Estimating the Health Effects of Environmental Exposures: Statistical Methods for the Analysis of Spatio-temporal Data

Estimating the Health Effects of Environmental Exposures: Statistical Methods for the Analysis of Spatio-temporal Data

A dissertation presented

by

Andrew William Correia

to

The Department of Biostatistics

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Biostatistics

Harvard University
Cambridge, Massachusetts

April 2013

Advisor: Professor Francesca Dominici Andrew William Correia

Estimating the Health Effects of Environmental Exposures: Statistical Methods for the Analysis of Spatio-temporal Data

Abstract

In the field of environmental epidemiology, there is a great deal of care required in constructing models that accurately estimate the effects of environmental exposures on human health. This is because the nature of the data that is available to researchers to estimate these effects is almost always observational in nature, making it difficult to adequately control for all potential confounders - both measured and unmeasured. Here, we tackle three different problems in which the goal is to accurately estimate the effect of an environmental exposure on various health outcomes.

In Chapter 1, we extend and expand upon a previous study examining the relationship between fine particle air pollution and life expectancy in the United States (US) by analyzing data from the period 2000 to 2007 from 545 counties across the US. Using straightforward regression techniques, we estimate the association between changes in air pollution levels and changes in life expectancy over the period from 2000 to 2007 for the entire US as well as for a number of subpopulations within the US.

Chapter 2 builds upon the previous chapter by developing a modeling approach for estimating the effects of monthly variations in fine particle air pollution on monthly variations in mortality while controlling for potential sources of confounding. We first show via a simulation study where previous approaches to estimating this relationship break down. We then propose a new model to overcome those deficiencies, and we evaluate this approach using a large Medicare dataset linked with air pollution exposure estimates from across the US.

In Chapter 3, we evaluate the impact of noise exposure from airports on hospitalizations for cardiovascular disease (CVD) among Medicare enrollees living in zip codes surrounding major airports in the continental US. We begin with a fully Bayesian hierarchical Poisson model for the expected number of CVD hospitalizations in each zip code as a function of exposure to noise as well as several other individual and area-level covariates. We then conduct a thorough sensitivity analysis, examining potential sources of confounding, spatial dependence, and the possibility of a threshold effect.

Contents

Acknowledgments

I would like to thank my advisor, Francesca Dominici, for being an excellent mentor, and for her support, positive energy, and commitment to not only me, but to everyone in our group.

I would also like to thank my committee members Jonathan Levy and Brent Coull for their guidance and feedback, as well as Arden Pope, Majid Ezzati, and Doug Dockery who all played a significant role in my development as a graduate student. I feel quite fortunate to have had a chance to work with each these very talented and dedicated people.

I am extremely grateful to my parents, Gil and Gabriela Correia, whose love and support helped me to become the person I am today. Thank you for always encouraging me, for always pushing me to do better, and for never letting me quit anything I started.

Lastly, I would like to thank my beautiful bride to be, Andrea Saunders. Thank you for always being there, no matter what.

Dedicated to my Dad.

The Effect of Air Pollution Control on Life Expectancy in the United States: An Analysis of 545 US counties for the period 2000 to 2007

Andrew W. Correia[1], C. Arden Pope III[2], Douglas W. Dockery[3], Yun Wang[1], Majid Ezzati[4], and Francesca Dominici[1]

[1] Department of Biostatistics, Harvard University

[2] Department of Economics, Brigham Young University

[3] Departments of Environmental Health and Epidemiology, Harvard School of Public Health

[4] MRC-HPA Centre for Environment and Health and Department of Epidemiology and Biostatistics, Imperial College London

1.1 Introduction

Since the 1970s, enactment of increasingly stringent air quality controls has led to improvements in ambient air quality in the United States at costs that the U.S. Environmental Protection Agency (EPA) has estimated as high as $25 billion per year {United States Environmental Protection Agency (1997)}. However, even with the well-established link between long-term exposure to air pollution and adverse effects on health {Pope (2007)}, the extent to which more recent regulatory actions have benefited public health remains in question.

Air pollutant concentrations have been generally decreasing in the U.S., with substantial differences in reductions across metropolitan areas. Levels of fine particulate matter air pollution (particulate matter $< 2.5\mu\mathrm{g}/\mathrm{m}^3$ in aerodynamic diameter, $\mathrm{PM}_{2.5}$) remain relatively high in some areas. In a 2010 study, the EPA estimated that 62 U.S. counties, accounting for 26% of their total study population, had $\mathrm{PM}_{2.5}$ concentrations not in compliance with the National Ambient Air Quality Standards (NAAQS) {Schmidt et al. (2010)}.

Reductions in particulate matter air pollution are associated with reductions in both cardiopulmonary and overall mortality {Pope (2007)}. In the mid-1990s, the Harvard Six Cities Study and the American Cancer Society (ACS) study reported associations of cardiopulmonary mortality risk with chronic exposure to fine particulate air pollution while controlling for smoking and other individual risk factors {Dockery et al. (1993); Pope et al. (1995)}. Reanalysis and extended analyses of these studies have confirmed that fine particulate air pollution is an important independent environmental risk factor for cardiopulmonary disease and mortality {Krewski et al. (2000); Pope et al. (2002, 2004); Jerrett et al. (2005); Laden et al. (2006); Krewski et al. (2005b,a)}. Additional cohort studies, population-based studies, and short-term time-series studies have also shown associations between reductions in air pollution and reductions in human mortality {Burnett et al. (2001); Samet et al. (2000); Schwartz et al. (2008); Evans et al. (1984); Özkaynak and Thurston (1987); Pope and Dockery (2006); Schwartz (1991, 1992); Dominici et al. (2003)}.

More recently, studies have suggested an association between $PM_{2.5}$ and life expectancy {Tainio et al. (2007); Pope et al. (2009)}, a well-documented and important measure of overall public health {Brunekreef (1997); McMichael et al. (1998); Rabl (2003)}.

As our primary analysis, we estimate the association between changes in $PM_{2.5}$ and in life expectancy in 545 U.S. counties during the period 2000 to 2007. This period is of particular interest, as the EPA restarted wide collection of $PM_{2.5}$ data in 1999 - 2000, after stopping the nationwide $PM_{2.5}$ monitoring program during the mid-1980s and most of the 1990s. In secondary analyses, we extended the data and statistical analysis originally reported by Pope et al. (2009) for the period 1980 - 2000 to 2007, and investigated whether the relationship reported by Pope et al. (2009) persists in the more recent years.

1.2 Methods

1.2.1 Data

We constructed and analyzed three data sets to estimate the association between changes in life expectancy and changes in $PM_{2.5}$ during the period 2000 to 2007 in 545 counties (Dataset 1), and to investigate whether the association previously reported by Pope et al. (2009) persists when the data on the same 211 counties are extended to the year 2007 (Datasets 2 and 3).

Dataset 1 included information on 545 U.S. counties for the years 2000 and 2007. These counties include all counties with available matching $PM_{2.5}$ data for 2000 and 2007. Additionally, unlike previous work in which counties were located only in metropolitan areas Pope et al. (2009), Dataset 1 is comprised of counties in both metropolitan and non-metropolitan areas. Figure 1.1 shows the counties in this dataset shaded according to life expectancy in 2000 and 2007. Variables in this dataset were available at the county level, for both 2000 and 2007, and included: life expectancy, $PM_{2.5}$, per capita income, population, proportions who were high school graduates, and propor-

tions who were white, black, or Hispanic. Because data on smoking prevalence were not available for all 545 counties, we used age-standardized death rates for lung cancer and chronic obstructive pulmonary disease (COPD) as proxy variables for smoking prevalence {Peto et al. (1992); Eftim et al. (2008)}. Death rates were calculated in 5-year age groups and age-standardized for the 2000 U.S. population of adults 45 years of age or older. Daily $\mathrm{PM}_{2.5}$ data were obtained from the EPA's Air Quality System (AQS - *http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm*). Daily $\mathrm{PM}_{2.5}$ levels for each county were averaged across monitors within that county using a trimmed mean approach; those daily county-level means were further averaged across days to obtain a county-specific yearly $\mathrm{PM}_{2.5}$ average {Peng and Dominici (2008)}.

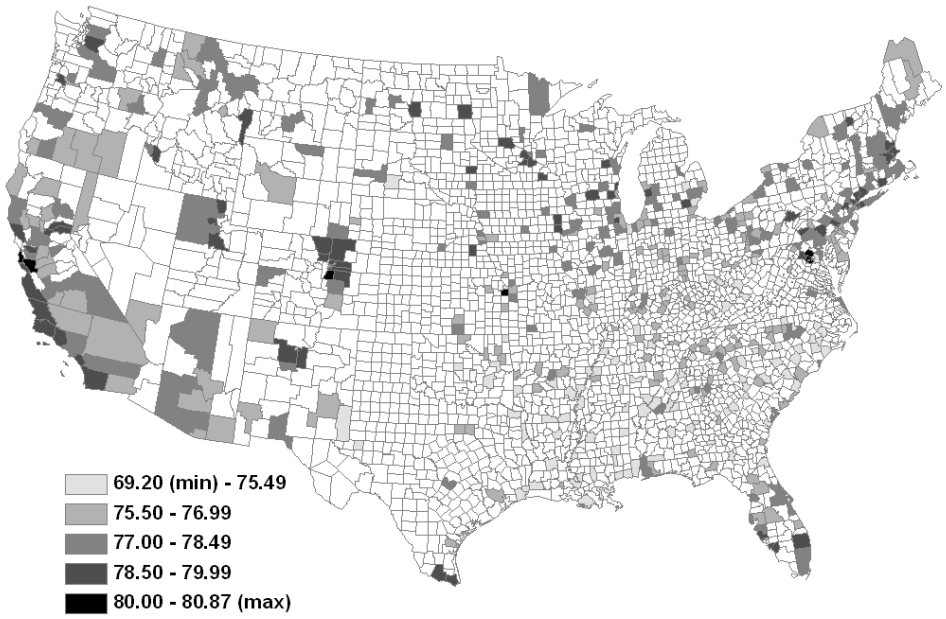
County-level life expectancies were calculated by applying a mixed-effects spatial Poisson model to mortality data from the National Center for Health Statistics (NCHS) and population data from the U.S. Census to obtain robust estimates of the number of deaths in each county {Kulkarni et al. (2011)}. These estimated counts were then used to calculate county life expectancies using standard life table techniques, which we discuss in more detail in the eAppendix (Section A).

Socioeconomic and demographic variables were obtained from the U.S. Census and the American Community Survey except per capita income, which was obtained from the Bureau of Economic Analysis. All yearly income variables were adjusted for inflation with 2000 as the base year. Age-standardized death rates for lung cancer and COPD were calculated using mortality data from NCHS using death rates for 2005 to serve as a proxy for 2007 (NCHS data for 2007 was not readily available). Lastly, data on smoking prevalence (proportion of the population who are current smokers) were available from the Behavioral Risk Factor Surveillance System in both 2000 and 2007 for 383 of the 545 counties.

