes from a separate regression. IV estimates in Column (2) are shown in Figure 2. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table A.3: IV estimates of effect of acute SO_2 exposure on 1-day mortality, for different age groups

	(1)	(2)
Age group	Absolute effect, deaths per million	Relative effect, percent
0–1	0.064	0.19
	(0.060)	(0.18)
1–19	0.0027	0.18
	(0.0022)	(0.15)
20–44	0.016**	0.37**
	(0.0052)	(0.12)
45–64	0.056**	0.21**
	(0.013)	(0.048)
65–69	0.30**	0.43**
	(0.046)	(0.066)
70–74	0.24**	0.22**
	(0.071)	(0.067)
75–79	0.48**	0.30**
	(0.097)	(0.061)
80–84	1.1**	0.43**
	(0.17)	(0.071)
85+	2.3**	0.51**
	(0.44)	(0.10)

Notes: Dependent variable is number of deaths per million people on the day of exposure. Each estimate comes from a separate regression. Relative effect is calculated as the percent of the age group's mean one-day mortality rate. Estimates are also shown in panels (a) and (b) of Figure 4. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table A.4: IV estimates of effect of acute SO_2 exposure on cumulative mortality, for different age groups and outcome windows

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Window	0–1	1–19	20–44	45–64	65–69	70–74	75–79	80–84	85+
1 day	0.064	0.0027	0.016**	0.056**	0.30**	0.24**	0.48**	1.1**	2.3**
	(0.060)	(0.0022)	(0.0052)	(0.013)	(0.046)	(0.071)	(0.097)	(0.17)	(0.44)
3 days	0.21**	0.00068	0.021**	0.058**	0.22**	0.36**	0.64**	1.4**	3.2**
	(0.080)	(0.0033)	(0.0063)	(0.014)	(0.066)	(0.089)	(0.11)	(0.25)	(0.56)
7 days	0.31*	–0.0043	0.022**	0.025	0.30**	0.55**	1.0**	1.5**	3.2**
	(0.15)	(0.0061)	(0.0077)	(0.023)	(0.10)	(0.11)	(0.19)	(0.35)	(0.63)
14 days	0.27	–0.0071	0.0099	0.014	0.54**	0.79**	1.3**	2.3**	4.3**
	(0.27)	(0.0074)	(0.010)	(0.038)	(0.15)	(0.18)	(0.26)	(0.54)	(1.1)
28 days	0.53	0.013	0.0060	0.067	0.85**	1.4**	1.9**	2.9**	6.5**
	(0.43)	(0.011)	(0.019)	(0.052)	(0.25)	(0.24)	(0.50)	(0.73)	(1.7)

Notes: Dependent variable is cumulative number of deaths per million people in the days following acute (1-day) exposure. Each estimate comes from a separate regression. Age group is given at the top of each column. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table A.5: IV estimates of effect of acute SO_2 exposure on cumulative mortality, for different causes of death and outcome windows

	(1)	(2)	(3)	(4)
Outcome window	Cardio	Cancer	Other	External
1 day	0.029**	0.027**	0.023**	0.0047**
	(0.0045)	(0.0044)	(0.0046)	(0.0016)
3 days	0.040**	0.024**	0.029**	0.0070*
	(0.0054)	(0.0039)	(0.0045)	(0.0030)
5 days	0.046**	0.015**	0.035**	0.0041
	(0.0078)	(0.0043)	(0.0071)	(0.0032)
7 days	0.054**	0.012**	0.037**	0.0034
	(0.0089)	(0.0046)	(0.0077)	(0.0037)
10 days	0.060**	0.016**	0.044**	0.0032
	(0.011)	(0.0055)	(0.0097)	(0.0040)
14 days	0.069**	0.0090	0.052**	0.000048
	(0.015)	(0.0067)	(0.012)	(0.0054)
21 days	0.080**	0.0055	0.066**	0.0068
	(0.019)	(0.0085)	(0.015)	(0.0074)
28 days	0.097**	0.0082	0.077**	0.0074
	(0.020)	(0.011)	(0.019)	(0.0086)