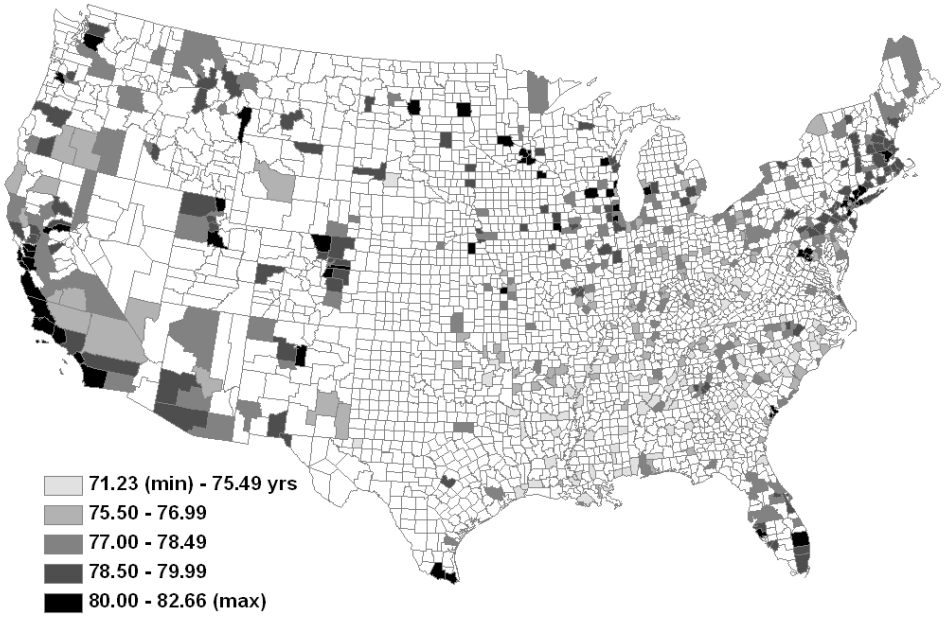
Dataset 2 included data for the year 1980 and the year 2000 for the same 211 U.S. counties included in the 51 metropolitan statistical areas (MSAs) previously analyzed by Pope et al. (2009). This dataset is identical to that in the paper by Pope et al. (2009) where

Figure 1.1: **United States County Maps Shaded by Life Expectancy:** Maps of the US with the 545 counties from Dataset 1 shaded according to life expectancy for the years 2000 (a) and 2007 (b).

Figure 1.1 (Continued)



(a) yr. 2000 county life expectancies



(b) yr. 2007 county life expectancies

it is described in more detail.

Dataset 3 extended Dataset 2 to 2007. All data were available at the county level except for $PM_{2.5}$, which for the year 1980 was available only at the MSA level and for the year 2007 was available at the county level for only 113 of the 211 counties originally included in Pope et al. (2009). Thus, for the year 2007, we assigned the same $PM_{2.5}$ values to all the counties that shared an MSA, consistent with the previous analysis by Pope et al. (2009). Details and results pertaining to Datasets 2 and 3 are summarized in the eAppendix (Section B1).

1.2.2 Statistical Analysis

Cross-sectional and first-difference linear regression models were fitted to all three datasets. Specifically, we regressed life expectancy versus $PM_{2.5}$ levels across counties separately for the years 1980 (Dataset 2), 2000 (Datasets 1 and 2), and 2007 (Datasets 1 and 3). We then regressed changes in life expectancy over the years 2000 to 2007 (Datasets 1 and 3), 1980 to 2000 (Dataset 2), and 1980 to 2007 (Dataset 3) versus changes in $PM_{2.5}$ over those same periods adjusted for changes in the socioeconomic, demographic, and proxy smoking variables outlined above. Additionally for our largest dataset (Dataset 1: 545 counties, 2000 to 2007), we also performed several stratified and weighted analyses. More specifically, we estimated the effect of changes in $PM_{2.5}$ on life expectancy in models stratified by: 1) percentage of the population with an urban residence in 2000; 2) population density in 2000; 3), land area in 2000; 4) $PM_{2.5}$ levels in 2000; 5) 5-year in-migration in 2000; and 6) change in average yearly temperature over the entire period. These stratified analyses allowed us to examine whether $PM_{2.5}$ effects on life expectancy were different in counties with particular demographic or weather characteristics. The sensitivity of our results to model specification was further assessed by fitting models weighted by: 1) total population; 2) year 2000 population density; and 3) inverse land area. We included direct measures of the change in prevalence of smoking for the subgroup of counties with matching data on smoking prevalence (383 out of 545), and

fit separate models for men and women to determine if effects differed by sex. To account for the correlation due to clustering of counties in the same MSA, robust clustered standard errors were calculated for all models {Pope et al. (2009); Diggle et al. (1994)}. Specifically, the variance of the vector of estimated regression coefficients, $\hat{\beta}$, is given by: $\mathrm{Var}(\hat{\beta}) = \left(X^T X\right)^{-1} \left(X^T \hat{V} X\right) \left(X^T X\right)^{-1}$, where $\hat{V}$ is a block-diagonal matrix with non-zero blocks $V_{0,j} = (y_j - \hat{\mu}_j)(y_j - \hat{\mu}_j)^T$, where j indexes the MSAs, y_j is the vector of observed outcomes in MSA j, and $\hat{\mu}_j$ is the vector of fitted values from a standard ordinary least squares (OLS) regression for MSA j. $\hat{\beta}$ is equal to the OLS estimator. Models were estimated using either REGRESS in Stata version 11.0, `lm()` in R version 2.11.1, or PROC SURVEYREG in SAS version 9.2.

1.3 Results

We report the results of our primary analysis, which estimated the cross-sectional relationship between life expectancy and PM2.5, and between changes in life expectancy and changes in PM2.5, for the period 2000 to 2007 in 545 US counties (Dataset 1). Results of the secondary analyses of the counties studied by Pope et al. (2009) using Datasets 2 and 3 are summarized in Appendix D (Tables 4.1 - 4.4). Table 1.1 lists the summary statistics for the variables in Dataset 1. In 2000, 189 of the 545 counties had a $\mathrm{PM}_{2.5}$ level greater than the current 3-year NAAQS level of $15 \mu\mathrm{g/m}^3$; by 2007 only 48 of those 189 were not in compliance with the NAAQS. On average, $\mathrm{PM}_{2.5}$ levels decreased at a rate of $0.22 \mu\mathrm{g/m}^3$ per year, a rate 33% lower than observed in the 211 counties analyzed for the period 1980 to 2000 ($0.33 \mu\mathrm{g/m}^3$ per year) {Pope et al. (2009)}.

Figures 1.2A and 1.2B show life expectancies plotted against $\mathrm{PM}_{2.5}$ levels for the years 2000 and 2007. Consistent with Pope et al. (2009) cross-sectional regression models showed a negative association between life expectancy and $\mathrm{PM}_{2.5}$ in both years. Details are summarized in the Appendix C. Figures 1.2C and 1.2D show changes in life expectancy plotted against changes in $\mathrm{PM}_{2.5}$ levels for 2000 to 2007. We also plotted the estimated regression lines under Models 1 and 3 of Table 1.2, defined below.

Table 1.1: **Summary Characteristics of the 545 Counties Analyzed for the Years 2000 to 2007:** $(^*)$, 2005 death rates are used as a proxy for 2007 death rates. COPD denotes chronic obstructive pulmonary disease.

Variable	Mean(SD)
Life Expectancy (yr.)	
2000	76.7 (1.7)
2007	77.5 (2.0)
Change	0.8 (0.6)
$PM_{2.5}$ ($\mu g/m^3$)	
2000	13.2 (3.4)
2007	11.6 (2.8)
Reduction	1.6 (1.5)
Per Capita Income (in thousands of $)	
2000	27.9 (7.4)
2007	30.4 (7.9)
Change	2.5 (2.3)
Population (in hundreds of thousands)	
2000	3.5 (6.3)
2007	3.8 (6.6)
Change	0.3 (0.6)
HS Graduates (% of pop.)	
2000	0.81 (0.07)
2007	0.85 (0.06)
Change	0.04 (0.02)
Black Population (% of pop.)	
2000	0.115 (0.138)
2007	0.117 (0.139)
Change	0.002 (0.017)
Hispanic Population (% of pop.)	
2000	0.119 (0.189)
2007	0.098 (0.135)
Change	-0.021 (0.057)
Deaths from Lung Cancer (no./ 10,000 pop.)*	
2000	16.4 (3.5)
2007	15.5 (3.8)
Change	-0.9 (2.2)
Deaths from COPD (no./ 10,000 pop.)*	
2000	12.8 (3.1)
2007	12.5 (3.5)
Change	-0.3 (2.1)

Figure 1.2: **Cross Sectional and First Difference Plots of $\mathbf{PM_{2.5}}$ vs. Life Expectancy:** Cross-sectional life expectancies plotted vs $\mathrm{PM}_{2.5}$ levels for (A) 2000 and (B) 2007 in Dataset 1. The slopes of the regression lines correspond to estimates from the simple model: LE $=$ intercept + slope*$\mathrm{PM}_{2.5}$ in both the 2000 and 2007 plots. In the second row on the left (C) the data are plotted as change in life expectancy vs change in $\mathrm{PM}_{2.5}$ over the period 2000 - 2007. The regression line corresponds to the simple model $\Delta\mathrm{LE} =$ intercept + slope*$\Delta\mathrm{PM}_{2.5}$ (Model 1 in Table 1.2). (D) On the right is the added variable plot for PM2.5 corresponding to Model 3 in Table 1.2.

Figure 1.2 (Continued)

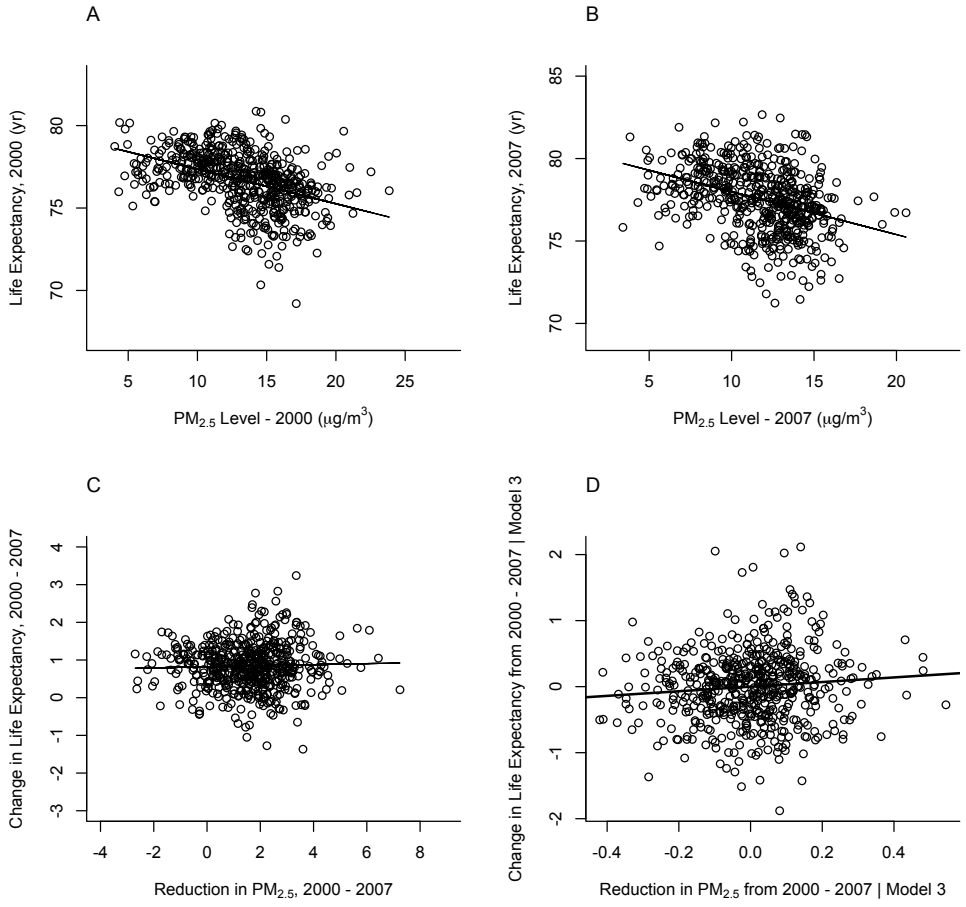


Table 1.2 summarizes estimated regression coefficients for the association between changes in $\mathrm{PM}_{2.5}$ and changes in life expectancy for 545 counties for 2000 to 2007 for selected regression models. When controlling for changes in all available socioeconomic and demographic variables as well as smoking prevalence proxy variables (Model 3), a $10\mu\mathrm{g/m}^3$ decrease in $\mathrm{PM}_{2.5}$ was associated with an estimated mean increase in life expectancy of 0.35 years (SE= 0.16 years, $p = 0.033$). The estimated effect of $\mathrm{PM}_{2.5}$ on life expectancy was consistent across models adjusting for various patterns of potentially confounding variables (e.g. Models 2 $\&$ 3). Models 4 - 8 of Table 1.2 show the results for select stratified and weighted regressions. In counties with a population density greater than 200 people per square mile, a $10\mu\mathrm{g/m}^3$ decrease in $\mathrm{PM}_{2.5}$ was associated with an increased life expectancy of 0.72 (0.22 years, $p < 0.01$) (Model 5), compared with -0.31 years (0.22 years, $p = 0.165$) in counties with less than 200 people per square mile (P difference < 0.01). In counties whose proportion of urban residences was greater than 90 percent, a $10\mu\mathrm{g/m}^3$ decrease in $\mathrm{PM}_{2.5}$ was associated with an increased life expectancy of 0.95 (0.31, $p < 0.01$) (Model 6), compared with -0.16 (0.16 years, $p = 0.299$) in counties with less than 90% urban residences (P difference < 0.01).