Notes: Dependent variable is cumulative number of deaths per million people in the days following acute (1-day) exposure. Each estimate comes from a separate regression. Cause of death is given at the top of each column. Estimates are also shown in Figure 3. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table A.6: IV estimates of effect of acute SO_2 exposure on 1-day mortality, controlling for all pollutants except TSP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SO_2, ppb	0.085**	0.077**	0.087**	0.082**	0.077**	0.084**	0.078**
	(0.0085)	(0.012)	(0.010)	(0.0091)	(0.012)	(0.011)	(0.013)
NO_2, ppb		0.010			0.011		0.013
		(0.010)			(0.015)		(0.015)
O_3, ppb			–0.0060			–0.0053	–0.0093
			(0.020)			(0.020)	(0.020)
CO, ppm				0.097	–0.012	0.095	–0.038
				(0.13)	(0.20)	(0.14)	(0.20)
First-stage F-statistic	225	72	34	91	41	34	30
Mean outcome	26	26	26	26	26	26	26
Sample size	275,690	275,690	275,690	275,690	275,690	275,690	275,690

Notes: Dependent variable is number of deaths per million people on the day of exposure. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level. National means for all air pollutants are available in Table 1.

Table A.7: IV estimates of effect of acute SO_2 exposure on 1-day mortality, using different fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
SO_2, ppb	0.084**	0.076**	0.078**	0.081**	0.079**	0.077**
	(0.013)	(0.0094)	(0.0090)	(0.0096)	(0.0096)	(0.0094)
Fixed effects	county-month, month-year	county, year, month	county, state-year-month	county-year, state-month	county, year, state-month	county, month-year, state-month
First-stage F-statistic	526	467	536	598	504	508
Mean outcome	25	25	25	25	25	25
Sample size	2,023,369	2,023,369	2,023,345	2,023,350	2,023,369	2,023,369

Notes: Dependent variable is number of deaths per million people on the day of exposure. All regressions include flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table A.8: IV estimates of effect of acute SO_2 exposure on 1-day mortality, clustering standard errors at different levels

	(1)	(2)	(3)	(4)
SO_2, ppb	0.084**	0.084**	0.084**	0.084**
	(0.013)	(0.013)	(0.016)	(0.012)
Clustering level(s)	County	County, geographic group by year	Geographic group	State
First-stage F-statistic	526	526	526	526
Mean outcome	25	25	25	25
Sample size	2,023,369	2,023,369	2,023,369	2,023,369

Notes: Dependent variable is number of deaths per million people on the day of exposure. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered at the level(s) indicated in each column, are reported in parentheses. Geographic groups are shown in Figure A.3. A */** indicates significance at the 5%/1% level.

Table A.9: IV estimates of effect of acute SO_2 exposure on 1-day mortality, using different numbers of instrument leads and lags

	(1)	(2)	(3)	(4)	(5)	(6)
SO_2, ppb	0.084**	0.087**	0.090**	0.084**	0.083**	0.083**
	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.013)
# of instrument leads	2	0	1	4	4	6
# of instrument lags	2	0	1	4	6	6
First-stage F-statistic	526	761	569	525	524	524
Mean outcome	25	25	25	25	25	25
Sample size	2,023,369	2,023,369	2,023,369	2,020,342	2,017,398	2,017,398

Notes: Dependent variable is number of deaths per million people on the day of exposure. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table A.10: IV estimates of effect of acute SO_2 exposure on 1-day mortality, using different first-stage specifications

	(1)	(2)	(3)	(4)	(5)	(6)
SO_2, ppb	0.084**	0.097**	0.073**	0.080**	0.077**	0.080**
	(0.013)	(0.014)	(0.012)	(0.014)	(0.015)	(0.016)
Num. geo. groups	50	25	100	50	50	50
Wind angle spec	Sines	Sines	Sines	40-degree bins	60-degree bins	90-degree bins
Num. instruments	100	50	200	400	250	150
First-stage F	526	886	310	142	194	263
Mean outcome	25	25	25	25	25	25
Sample size	2,023,369	2,023,369	2,023,369	2,023,369	2,023,369	2,023,369

Notes: Dependent variable is number of deaths per million people on the day of exposure. Column (1) reports our main specification, which allows the effect of wind direction to vary across 50 different geographic groups and assumes that the effect of wind direction on pollution, $f^g(\cdot)$, follows the sine parameterization given by Equation (2). Columns (2)–(3) vary the number of geographic groups employed in the first stage. Columns (4)–(6) use different non-parametric parameterizations for $f^g(\cdot)$. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table A.11: 2SLS and LIML estimates of effect of acute SO_2 exposure on cumulative mortality, for different outcome windows

	(1)	(2)	(3)	(4)	(5)	(6)
SO_2, ppb	0.084**	0.084**	0.11**	0.10**	0.19**	0.20**
	(0.013)	(0.013)	(0.017)	(0.013)	(0.039)	(0.028)
IV method	2SLS	LIML	2SLS	LIML	2SLS	LIML
Outcome window (days)	1	1	7	7	28	28
First-stage F-statistic	526	752	524	745	512	729
Mean outcome	25	25	173	173	691	691
Sample size	2,023,369	2,022,046	2,023,369	2,022,046	2,023,369	2,022,046

Notes: Dependent variable is cumulative number of deaths per million people in the days following acute (1-day) exposure. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table A.12: Placebo and falsification tests of the effect of acute SO_2 exposure on mortality

	(1)	(2)	(3)	(4)
SO_2, ppb	–0.079	0.18	–0.040	
	(0.062)	(0.23)	(0.49)	
SO_2 on day $t+1$, ppb				–0.0033
				(0.0048)
Outcome window, days	1	7	28	1
First-stage F-statistic	2.0	1.9	1.9	528
Mean outcome	25	173	691	25
Sample size	2,023,369	2,023,369	2,023,369	1,846,528
Placebo test	X	X	X	
Falsification test				X

Notes: Dependent variable is cumulative number of deaths per million people in the days following acute (1-day) exposure. Columns (1)–(3) report results of placebo regressions where the wind direction instrument is a randomly generated variable. Column (4) reports a falsification test of the effect of SO_2 exposure on the previous day's mortality rate. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table A.13: ICD-8 and ICD-9 codes for subcategories of cardiovascular and other diseases

ID	Disease	ICD-8 codes (1968–1978)	ICD-9 codes (1979–1998)	Prevalence (%)
	Cardiovascular diseases			
1	Heart diseases	390–398.9, 402–402.9, 404–429.9, 410–429.9	390–398.9, 402–402.9, 404–429.9	37.58
2	Hypertension	400–400.9, 401–401.9, 403–403.9	401–401.9, 403–403.9	0.37
3	Cerebrovascular disease	Same as ICD-9	430–438.9	8.69
4	Atherosclerosis	Same as ICD-9	440–440.9	1.38
5	Other cardiovascular diseases	Same as ICD-9	441–448.9	1.17
	Other diseases			
6	Infectious and parasitic diseases	Same as ICD-9	001–139.9	1.13
7	Benign neoplasms	Same as ICD-9	210–239.9	0.29
	Endocrine, nutritional and metabolic diseases, and immunity disorders			
8	Diabetes	Same as ICD-9	250–250.9	1.81
9	Other endocrine, nutritional and metabolic diseases, and immunity disorders	Same as ICD-9	240–249.0, 260–279.9	0.57
10	Diseases of blood and blood forming organs	Same as ICD-9	280–289.9	0.32
11	Mental disorders	Same as ICD-9	290–319	0.71
	Diseases of the nervous system and sense organs			
12	Meningitis	Same as ICD-9	320–322.9	0.08
13	Parkinson's disease (paralysis agitans)	342	332–332.1	0.21
14	Other diseases of nervous system and sense organs	320–341.9, 343–389.9	323–331.9, 332.2–389.9	0.89
	Diseases of the respiratory system			
15	Acute bronchitis and bronchiolitis	Same as ICD-9	466–466.9	0.03
16	Pneumonia and influenza	470–474.9, 480–486.9	480–487.9	2.97
17	COPD and allied conditions	490–493.9	490–496.9	2.52
18	Other respiratory diseases	460–469.9, 475–479.9, 487–489.9, 494–519.9	460–465.9, 467–479.9, 488–489.9, 497–519.9	1.25
	Diseases of the digestive system			
19	Ulcer of stomach and duodenum	Same as ICD-9	531–533.9	0.33
20	Appendicitis	Same as ICD-9	540–543.9	0.03
21	Hernia	Same as ICD-9	550–553.9, 560–560.9	0.29
22	Chronic liver diseases	Same as ICD-9	571–571.9	1.48
23	Other digestive diseases	Same as ICD-9	520–530.9, 534–539.9, 544–549.9, 554–559.9, 561–570.9, 572–579.9	1.54
	Diseases of the genitourinary system			
24	Nephritis and kidney infections	Same as ICD-9	580–590.9	0.91
25	Other diseases of the genitourinary system	Same as ICD-9	591–629.9	0.64
26	Residual: Complications of pregnancy, childbirth, and the puerperium; diseases of veins and lymphatics, and other diseases of circulatory system; diseases of the skin and subcutaneous tissue; diseases of the musculoskeletal system and connective tissue; congenital anomalies; certain conditions originating in the perinatal period; and ill-defined conditions	Same as ICD-9	450–459.9, 630–799.9	4.28

Notes: This table lists the ICD-8 and ICD-9 codes that define the different causes of death shown in Figure A.8. Death certificates used ICD-8 codes during the years 1968–1978, and used ICD-9 codes during the years 1979–1998. ICD-8 codes are shown in the table only when they differ from the ICD-9 codes. Prevalence reports the number of deaths from the disease during 1972–1988 as a fraction of total deaths.