Table 1.2: **Results of Selected Regression Models for County-Level Analysis, 2000 - 2007:** (a), Included only counties with the largest year 2000 population in their respective MSA; (b), Included only counties with a year 2000 population density > 200 people per square mile; (c), Included only counties with a year 2000 urban rate $> 90\%$; (d), Weighted by the square root of the year 2000 population density; (e), Weighted by the inverse of county land area. Estimate for the effect of $\mathrm{PM}_{2.5}$ is for a $10\mu\mathrm{g}/\mathrm{m}^3$ reduction; change in income is given in thousands of dollars. Changes in LC ASDR and COPD ASDR are changes in the age standardized death rate for lung cancer and chronic obstructive pulmonary disease, respectively.

Variable	Model 1	Model 2	Model 3	Model 4[a]	Model 5[b]	Model 6[c]	Model 7[d]	Model 8[e]
Intercept	0.82±0.04	1.08±0.08	1.00±0.08	0.97±0.10	0.91±0.11	0.84±0.15	0.79±0.15	0.67±0.15
Reduction in $\mathrm{PM}_{2.5}$	0.14±0.19	0.35±0.17	0.35±0.16	0.30±0.23	0.72±0.22	0.95±0.31	0.74±0.24	0.96±0.28
Change in income	–	0.013±0.017	0.017±0.018	0.005±0.018	0.02±0.02	-0.01±0.03	0.03±0.02	0.05±0.02
Change in pop.	–	0.13±0.05	0.11±0.05	0.07±0.05	0.06±0.04	0.02±0.04	0.07±0.06	0.34±0.12
Change in HS%	–	-9.12±1.61	-7.98±1.56	-7.27±1.95	-4.42±2.60	-4.04±3.20	-1.94±3.35	-3.30±3.45
Change in black%	–	-6.55±2.05	-6.34±1.97	-7.86±3.07	-12.56±3.59	-8.12±2.84	-11.14±3.00	-6.21±2.97
Change in Hisp%	–	-2.16±0.47	-2.03±0.47	-2.12±0.59	-0.95±0.62	5.28±3.58	-3.25±0.63	-4.57±0.75
Change in LC ASDR	–	–	-0.02±0.02	-0.02±0.02	-0.01±0.05	-0.05±0.05	-0.07±0.02	-0.07±0.03
Change in COPD ASDR	–	–	-0.05±0.01	-0.05±0.02	-0.06±0.03	-0.06±0.05	-0.08±0.02	-0.06±0.02
No. of county units	545	545	545	257	307	169	545	545

When we re-estimated Model 3 of Table 1.2 using the square root of population density as the weight (Model 7), the estimated effect of a $10\mu g/m^3$ reduction of $PM_{2.5}$ on life expectancy was more than double that observed in our un-weighted analysis (0.74 [0.24] vs. 0.35 [0.16]). When that same model was weighted by the inverse of county land area (Model 8), the effect was nearly triple that of the un-weighted analysis (0.96 [0.27]). Table 1.3 summarizes a number of our stratified and weighted analyses.

Table 1.3: **Summary of Selected Stratified Regression Analyses for 545 Counties (Dataset 1, 2000 - 2007):** $(^*)$ Corresponds to the covariate pattern in Model 3 of Table 1.2. Covariates include change in income, change in population, change in proportion of high-school graduates, change in proportion of black population, change in proportion of Hispanic population, change in lung cancer mortality rate, and change in COPD mortality rate. Analysis used: SAS 9.2, PROC SURVEYREG, clustered by MSA, using the "weight" statement, and Stata 11.0, REGRESS using the "cluster" option.

Selected counties and analysis	Number of Counties	$\hat{\beta}$ (SE, p) for $10\mu g/m^3$ Reduction in $PM_{2.5}$ (full model)*
2000 Pop. Den. >1000	96	0.86(0.45, 0.061)
2000 Pop. Den. >800	116	0.62(0.41, 0.139)
2000 Pop. Den. >600	145	0.81(0.32, 0.014)
2000 Pop. Den. >400	197	0.84(0.27, 0.003)
2000 Pop. Den. >200	307	0.72(0.22, 0.001)
2000 Pop. Den. < 200	238	-0.31(0.22, 0.165)
2000 urban rate >90%	169	0.95(0.31, 0.003)
2000 urban rate >95%	109	1.12(0.32, 0.001)
2000 Pop. Den. >200 & 2000 urban rate >90%	159	0.96(0.28, 0.001)
2000 urban rate <90%	376	-0.16(0.16, 0.299)
All counties, regression weighted by square root of 2000 Pop. Den.	545	0.74(0.24, 0.002)
All counties, regression weighted by inverse of county land area	545	0.96(0.27, 0.001)

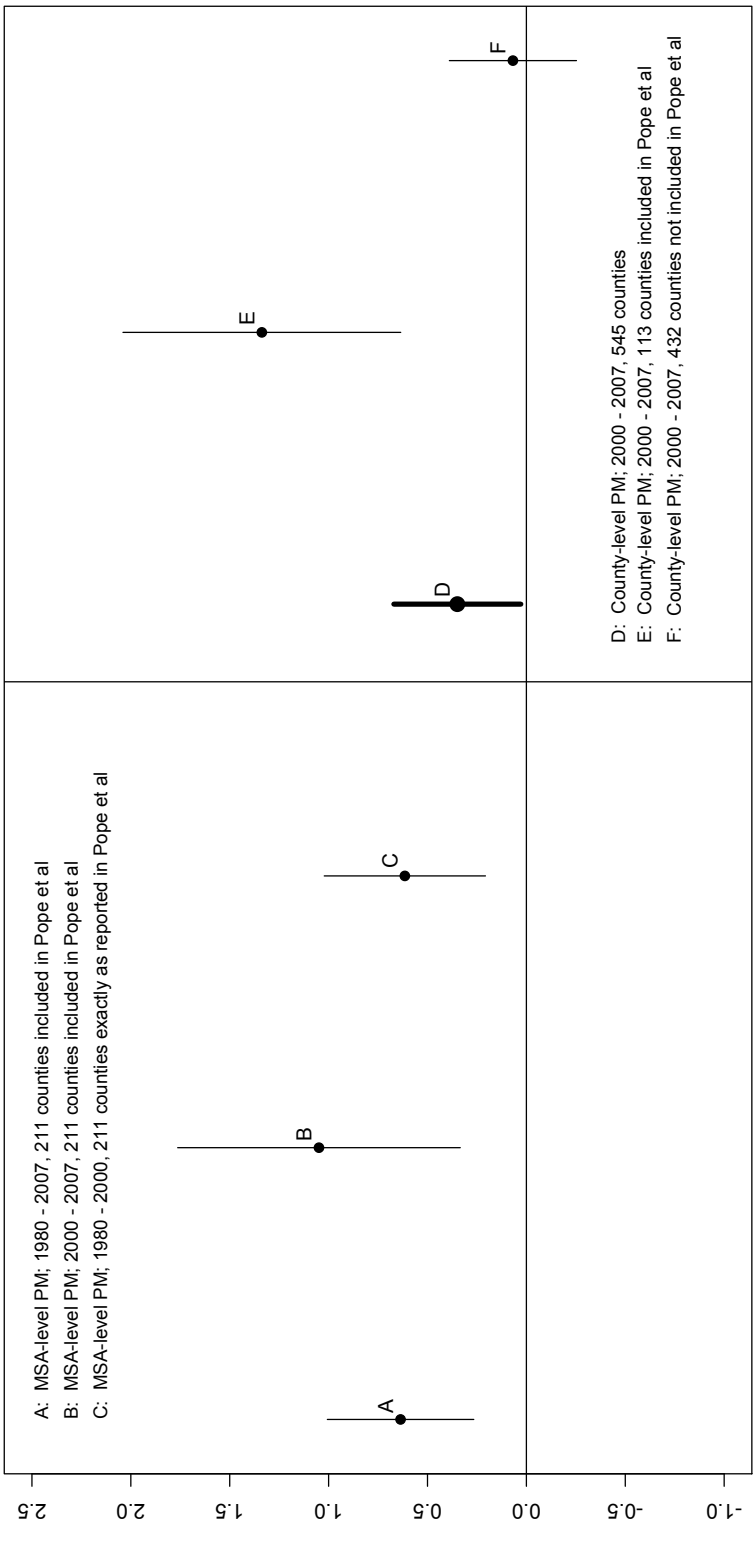
We conducted similar analyses for the 211-county dataset for 1980 to 2007 and from 2000 to 2007, the results of which are presented in Tables 4.3 and 4.4 of Appendix D,

respectively. Results for the period from 1980 to 2000 were identical to those reported by Pope et al. (2009).

Figure 1.3 summarizes the point estimates and 95% confidences interval for the effect of a $10\mu\mathrm{g}/\mathrm{m}^3$ decrease in $\mathrm{PM}_{2.5}$ on life expectancy for select un-weighted and un-stratified regression models in each dataset/time period. Models fitted using Datasets 2 and 3 (left) controlled for changes in income, population, proportion of the population that is black, lung cancer death rate, and COPD death rate, corresponding to Model 4 in eTables 2a,b. Models fitted using Dataset 1 controlled for all available variables and correspond to Model 3 in Table 1.2. These estimates were fairly consistent, though estimates corresponding to the counties from Pope et al. (2009) for the period 2000 to 2007 appeared slightly larger than those from other analyses.

Figure 1.3: **Effect Estimates and Confidence Intervals for the Effect of a $10\mu g/m^3$ Decrease in $PM_{2.5}$ on Life Expectancy:** Estimates A and B were obtained from Dataset 3; Estimate C was obtained from Dataset 2. Estimates A, B, and C were adjusted for changes in income, population, proportion of the population that is black, lung cancer death rate, and COPD death rate (Model 4, eTables 2a,b). Estimates D, E, and F were obtained from Dataset 1, adjusted for changes in income, population, proportion of high school graduates, proportion of the population that is black, proportion of the population that is Hispanic, lung cancer death rate, and COPD death rate (Model 3, Table 1.2). "Pope et al" refers to Pope et al. (2009).

Figure 1.3 (Continued)



In the analyses stratified by sex, the estimated effect of a $10\mu g/m^3$ reduction in $PM_{2.5}$ for the covariate pattern corresponding to Model 3 of Table 1.2 was an additional 0.59 (0.17) years of life expectancy for women and 0.08 (0.20) years for men (P difference = 0.027). Differences by sex were also observed in stratified and weighted models, although with less precision. Sex differences were smaller in the most urban counties (urban rate $> 90\%$). Similar results were observed for the period 1980 to 2000 in Dataset 2. Sex-specific results are presented in Table 1.4.

Table 1.4: **Comparison of Results of Select Models for Males vs. Females (Dataset 1, 2000 - 2007):** (*) Covariates include change in income, change in population, change in proportion of high-school graduates, change in proportion of black population, change in proportion of Hispanic population, change in lung cancer mortality rate, change in COPD mortality rate. Analysis used: SAS 9.2, PROC SURVEYREG, clustered by MSA, using the "weight" statement, and Stata 11.0, REGRESS using the "cluster" option. (†) Indicates that the estimate for males was statistically significantly different than the estimate for females for the model specified in that row.

Selected counties and analysis	Males: $\hat{\beta}$ (SE, p) for $10\mu g/m^3$ Dec. in $PM_{2.5}$ (full model)*	Females: $\hat{\beta}$ (SE, p) for $10\mu g/m^3$ Dec. in $PM_{2.5}$ (full model)*
All counties	0.08(0.20, 0.681)†	0.59(0.17, 0.001)†
2000 Pop. Density > 200	0.44(0.25, 0.084)	0.85(0.24, 0.001)
2000 Pop. Density < 200	-0.55(0.27, 0.043)	-0.06(0.24, 0.805)
2000 urban rate > 90%	0.81(0.37, 0.033)	1.07(0.28, <0.001)
2000 urban rate < 90%	-0.44(0.20, 0.025)	0.08(0.19, 0.664)
All counties, regression weighted by square root of 2000 Pop. Den.	0.57(0.29, 0.047)	0.87(0.22, <0.001)
All counties, regression weighted by inverse of county land area	0.74(0.30, 0.013)	1.14(0.30, <0.001)

Effect estimates were not highly sensitive to the inclusion of the estimated change in smoking prevalence. Table 1.5 summarizes the results for the inclusion/exclusion of the smoking prevalence variable across several models. For example, when Model 3 in Table 1.2 was re-estimated for the 383 counties with matching smoking prevalence data,

a reduction of $10\mu\mathrm{g}/\mathrm{m}^3$ was associated with an increase in life expectancy of 0.49 (0.19) years without including change in smoking prevalence in the model, and 0.47 (0.19) when including those changes. Similar results for smoking were observed in our stratified and weighted models, as well as in our models for men and women separately.

Table 1.5: **Comparison of $PM_{2.5}$ Effect Estimates from Selected Models for Inclusion of Smoking Variable Versus No Inclusion of Smoking Variable:** Covariates include change in income, change in population, change in high-school graduates, change in proportion of black population, change in proportion of Hispanic population, change in lung cancer mortality rate, change in COPD mortality rate. Analysis used: SAS 9.2, PROC SURVEYREG, clustered by MSA, using the "weight" statement, and Stata 11.0, REGRESS using the "cluster" option.

Selected counties and analysis	No. Counties	Full model, with smoking: $\hat{\beta}$ (SE, p) per $10\mu g/m^3$ $PM_{2.5}$	Full model, no smoking: $\hat{\beta}$ (SE, p) per $10\mu g/m^3$ $PM_{2.5}$
All Counties	383	0.47(0.19, 0.013)	0.49(0.19, 0.011)
2000 population density (persons per square mile)			
>800	110	0.52(0.43, 0.230)	0.53(0.43, 0.221)
>600	139	0.68(0.30, 0.028)	0.68(0.30, 0.027)
>400	187	0.71(0.26, 0.007)	0.70(0.25, 0.007)
>200	272	0.67(0.22, 0.003)	0.65(0.22, 0.004)
<200	111	-0.50(0.30, 0.100)	-0.39(0.30, 0.193)
2000 urban rate			
>90%	157	0.76(0.28, 0.009)	0.76(0.28, 0.008)
>95%	101	1.01(0.31, 0.002)	0.98(0.32, 0.003)
<90%	226	-0.14(0.20, 0.483)	-0.13(0.20, 0.513)
2000 population density & 2000 urban rate			
>200 & >90%	100	0.95(0.32, 0.004)	0.93(0.32, 0.005)
Regression weighted by square root of 2000 population density (All counties)	383	0.77(0.24, 0.002)	0.76(0.25, 0.003)
Regression weighted by inverse of county land area (All counties)	383	0.81(0.26, 0.002)	0.74(0.27, 0.007)
Sex			
Men	383	0.20(0.23, 0.389)	0.22(0.23, 0.343)
Women	383	0.71(0.20, 0.001)	0.72(0.20, <0.001)

1.4 Discussion

Data on air pollution and life expectancy from 545 US counties in 2000 and 2007 show that recent declines in $\mathrm{PM}_{2.5}$ to relatively low levels continue to prolong life expectancy in the US. These benefits are largest among the most urban and densely populated counties. These associations were estimated controlling for socioeconomic and demographic variables as well proxy variables for and direct measures of smoking prevalence.

In previous studies, a $10\mu\mathrm{g}/\mathrm{m}^3$ decrease in $\mathrm{PM}_{2.5}$ has been associated with gains from 0.42 to 1.51 years of life expectancy{Tainio et al. (2007); Pope et al. (2009)}. Here, a decrease of $10\mu\mathrm{g}/\mathrm{m}^3$ in $\mathrm{PM}_{2.5}$ was associated with an increase in life expectancy of 0.35 (0.16) for 545 counties for the period from 2000 to 2007. An increase in life expectancy of 0.56 (0.19) was estimated for the same 211 counties included in the Pope et al. (2009) analysis but extended to the period 1980 to 2007. The estimated effect in those 211 counties from 2000 to 2007 was equal to 1.00 (0.32). Stratified and weighted analyses within the 545 counties from 2000 to 2007 yielded larger estimates between 0.72(0.22) and 1.12(0.32) - broadly in agreement with those previously reported.

From 2000 to 2007, the average increase in life expectancy across the counties in this study was 0.84 years, and the average decrease in $\mathrm{PM}_{2.5}$ in those same counties was $1.56\mu\mathrm{g}/\mathrm{m}^3$. While $\mathrm{PM}_{2.5}$ reductions presumably account for some of the improvements in life expectancy over this period, it is only one of many contributing factors. Other factors may include improvements in the prevention and control of the chronic diseases of adulthood, particularly cardiovascular diseases (CVD) and stroke {Yeh et al. (2011); Shrestha (2005)}, and changes in the risk factors associated with them, including medical advances, declines in smoking, and decreases in blood pressure and cholesterol {Shrestha (2005)}. Given the well-established link between air pollution and CVD mortality{Pope et al. (1995, 2002, 2004)}, and changes in other CVD risk factors, issues of multicausality and competing risk make it difficult to quantify exactly the changes in life expectancy

attributable to reductions in $\mathrm{PM}_{2.5}$. However, if we consider one of our more conservative effect estimates (Model 3, Table 1.2) the $1.56\mu\mathrm{g/m}^3$ reduction in $\mathrm{PM}_{2.5}$ accounts for about 0.055 years (1.56×0.0354) of additional life expectancy, or roughly 7% of the increase in life expectancy. Using the estimate from our most urban counties (Model 6, Table 1.2), the increase in life expectancy attributable to the average reduction in $\mathrm{PM}_{2.5}$ was 0.148 years (1.56×0.095), or as much as 18% of the total increase.

An interesting aspect of this study was how pronounced the $\mathrm{PM}_{2.5}$ effect was for the original 211 counties from 2000 to 2007. Given that they were originally selected simply on the availability of matching pollution data, what is special about these counties that results in larger estimates of the effect of $\mathrm{PM}_{2.5}$ on life expectancy? The stratified and weighted analyses suggest plausible explanations. For instance, the 211 counties were all in metropolitan areas, and the stratified analyses suggest that the effect of $\mathrm{PM}_{2.5}$ on life expectancy is greatest in the most urban counties. One possible reason is that the composition of $\mathrm{PM}_{2.5}$ is different in urban areas {Louie et al. (2005)}, causing $\mathrm{PM}_{2.5}$ to have a larger health impact. Another possibility is the "non-metropolitan mortality penalty" - the recent phenomenon in which mortality rates are higher in rural compared with urban areas {Cossman et al. (2010)}. While it is not clear why the mortality gap between metro and non-metro areas has widened, some hypotheses include greater improvements in standards of care in metro areas, changes in uninsurance rates, changes in disease incidence, and changes in health behaviors {Cossman et al. (2010)}. These, however, would be valid explanations only if they occurred at different rates in metropolitan areas compared with rural areas. If so, then perhaps failure to include variables that captured one or more of these differences could explain the different estimates of the effect of $\mathrm{PM}_{2.5}$ on life expectancy.

Alternatively, metropolitan areas are more densely populated than non-metro areas. Our models that stratified by population density showed that the effect of $\mathrm{PM}_{2.5}$ on life expectancy is greatest in the most densely populated study areas (those with a population density of at least 200 people per square mile) - possibly suggesting a role

for differential exposure misclassification. That is, in densely populated areas, it is more likely that any two people from the same area are exposed to the same level of $\mathrm{PM}_{2.5}$ with perhaps less exposure misclassification. This possibility was supported in our models weighted by the square root of population density and the inverse of land area, which placed more weight on the most densely populated counties and the smallest counties. In these models the effect of a $10\mu\mathrm{g/m}^3$ decrease in $\mathrm{PM}_{2.5}$ on life expectancy was much larger than the equivalent un-weighted analysis.

Another interesting finding was the difference in the effect of changes in $\mathrm{PM}_{2.5}$ on men and women. Findings in the literature regarding the effects of air pollution by sex for long-term exposure have been mixed. Studies using the ACS and Harvard Six-Cities cohorts show no significant difference in pollution-related mortality between men and women {Dockery et al. (1993); Pope et al. (1995); Krewski et al. (2000); Pope et al. (2002, 2004); Laden et al. (2006)}. Studies using a Medicare cohort have reported different effects by age and region, but did not stratify by sex {Eftim et al. (2008); Greven et al. (2011); Zeger et al. (2008)}. In a study using the Adventist Health cohort, Chen et al. (2005) reported a large effect of $\mathrm{PM}_{2.5}$ on fatal coronary heart disease (CHD) in women but no association in men. Similarly, in separate studies, Ostro et al. (2010) using a cohort of women (California Teachers' Study), reported associations between particulate matter and cardiovascular mortality, while Puett et al. (2011) using a cohort of men (Male Health Professionals), found no association with all-cause mortality or fatal CHD. For our main analysis using all 545 counties, we find a larger effect of $\mathrm{PM}_{2.5}$ on women, suggesting that reductions in $\mathrm{PM}_{2.5}$ are more beneficial to gains in life expectancy for women. Models fitted using data for the period from 1980 - 2000 as in Pope et al. (2009) showed similar results. Future work should investigate more thoroughly the possibility of different $\mathrm{PM}_{2.5}$-mortality associations for men versus women.

One factor that appeared to play no role in the $\mathrm{PM}_{2.5}$ and life expectancy relationship, however, was baseline $\mathrm{PM}_{2.5}$ level. This is in agreement with the findings by Pope et al. (2009) and implies that, while we may see differences across levels of population

density, urban rate, and land area, this is not due to these areas having a higher or lower baseline $\mathrm{PM}_{2.5}$ level. Furthermore, this finding suggests that there is no clear threshold below which further reductions in $\mathrm{PM}_{2.5}$ levels provide no benefit (eAppendix, eTable 3). The fact that our results were not sensitive to the inclusion of direct measures of change in smoking prevalence suggests that the estimated gains in life expectancy for a $10\mu\mathrm{g/m}^3$ reduction in $\mathrm{PM}_{2.5}$ are not a result of confounding due to changes in smoking prevalence.

Table 1.6: **Summary of Selected Regression Analyses Stratified by Baseline $\mathbf{PM_{2.5}}$ Levels for 545 Counties (Dataset 1, 2000 - 2007):** $(^*)$ Corresponds to the covariate pattern in Model 3 of Table 1.2. Covariates include change in income, change in population, change in proportion of high-school graduates, change in proportion of black population, change in proportion of Hispanic population, change in lung cancer mortality rate, and change in COPD mortality rate. Analysis used: SAS 9.2, PROC SURVEYREG, clustered by MSA, using the "weight" statement, and Stata 11.0, REGRESS using the "cluster" option.