Table A.14: Baseline parameter values for the dynamic production model of health

	(1)	(2)
Parameter	Annual data	Daily data
I	0.74773	0.0020521
α	1.53762	1.537619
$\ln \delta$	–5.83878	–11.74124
μ_H	10.39737	11.43803
σ_e	2.25247	0.1178985
N	1,000,000	100,000
SSE	57.80479	20880.82

Notes: This table reports baseline parameter values for the dynamic production model of health given by Equation (3). Column (1) reports parameter estimates when the model is fitted to annual survival data from a 1972 period life table. Column (2) reports corresponding estimates for daily data. The parameters $\underline{H}$ and σ_H are normalized to 0 and 1, respectively. N is the number of individuals in the simulation, and SSE is the sum of squared errors. The resulting fit for Column (2) is shown in Figure A.10. The mortality data underlying the life table are counts, not samples. Because there is no sampling error, we do not report standard errors for these parameter estimates.

Table A.15: IV estimates of effect of acute SO_2 exposure on all-cause and cancer-related 1-day mortality, ages 65 and over

	(1)	(2)
Age group	All causes	Cancer-related causes
65–69	0.30**	0.17**
	(0.046)	(0.028)
70–74	0.24**	0.14**
	(0.071)	(0.034)
75–79	0.48**	0.13**
	(0.097)	(0.039)
80–84	1.1**	0.18**
	(0.17)	(0.066)
85+	2.3**	0.17*
	(0.44)	(0.085)

Notes: These estimates are used to calibrate the effect of air pollution exposure on mortality in the dynamic production model of health given by Equation 3. Dependent variable is number of deaths per million people on the day of exposure. Each estimate comes from a separate regression. All regressions include county-by-month and month-by-year fixed effects, as well as flexible controls for maximum temperature, precipitation, and wind speed; leads of these weather controls; and two leads and two lags of the instruments. Estimates are weighted by the county population. Standard errors, clustered by county, are reported in parentheses. A */** indicates significance at the 5%/1% level.

Table A.16: Projected effect of permanent change in SO_2 on survival gains (years)

	(1)	(2)	(3)
	IV extrapolation	Aging model 1 (α)	Aging model 2 (δ)
1-ppb decrease	0.17	1.18	1.32
	[–0.02, 0.36]	[0.35, 2.21]	[0.38, 2.51]
2-ppb decrease	0.34	2.41	2.67
	[–0.04, 0.73]	[0.67, 4.89]	[0.73, 5.56]
3-ppb decrease	0.51	3.76	4.12
	[–0.06, 1.10]	[0.95, 7.55]	[1.04, 8.50]
1-ppb increase	–0.17	–1.05	–1.20
	[–0.36, 0.02]	[–1.99, –0.31]	[–2.33, –0.34]
2-ppb increase	–0.33	–2.15	–2.46
	[–0.71, 0.04]	[–3.77, –0.68]	[–4.43, –0.76]
3-ppb increase	–0.50	–3.12	–3.58
	[–1.05, 0.06]	[–5.30, –0.94]	[–6.22, –1.04]

Notes: Each value in this table reports the projected change in life expectancy (years) caused by a permanent change in SO_2 of up to 3 part per billion (ppb) for the cohort of US individuals born in 1972. Baseline life expectancy in 1972 is 71.32 years. 90% bootstrap confidence intervals, representing the range from the 5th to 95th percentiles of the bootstrapped statistics, are reported in brackets. Values in Column (1) are calculated by extrapolating our age-specific 28-day IV estimates to the whole life-cycle. Values in Columns (2) and (3) come from the dynamic production model of health described by Equation (3). Column (2) reports projections under the assumption that non-cancer-related pollution deaths are governed by changes in the model's depreciation parameter α. Column (3) reports projections under the alternative assumption that they are governed by the depreciation parameter δ. Figure 7 shows how the survival gains reported in the first row are distributed across the life cycle.