Selected counties and analysis	Number of Counties	$\hat{\beta}$ (SE, p) for $10\mu\mathrm{g/m}^3$ Reduction in $\mathrm{PM}_{2.5}$ (full model)*
2000 PM2.5 < 10mg/m3	100	-0.28(0.39, 0.482)
2000 PM2.5 < 12mg/m3	186	0.50(0.27, 0.065)
2000 PM2.5 < 14mg/m3	301	0.61(0.21, 0.004)
2000 PM2.5 < 16mg/m3	430	0.36(0.19, 0.064)
2000 PM2.5 < 18mg/m3	511	0.47(0.18, 0.009)
2000 PM2.5 > 18mg/m3	34	0.85(0.82, 0.314)
2000 PM2.5 > 16mg/m3	115	0.87(0.38, 0.023)
2000 PM2.5 > 14mg/m3	244	0.28(0.27, 0.305)
2000 PM2.5 > 12mg/m3	359	0.15(0.21, 0.462)
2000 PM2.5 > 10mg/m3	445	0.27(0.18, 0.126)

Unlike previous cross-sectional analyses {Evans et al. (1984); Özkaynak and Thurston (1987)}, we were able to estimate the association between county-specific temporal changes in $\mathrm{PM}_{2.5}$ levels and county-specific temporal changes on life expectancy adjusted by temporal changes in several potential confounding factors. By looking at within-county temporal changes, we reduce the potential bias due to unmeasured confounding. Further, by estimating clustered robust standard errors at the MSA level, we took a conservative approach in accounting for potential spatial correlation between neighboring counties.

Our analysis has the strengths of using some of the largest available datasets, and applying relatively simple analyses. Additionally, we improved on the original analysis by constructing a dataset with $\mathrm{PM}_{2.5}$ measured at the county level, in contrast to the more coarse MSA-level readings used in previous studies {Pope et al. (2002, 2009)}.

The analysis is limited, however, in its ability to control for all potential unmeasured confounding. Additionally, in comparing selected years, we do not fully exploit potentially informative data between those years. Furthermore, sophisticated analyses of the U.S. Medicare population by Greven et al. (2011) did not observe associations between "local" trends in $\mathrm{PM}_{2.5}$ levels and "local" trends in mortality in 814 zip code level locations in the U.S. for the period 2000 - 2006. "Local" trends were defined as the difference between monitor-specific trends and national trends. The Medicare cohorts, however, consisted only of people age 65 and older, whereas our life expectancy calculations integrate over all ages. Also, other studies using Medicare based cohorts have found significant associations between $\mathrm{PM}_{2.5}$ and overall mortality {Eftim et al. (2008); Zeger et al. (2008)}. Future work is needed to investigate whether these differences among studies are due to differences in statistical models, data sources, or populations studied.

It is also worth considering whether life expectancy was the most appropriate outcome to consider in our model. Because life expectancies are calculated from age-specific mortality rates, perhaps a model with age-specific mortality rates as the outcome would be more appropriate, allowing the age groups most affected by $\mathrm{PM}_{2.5}$ exposure to be pinpointed precisely.

In summary, our study reports strong evidence of an association between recent further reductions in fine-particulate air pollution and improvements in life expectancy in the United States, especially in small, densely populated urban areas.

A Closer Look at Exposure Decomposition in Long-term Air Pollution Studies: A Distributed Lag Approach

Andrew W. Correia and Francesca Dominici

Department of Biostatistics

Harvard University

2.1 Background

In the environmental epidemiology literature, there has been a great deal of work on assessing the impacts of air pollution exposure on various health outcomes. These studies range from short-term daily time-series studies {Dominici et al. (2002, 2003)}, to long-term cohort studies {Pope et al. (1995); Dockery et al. (1993)}, to long-term population based studies {Pope et al. (2009); Correia et al. (2013)}. Due to the nature of the research question of interest and the data available to researchers to answer those questions, the majority of these studies are observational. In general, caution is urged when interpreting parameter estimates from observational studies, as it is very difficult to properly control for every potential confounder (measured and unmeasured) in an observational study {Christenfeld et al. (2004); Greenland and Morgenstern (2001)}. Therefore, a major concern in environmental epidemiology is bias due to residual confounding, and indeed in both long-term and short-term studies, some critics argue that the estimated effect estimates of air pollution on mortality and morbidity are unreliable due to the difficulty of fully controlling for all potential confounders {Vedal (1997); Moolgavkar (1994, 2005)}.

Recent work by Janes et al., 2007 and Greven et al., 2011 has attempted to overcome issues of residual confounding by decomposing the air pollution exposure variable into a "local" term and a "global" term, where under certain assumptions, differences in the estimated effects of the "local" and "global" terms implies unmeasured confounding. The approach taken in these two papers has been somewhat controversial, as the models fitted via the exposure decompostion approach have estimated a null effect of air pollution on mortality at the "local" level - quite a contrast to the majority of literature on the subject. However, a thorough investigation of the modeling approach in Janes et al., 2007 and Greven et al., 2011 and a discussion as to why the results in those papers stand in contrast to the majority of the literature on air pollution and mortality has, to our knowledge, not been undertaken. The outline of this paper is as follows: 1) We will begin by focusing on the simpler approach presented in Janes et al., 2007, and discussing its methodology and modeling assumptions; 2) we illustrate the implications of those modeling assumptions

via simulation studies; 3) we then discuss the approach in Greven et al., 2011, showing how the modeling assumptions in that paper relate to those in Janes et al., 2007, and also how the simulation results apply to the Greven model; 4) we propose a model that is a combination of the Janes and Greven models, which also integrates a distributed lag on the "local" exposure term to more adequately model the temporal relationship between $\mathrm{PM}_{2.5}$ and mortality; and 5) we close with a discussion.

2.2 Overview of Methodology and Modeling Assumptions

In Janes et al., 2007, the authors' aim is to estimate the effect of $\mathrm{PM}_{2.5}$ on mortality in 113 US counties from 1999 to 2002 in the Medicare population. A particular county's monthly $\mathrm{PM}_{2.5}$ level is calculated as the average of $\mathrm{PM}_{2.5}$ over the preceding year - that is, the average $\mathrm{PM}_{2.5}$ level over the past 12 months, including the current month. Mortality counts in a given month for any county are simply the sum of the number of deaths in that county for the given month; these monthly counts are not given by an average of mortality counts over the preceding year. The authors then stratify individuals into one of six different age-sex strata. It is assumed that the causal model for the effect of $\mathrm{PM}_{2.5}$ on mortality in each age-sex stratum is given by:

$$\log \mathrm{E}(Y_t^c) = \log(N_t^c) + \delta_0^c + \delta_1 \mathrm{PM}_t^c, \tag{2.1}$$

where Y_t^c and N_t^c are the mortality counts and number of people at risk, respectively, for each county c and month t; the δ_0^c are county-specific random intercepts; and δ_1 is the association between month-to-month variation in PM_t^c and month-to-month variation in mortality.

Because estimates from Model 2.1 are likely to be confounded by variables trending in a similar fashion to $\mathrm{PM}_{2.5}$ and mortality, Janes and colleagues introduce another popular model in the environmental epidemiology literature in which temporal con-

founding is addressed, at least in part, by a smooth function of time. Specifically:

$$\log \mathrm{E}(Y_t^c) = \log(N_t^c) + \beta_0^c + \beta_1 \mathrm{PM}_t^c + s(t; d), \tag{2.2}$$

where β_0^c and β_1 are defined analogously to δ_0^c and δ_1, respectively, and $s(t; d)$ is a natural cubic spline with d degrees of freedom. From Model 2.2, Janes et al., 2007 propose:

$$\log \mathrm{E}(Y_t^c) = \log(N_t^c) + \eta_0^c + \eta_1 \widehat{\mathrm{PM}}_t + \eta_2(\mathrm{PM}_t^c - \widehat{\mathrm{PM}}_t) + s^*(t; d-1), \tag{2.3}$$

where $\widehat{\mathrm{PM}}_t$ is the annual average in $\mathrm{PM}_{2.5}$ for month t across all counties and $s^*(t; d-1)$ is orthogonal to both $\widehat{\mathrm{PM}}_t$ and PM_t^c.

Predicted values from Models 2.2 and 2.3 are equivalent. However, Janes et al., 2007 point out that because Model 2.3 estimates the association between $\mathrm{PM}_{2.5}$ and mortality at both a national scale and a local scale, we are able to detect unmeasured confounding via large differences between the estimates of η_1 and η_2 - if there is no confounding or measurement error, η_1 and η_2 should be equal; if there is confounding, it is more likely to be at the global-level (η_1) than at the local level (η_2). Additionally, the authors state that the random, county-specific intercepts in the model control for unmeasured county-specific characteristics that do not vary with time (i.e. they control for unmeasured spatial confounding).

In summary, the modeling assumptions given in Janes et al., 2007 (and, generally, in Greven et al., 2011 as well) are as follows:

1. The true causal model for the effect of $\mathrm{PM}_{2.5}$ on mortality is given by Model 2.1.
2. Absent confounding and measurement error, if the causal link between mortality and $\mathrm{PM}_{2.5}$ is given by Model 2.1 then the estimates of η_1 and η_2 in Model 2.3 should be equal.

3. Estimates should not be biased as a result of spatial confounding due to the inclusion of random, county-specific intercepts.

4. Mortality in month t has a causal relationship with $\mathrm{PM}_{2.5}$ levels averaged over the past twelve months up to and including t.

5. The estimate of η_2 is less likely to be confounded than the estimate of η_1.

Since one cannot ever know the true underlying causal model, we assume that item 1 is correct. Further, for reasons outlined in Janes et al., 2007 and Greven et al., 2011, we will proceed under the assumption that item 5 is correct as well. Then, given the proposed causal model (2.1), we test assumptions 2, 3 and 4 via simulation in the following section. Because the model in Greven et al., 2011 is much more computationally intensive, we conduct our simulations based on the model presented in Janes et al., 2007 and then discuss how those simulation results relate to the Greven model.

2.3 Simulation Study

2.3.1 Equality of η_1 and η_2

To test assumption 2, we first simulated data based on Model 2.1 with $\delta_1 = 0.009$ and with $\delta_0^c = -5.75$ for all c using real $\mathrm{PM}_{2.5}$ and population data - the same $\mathrm{PM}_{2.5}$ and population data used in Greven et al., 2011, where it is described in more detail. We then analyzed the simulated data with Model 2.3 to be sure that we do indeed observe $\eta_1 \approx \eta_2$ in this simplest scenario. Note that because we assume $\delta_0^c = -5.75$ for all c, we fit Model 2.3 with a fixed intercept (η_0) instead of a random intercept (η_0^c). This is only to improve computational speed and has no impact on assessing the validity of assumption 2. Also note that we set $\delta_1 = 0.009$, as this corresponds roughly to the average estimate across strata in the non-decomposed model in Janes et al., 2007, and it is also roughly the average effect-estimate of a number of short-term studies summarized in Table 1 of Pope and Dockery, 2006. As this study explores mostly temporal variability, much like many

short-term studies, we believe this is a realistic, though perhaps conservative value for δ_1 given the results in the literature for other long-term studies {Table 2, Pope and Dockery, 2006}. Throughout this paper, the degrees of freedom parameter for the cubic spline in Models 2.2 and 2.3, d, is taken to be 16 as in Janes et al., 2007.

Results under this basic simulation assuming no confounding are given in Table 2.1 below. Indeed, though estimates of η_1 are a bit more variable than those of η_2, we see that, on average, $\eta_1 \approx \eta_2$ when there is no confounding.

Table 2.1: **Performance of Model 2.3 Under the Assumption of no Confounding:** η_1 is the global parameter and η_2 is the local parameter.

Parameter	Avg. Estimate	Avg. SE	Avg. p-value
η_1	9.02×10^{-3}	4.61×10^{-3}	0.150
η_2	9.02×10^{-3}	8.31×10^{-4}	< 0.001

Temporal Confounding

We then simulate under Model 2.1 again, but with the addition of a confounder, U_t, trending only at the national level. That is, U_t was generated to be correlated with both the outcome and $\widehat{\mathrm{PM}}_t$, but U_t is not correlated with $(\mathrm{PM}_t^c - \widehat{\mathrm{PM}}_t)$. Outcome data was generated via the following model:

$$\log \mathrm{E}(Y_t^c) = \log(N_t^c) + \delta_0^c + \delta_1 \mathrm{PM}_t^c + \delta_2 U_t, \tag{2.4}$$

where $\delta_2 = 0.15$. Simulated data is again modeled under Model 2.3. Results are summarized in Table 2.2 below:

Here, we see that with a strong influence from a "global" confounder, the estimate of η_1 is indeed inflated, though the estimate of η_2 remains an accurate estimate of the

Table 2.2: **Performance of Model 2.3 Under the Assumption of Global-scale Temporal Confounding:** η_1 is the global parameter and η_2 is the local parameter.

Parameter	Avg. Estimate	Avg. SE	Avg. p-value
η_1	8.44×10^{-2}	7.52×10^{-3}	< 0.001
η_2	9.09×10^{-3}	1.38×10^{-3}	< 0.001

true effect. The same is true when outcomes are generated from a model with a positive interaction effect between $\widehat{\mathrm{PM}}_t$ and U_t.

Spatial Confounding

In this section we test assumption 3 - that the estimates of η_1 and η_2 should not be impacted by confounders that do not vary with time to due to the inclusion of location-specific intercepts. Thus, consider a new confounder, U^c, that only varies spatially and is constant across time. Outcome data is generated by:

$$\log \mathrm{E}(Y_t^c) = \log(N_t^c) + \delta_0^c + \delta_1 \mathrm{PM}_t^c + \delta_2 U^c, \tag{2.5}$$

where $\delta_0^c = -6.25$ for all c, $\delta_1 = 0.009$, and $\delta_2 = -0.15$. U^c is generated such that $\mathrm{Corr}(U^c, \overline{\mathrm{PM}}^c) \approx 0.25$, and $\overline{\mathrm{PM}}^c = (1/T)\sum_t \mathrm{PM}_t^c \ \ \forall c$. We again model this outcome data via Model 2.3, but this time allowing for estimation of location-specific intercepts, η_0^c. Results are summarized in Table 2.3 below.

Table 2.3: **Performance of Model 2.3 Under the Assumption of Spatial Confounding:** η_1 is the global parameter and η_2 is the local parameter.

Parameter	Avg. Estimate	Avg. SE	Avg. p-value
η_1	1.04×10^{-2}	7.57×10^{-3}	0.166
η_2	-6.12×10^{-3}	3.04×10^{-3}	0.052

We can see that, despite allowing for the estimation of location-specific intercepts, the average estimate of η_2 is severely biased, while the estimate of η_1 is only slightly

biased, suggesting that assumption 3 - that estimates are not susceptible to spatial confounding - does not hold. In other words, in the presence of unmeasured spatial confounding, even if we introduce into the model a county-specific random intercept, the estimate of the local effect can be severely biased.

Spatial and Global Temporal Confounding

In this section, we assume there exists both a spatial confounder, U^c, and a "global" temporal confounder, U_t. Outcome data is generated by:

$$\log \mathrm{E}(Y_t^c) = \log(N_t^c) + \delta_0^c + \delta_1 \mathrm{PM}_t^c + \delta_2 U^c + \delta_3 U_t, \tag{2.6}$$

where $\delta_0^c = -8.22$ for all c, $\delta_1 = 0.009$, $\delta_2 = -0.15$, and $\delta_3 = 0.15$. Confounders U^c and U_t are generated such that $\mathrm{Corr}(U^c, \overline{\mathrm{PM}}^c) \approx 0.25$, and $\mathrm{Corr}(U^t, \overline{\mathrm{PM}}_t) \approx 0.3$. Outcome data is again modeled via Model 2.3, again allowing for the estimation of location-specific intercepts, η_0^c. Results are summarized in Table 2.4 below:

Table 2.4: **Performance of Model 2.3 Under the Assumption of Spatial and Global-scale Temporal Confounding:** η_1 is the global parameter and η_2 is the local parameter.

Parameter	Avg. Estimate	Avg. SE	Avg. p-value
η_1	6.83×10^{-2}	7.40×10^{-3}	< 0.001
η_2	-6.09×10^{-3}	3.04×10^{-3}	0.054

Results for the local term, η_2, are very similar to the previous case with only spatial confounding, which is to be expected since the local and global $\mathrm{PM}_{2.5}$ terms are orthogonal. We also observe that the global term, η_1, becomes inflated due to the global temporal confounder, as in the earlier simulation with only temporal confounding at the global level. Thus, in the case of global temporal confounding together with spatial confounding, we observe that both the local and global terms can be quite biased. It is also possible in this situation, since the local and global terms are each affected separately by the spa-

tial and global temporal confounding, that both estimates could similar in magnitude but *both* be biased as a result of different sources of confounding.

Spatio-temporal Confounding

Now, suppose instead that confounding takes place at the local level and varies with time. Specifically, we generate U_t^c to be correlated with PM_t^c ($\mathrm{Corr}(\mathrm{PM}_t^c, U_t^c) \approx 0.25$) and generate outcome data via the following model:

$$\log \mathrm{E}(Y_t^c) = \log(N_t^c) + \delta_0^c + \delta_1 \mathrm{PM}_t^c + \delta_2 U_t^c, \tag{2.7}$$

again with $\delta_1 = 0.009$, $\delta_2 = 0.15$, and $\delta_0^c = -7.35$ for all c . Analyzing this data under Model 2.3 allowing for the estimation of location-specific intercepts yields the following results (Table 2.5):

Table 2.5: **Performance of Model 2.3 Under the Assumption of Local Spatio-temporal Confounding:** η_1 is the global parameter and η_2 is the local parameter.

Parameter	Avg. Estimate	Avg. SE	Avg. p-value
η_1	3.27×10^{-2}	7.86×10^{-3}	0.002
η_2	2.99×10^{-2}	1.46×10^{-3}	< 0.001

In this instance, we observe that $\eta_1 \approx \eta_2$. Because of this, however, we would incorrectly assume that the estimates are *not* confounded, though they are, in fact, biased upward - both more than $3\times$ higher than the truth, 0.009.

2.3.2 Size of the Window for Averaging Monthly Air Pollution

Two-month Rolling Mean vs 12-month Rolling Mean

In this section, we test the implications of assumption 4 being incorrect. Recall that assumption 4 assumes that mortality at time t is associated with the average $\mathrm{PM}_{2.5}$ over

the past 12 months. First, we investigate what happens if the true causal relationship between $\mathrm{PM}_{2.5}$ and mortality only exists at, say, a two-month window as opposed to the 12-month window assumed in Janes et al., 2007 and Greven et al., 2011.

Consider, again, outcome data generated via Model 2.1 with no confounding and with δ_0^c and δ_1 as described above. However, we generate that outcome data with $\mathrm{PM}_t^c = (\mathrm{PM}_{t,raw}^c + \mathrm{PM}_{t-1,raw}^c)/2$, where $\mathrm{PM}_{t,raw}^c$ is the raw, observed $\mathrm{PM}_{2.5}$ level in county c at time t, not averaged over any previous or future months' values. We then model this outcome data using Model 2.3, but with $\mathrm{PM}_t^c = \frac{1}{12}\sum_{i=0}^{11}\mathrm{PM}_{t-i,raw}^c$, as in Janes et al., 2007 and Greven et al., 2011. Results are given in Table 2.6 below:

Table 2.6: **Performance of Model 2.3 Under the Assumption of a Mis-specified Rolling Mean:** η_1 is the global parameter and η_2 is the local parameter.

Parameter	Avg. Estimate	Avg. SE	Avg. p-value
η_1	9.21×10^{-3}	5.95×10^{-4}	< 0.001
η_2	8.82×10^{-3}	2.55×10^{-4}	< 0.001

Surprisingly, the mis-specified PM_t^c that incorporates information from ten extra, uninformative months actually performs quite well, and η_1 and η_2 are indeed nearly the same and very near the truth of 0.009. As in the previous section, the estimate of the local term, η_2, was very robust to the inclusion of a "global" confounder, and remained unbiased. Thus, it appears that mis-specifying the length of the rolling mean is not terribly serious offense with regards to accurately estimating the regression coefficients.

Generating Data with a Distributed Lag

Suppose now that the association between $\mathrm{PM}_{2.5}$ and mortality is best captured by a distributed lag model:

$$\log\mathrm{E}(Y_t^c) = \log(N_t^c) + \delta_0^c + \sum_{l=0}^{q}\delta_l^*\mathrm{PM}_{t-l}^c, \tag{2.8}$$

where q is the maximum lag we will consider and l is the amount of lag, ranging from 0

to q. In other words, we assume not only that pollution at time t affects mortality at time t, but also that pollution levels at times $t-1, t-2, \ldots, t-q$ impact mortality at time t. Model 2.8 above is known as an *unconstrained* distributed lag model - one simply calculates the lagged values of the exposure and plugs them directly into the regression. If the successive values of the exposure are highly correlated, however, estimates of the the δ_l's will be very unstable. To overcome this, it's useful to constrain the δ_l's to increase efficiency of the estimated lag parameters {Schwartz (2000)}. The most popular approach for constraining the regression coefficients is that of Almon, 1965, where the shape of the distributed lag (see figure below for an example) is fit by some polynomial function of degree p. Alternative choices include constraining the shape of the distributed lag with a piecewise natural cubic spline {Corradi and Gambetta, 1976; Zanobetti et al., 2000}, B-splines of an arbitrary degree, or simple moving averages (as was done in Janes et al. (2007) and Greven et al. (2011)), among others {Gasparrini et al. (2010)}.

Now, let's suppose that the relationship between $\mathrm{PM}_{2.5}$ and the relative risk (RR) of mortality is described by Figure 2.1. Here, we see some initial mortality displacement, or "harvesting" {Schwartz (2001)}, followed by a (only partially pictured) sustained but modest long-term increase in the RR of mortality after around 13 months.

We generated outcome data under Model 2.8 so that the relationship between $\mathrm{PM}_{2.5}$ and mortality is described by Figure 2.1, with δ_0^c constant across all c; we then fit that data under Model 2.3. Results are summarized below:

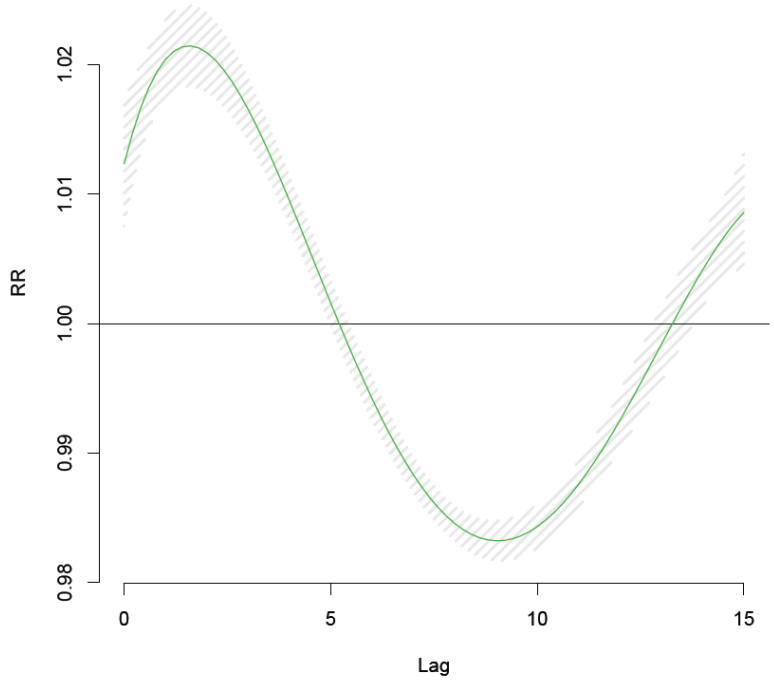
Table 2.7: **Performance of Model 2.3 Under the Assumption of an Underlying Lagged Relationship:** η_1 is the global parameter and η_2 is the local parameter.

Parameter	Avg. Estimate	Avg. SE	Avg. p-value
η_1	6.93×10^{-3}	2.74×10^{-3}	0.066
η_2	3.71×10^{-4}	2.76×10^{-4}	0.278

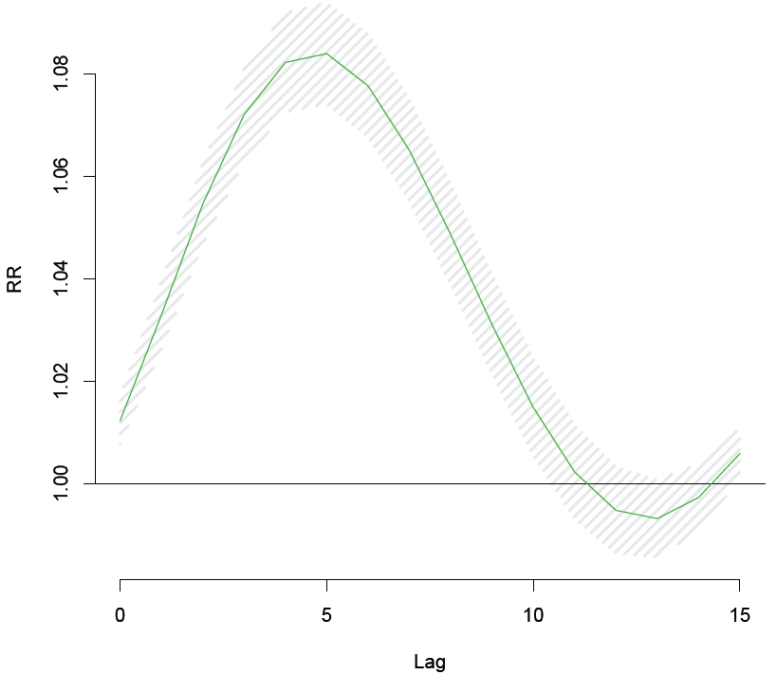
In this simulation, we observe $\eta_1 >> \eta_2$ (more than $18\times$ greater). However, there is no confounding under this simulation, only a distributed lag where the relationship is given by that of Figure 2.1. Model 2.3 does relatively well in capturing the overall

Figure 2.1: **Example of Distributed Lag Relationship and Corresponding Cumulative Distributed Lag:** Plots of the relative risk of mortality by lag (a) and cumulative relative risk of mortality by lag (b).

Figure 2.1 (Continued)



(a) Relative risk for a $10\mu g/m^3$ increase in $PM_{2.5}$ by lag



(b) Cumulative relative risk for a $10\mu g/m^3$ increase in $PM_{2.5}$ by lag

RR over the entire 15-month lag (estimated, on average, to be 6.36×10^{-4}), but it is unable to tell us anything about the large bump in the RR of mortality that exists out until just after 5 months. Thus, if the true relationship between $\mathrm{PM}_{2.5}$ and mortality over time is given by a similar distributed lag, the model specified by Janes et al., 2007 - and the accompanying assumptions for that model - would lead to a conclusion that the estimates are confounded, and that, on average, the local effect estimate is not statistically significant. The model would also fail to identify any bumps in mortality that are a function of mortality displacement over the duration of the lag window.

2.4 Extensions to Greven et al

In Greven et al., 2011, the authors begin with individual-level data, and wish to fit the proportional hazards model:

$$h^c(a,t) = h^c(a)\exp(x_t^c\beta),$$

where $h^c(a,t)$ denotes the hazard of dying at age a and time t for location c, $h^c(a)$ is a location-specific baseline hazard, and x_t^c is average $\mathrm{PM}_{2.5}$ exposure for county c at time t as described above. Age a takes on integer values from 65 to 89; subjects age 90 or older are pooled into the same age group, $\mathrm{I}(a \geq 90)$. Due to computational constraints given the size of the data set (18.2 million individuals across 814 different locations), the authors instead opt to fit the log-linear regression model:

$$\log\mathrm{E}(Y_{at}^c) = \log(N_{at}^c) + \log(h^c(a)) + x_t^c\beta, \tag{2.9}$$

which is equivalent to the originally proposed survival model, under a piecewise exponential assumption, with regard to likelihood-based inference {Holford (1980);

Laird and Olivier (1981)}. From Eq. 2.9, the authors then propose a model where x_t^c is decomposed:

$$\log \mathrm{E}(Y_{at}^c) = \log(N_{at}^c) + \log(h^c(a)) + (x_t^c - \bar{x}_t - \bar{x}^c + \bar{x})\beta_1 + (\bar{x}_t - \bar{x})\beta_2, \tag{2.10}$$

where the goal is to, as in Janes et al., 2007, identify unmeasured confounding via large differences between the estimates of β_1 and β_2. Estimating the parameters in Model 2.10 directly is computationally demanding, as there are still roughtly 1.4 million observations across all locations and times combined, and there is a need to directly estimate the log-hazard $\log(h^c(a))$ for all 814 locations, c, separately to control for spatial confounding. Thus, the model is fitted using a backfitting algorithm {Buja et al. (1989)}, which iterates between *Step 1:* estimating the $\mathrm{PM}_{2.5}$ effect for all locations - β_1 and β_2 - including the previous iteration's estimated hazard as an offset, and *Step 2:* separately estimating the log-hazard function for each location with $(x_t^c - \bar{x}_t - \bar{x}^c + \bar{x})\beta_1 + (\bar{x}_t - \bar{x})\beta_2$ as an offset.

Model 2.10 is very similar to Model 2.3; all of the simulation results above for the model in Janes et al. apply to this model except for one: Because a log-hazard function is estimated separately for each location via the backfitting algorithm, the location-specific hazard functions eliminate all purely spatial variation, which means that the estimate of β_1 can not be confounded by variables that vary only across locations. This approach is similar to fitting a separate model for each location, c, and then pooling the β_1^c estimates across all locations; clearly no location-specific variables can be identified in that case because they would be absorbed into the intercept term since they are constant over time.

Consider the following – for any fixed time point, $t = T$, Model 2.10 is given by:

$$\log \mathrm{E}(Y_{a,t=T}^c) = \log(N_{a,t=T}^c) + \log(h^c(a)) + (x_{t=T}^c - \bar{x}_{t=T} - \bar{x}^c + \bar{x})\beta_1 + (\bar{x}_{t=T} - \bar{x})\beta_2. \tag{2.11}$$

Now, $\bar{x}_{t=T}$ and $\bar{x}$ are constant across c. Thus, Eq. 2.11 can be rewritten as:

$$\log \mathrm{E}(Y^c_{a,t=T}) = \log(N^c_{a,t=T}) + \log(h^c(a)) + (x^c_{t=T} - \bar{x}^c - \Delta\bar{x}_{t=T})\beta_1. \tag{2.12}$$

However, this model contains two terms that are constant with respect to c – the location specific indicator, $\log(h^c(a))$, and, because t is fixed at $t = T$, $(x^c_{t=T} - \bar{x}^c - \Delta\bar{x}_{t=T})$ – which means β_1 is not identifiable. Thus, β_1 is only estimable via temporal variations, which means that it is not susceptible to bias via purely spatial confounders. Therefore, the results from the "spatial confounding" and "spatio-temporal confounding" sections above do not apply to the Greven et al., 2011 model.

For some confounder U^t_c to bias the local effect β_1, it would have to be associated with county-specific deviations in *both* PM^c_t and mortality from each of their respective national trends. An example would be if communities which showed larger decreases in $\mathrm{PM}_{2.5}$ than the national average also consistently showed larger decreases in smoking rates than the national average, and vice versa {Greven et al., 2011}. While possible, this type of confounding is certainly less likely than variables trending in a similar fashion on the national level, which is why these analyses focus more on the local effect estimates.

Another key distinction to make between the two models is the decomposition of PM^c_t in each model. In Janes et al. (2007), the modeling approach decomposes the exposure x^c_t into $(x^c_t - \bar{x}_t)$ and $(\bar{x}_t - \bar{x})$. In Greven et al., the exposure x^c_t is decomposed into $[(x^c_t - \bar{x}_t) - (\bar{x}^c - \bar{x})]$, $(\bar{x}_t - \bar{x})$, and $(\bar{x}^c - \bar{x})$, where $(\bar{x}^c - \bar{x})$ gets absorbed into the location-specific hazard. Interestingly, while the "local" terms in each model should have similar interpretations {Greven et al., 2011}, it appears that they do not. For the data used by Greven and colleagues, we calculated both exposure decompositions - call $(x^c_t - \bar{x}_t) \equiv x_{J,local}$ and call $[(x^c_t - \bar{x}_t) - (\bar{x}^c - \bar{x})] \equiv x_{G,local}$. $\mathrm{Corr}(x_{J,local}, x_{G,local}) = 0.27$ and $\mathrm{Corr}(x_{G,local}, \mathrm{PM}^c_t) = 0.26$, while $\mathrm{Corr}(x_{J,local}, \mathrm{PM}^c_t) = 0.98$. Clearly, the local term in the Janes decomposition is more representative of the raw $\mathrm{PM}_{2.5}$ levels that we want to make inference about. Also consider the Table 2.8 below, in which we present the results for our most basic simulation under no confounding, but using the Greven decomposition instead of the Janes decomposition.

Table 2.8: **Performance of Model 2.10 Under no Confounding:** β_2 is the global parameter and β_1 is the local parameter.

Parameter	Avg. Estimate	Avg. SE	Avg. p-value
β_2	9.29×10^{-3}	6.71×10^{-3}	0.284
β_1	8.72×10^{-3}	4.95×10^{-3}	0.188

Here, we see that while the local estimate is generally unbiased for the true effect of 0.009, it is not nearly as precise as the estimate obtained using the Janes decomposition. In fact on average, neither the local nor the global term in this decomposition is statistically significant - and this is absent any confounding at all! Thus, we propose an approach using the Janes decomposition but with the modeling approach of Greven to reanalyze the data from Greven et al., 2011. Additionally, because our parameter of interest - the coefficient on the local exposure term - is estimated solely via temporal variations, we propose the inclusion of a distributed lag for the local exposure to more precisely determine how and when longer-term $\mathrm{PM}_{2.5}$ exposure can impact mortality.

2.5 Methods

2.5.1 Statistical Approach

Due to the large number of individuals in the data set (described briefly in the previous section, and in slightly more detail in the next section), fitting an individual-level model is not feasible. Thus, the data was aggregated to the zip code level, pooling the number of individuals and death counts for each zip code, enabling us to fit:

$$\log \mathrm{E}(Y_{at}^c) = \log(N_{at}^c) + \log(h^c(a)) + (x_t^c - \bar{x}_t)\beta_1 + (\bar{x}_t - \bar{x})\beta_2, \tag{2.13}$$

where $\bar{x}_t$ denotes the national trend in annual average $\mathrm{PM}_{2.5}$, calculated as the fitted values of a linear regression with PM_t^c as the outcome and a smooth function of time with 16 degrees of freedom as the predictor {Janes et al., 2007}. In this data set, the

correlation between $(x_t^c - \bar{x}_t)$ and $(\bar{x}_t - \bar{x})$ is less than 10^{-15}, implying that the terms are essentially orthogonal. Therefore, potential confounders trending at the national-level cannot bias the estimate of β_1. While it is possible for β_1 to be confounded by local trends, this - as we discussed in the previous section - is less likely. Because β_2 is likely to be confounded, and because we believe it is less likely for β_1 to be confounded, we focus our attention primarily on the estimate of β_1.

Further, it is reasonable to believe that the current month's $\mathrm{PM}_{2.5}$ exposure is associated not only with the current month's mortality, but that it may be associated with the mortality rate in subsequent months as well. To account for this, we impose a lag on $(x_t^c - \bar{x}_t)$. Call $(x_t^c - \bar{x}_t) \equiv x_{t,local}^c$; then the model we fit is given by:

$$\log \mathrm{E}(Y_{at}^c) = \log(N_{at}^c) + \log(h^c(a)) + \sum_{l=0}^{q} \eta_l x_{t-l,local}^c + (\bar{x}_t - \bar{x})\gamma. \tag{2.14}$$

Because lagged values of the exposure $x_{t-l,local}^c$ will be highly correlated with one another, directly estimating the η_l will be difficult. To overcome this, we assume the shape of the distributed lag is given by some smooth function, which we estimate via piecewise natural cubic splines {Corradi and Gambetta, 1976; Zanobetti et al., 2000}. Because previous work has suggested that the effects of long-term exposure to $\mathrm{PM}_{2.5}$ are primarily felt within 0 to 2 years {Schwartz et al. (2008)}, we specified a lag of $q = 20$ months. Additionally, because successive monthly $\mathrm{PM}_{2.5}$ levels are so highly correlated and because we want to avoid an unrealistically wiggly function of the η_l, we only place internal knots at $l = (5, 10, 15)$. Further, note that because we believe the estimate for the global term will be biased, we do not impose a lag on that term since the estimates of the lagged "global" effects will also likely be biased.

In Model 2.14 we estimate the log-hazard, $\log(h^c(a))$, separately for each location, allowing us to eliminate spatial variability from the problem and thus eliminate the possibility of confounding due to factors that vary only across locations. Because this greatly increases the dimensionality of the problem, however, we estimate the parameters in

Model 2.14 via a backfitting algorithm {Greven et al. (2011); Buja et al. (1989)}. The algorithm consists of two steps at each iteration, j - the first, in which the parameters η_l and γ are estimated given the previous iteration's log-hazard, $\log(h^c(a))^{(j-1)}$, and the log of the risk set, $\log(N^c_{at})$, as an offset; and the second, in which the location-specific hazards are estimated separately for each location with $\log(N^c_{at}) + \sum_{l=0}^{q} \eta_l^{(j)} x^c_{t-l,local} + (\bar{x}_t - \bar{x})\gamma^{(j)}$ included as the offset. We run the algorithm until a stopping criteria is reached; here, we stop the algorithm if the $l2$ norm of the difference between the current iteration's parameter estimates and the previous iteration's parameter estimates is less than 10^{-6}. Once convergence is achieved, conclude with step 1 one last time.

2.6 Results

The dataset used in this study is identical to that used in Greven et al., 2011. Specifically, our data consists of monthly ambient $\mathrm{PM}_{2.5}$ measures linked with Medicare mortality data for the period 2000 - 2006. $\mathrm{PM}_{2.5}$ data was obtained from the EPA monitoring network and is from 814 monitor locations across the continental United States, chosen solely on the basis of availability. Each of the 814 locations has $\mathrm{PM}_{2.5}$ readings for at least four years, where each of the four years includes data from at least 10 months. Long-term $\mathrm{PM}_{2.5}$ exposure in month t is defined as the average exposure over the previous 12 months up to and including t.

Air pollution data was linked to Medicare mortality data as follows: The same monitor-level reading was assigned to all Medicare enrollees residing in a zip code whose geographical centroid is within a 6 mile radius of the monitor. The Medicare data consists of time of death, precise up to the month, as well as demographic data - age, gender, and race. The study population consists of 18.2 million Medicare enrollees and 3.2 million deaths over the entire period 2000 - 2006. More information can be found about this data in Greven et al., 2011 and at *http://www.resdac.umn.edu/Medicare*.

The estimated lagged RR of local monthly $\mathrm{PM}_{2.5}$ exposure on mortality is given by

Figure 2.2 below. The local relative risk (RR) associated with a $10\mu\text{g}/\text{m}^3$ increase in $\text{PM}_{2.5}$ does not drop to zero until just before 5 months, though the only lag that is statistically significantly different from zero is the 2-month lag. The cumulative RR of a $10\mu\text{g}/\text{m}^3$ increase in $\text{PM}_{2.5}$ over the entire 20 month lag, however, is 1.055, CI = $(1.052, 1.059)$; this corresponds to a log-RR of $0.0540(\pm 0.0018)$. Consistent with previous results under this type of model specification, the estimate for the global term is much larger. Here, the log-RR associated with a $10\mu\text{g}/\text{m}^3$ increase in $\text{PM}_{2.5}$ is $0.3780(\pm 0.0096)$, very comparable to the original estimate of $0.4313(\pm 0.0084)$ reported in Greven et al., 2011.

Note also that these results were not highly sensitive to different specifications of the knots in the distributed lag. As a check, we re-estimated the cumulative RR over the entire $q = 20$ month lag with only two internal knots at $l = (6.66, 13.33)$. The shape of the distributed lag curve was very similar to that observed in Figure 2.2, and the estimated cumulative RR was 1.054, CI = $(1.051, 1.058)$. Even completely doing away with the lag did not do much to change the estimated RR. That is, we used Model 2.13 to estimate the local effect, β_1, and found that for a $10\mu\text{g}/\text{m}^3$ increase in $\text{PM}_{2.5}$, the estimated RR of mortality was 1.052, CI = $(1.049, 1.056)$, a very slight decrease from the estimates obtained using the distributed lag models.

We also fit models that additionally controlled for local seasonality by including in Model 2.14 location-specific indicators for month. However the estimated relative risks were not impacted at all by the inclusion of these terms, suggesting that our estimates of the local association between $\text{PM}_{2.5}$ and mortality are not biased due to local seasonality issues.

Figure 2.2: **Plot of the Estimated RR of Mortality Associated with a $10\mu\mathrm{g}/\mathrm{m}^3$ Increase in $\mathrm{PM}_{2.5}$:** Here we see the estimated RR of mortality associated with a $10\mu\mathrm{g}/\mathrm{m}^3$ increase in $\mathrm{PM}_{2.5}$ plotted as a function of the lagged $\mathrm{PM}_{2.5}$ exposure. The association fades relatively quickly, suggesting that a lag shorter 20 months is likely to be sufficient.

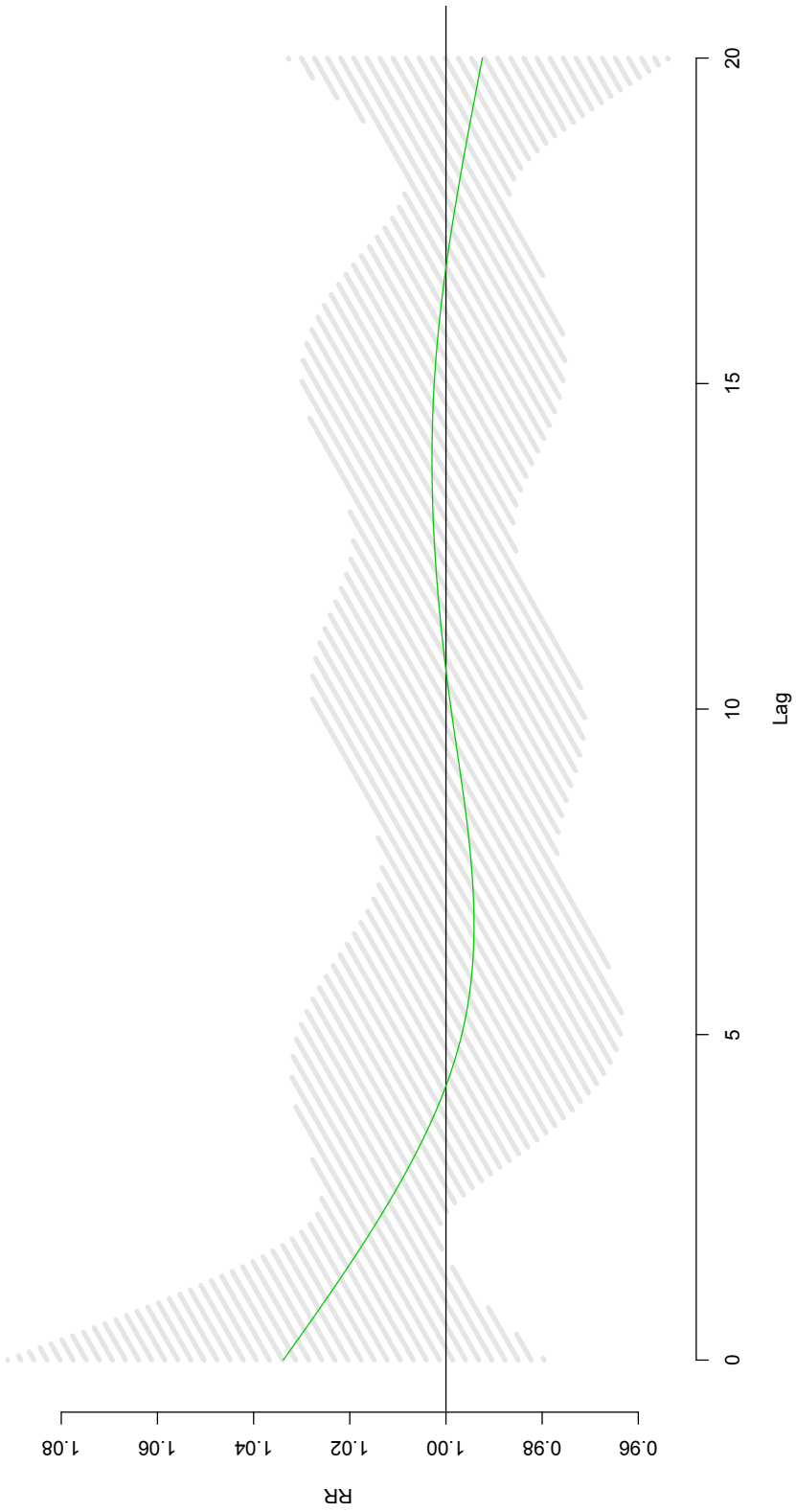


Figure 2.2 (Continued)

Another concern in fitting this model is that the local and global terms are approximately orthogonal to each other. Consider Table 2.9 below, which is the model based correlation matrix corresponding to Model 2.14. Focusing on only the second column in Table 2.9 we see that the largest correlation between a "Local PM" cubic spline basis vector and the "Global PM" variable is 0.033. All other spline bases have a correlation with Global PM less than 0.01. Table 2.10 shows the model-based correlation matrix corresponding to Model 2.13. Here, the correlation between the local and global $\mathrm{PM}_{2.5}$ variables is -0.019. By comparison, the model-based correlation between the local and global $\mathrm{PM}_{2.5}$ variables in Greven et al. (2011) is -0.017.

Table 2.9: **Model-based Correlation Matrix Corresponding to Model 2.14:** The five "Local PM" variables correspond to the five vectors that define the location of the knots for the piecewise cubic spline used to implement the distributed lag in Model 2.14.

	Intercept	Global PM	Local PM 1	Local PM 2	Local PM 3	Local PM 4	Local PM 5
Intercept	1						
Global PM	-0.015	1					
Local PM Basis 1	0.002	0.004	1				
Local PM Basis 2	-0.001	0.002	-0.827	1			
Local PM Basis 3	0.003	-0.003	0.659	-0.871	1		
Local PM Basis 4	-0.007	-0.007	-0.707	0.309	-0.425	1	
Local PM Basis 5	0.002	0.033	0.088	0.343	-0.608	-0.172	1

Table 2.10: **Model-based Correlation Matrix Corresponding to Model 2.13:** We observe virtually no correlation between the local and global $\mathrm{PM}_{2.5}$ variables in our non-lagged model, (2.13)

	Global PM	Local PM
Global PM	1.000	-0.019
Local PM	-0.019	1.000

2.7 Discussion

In this work, we investigated why recent models proposed by Janes et al. (2007) and Greven et al. (2011) that attempt to adjust for and identify confounding in air pollution studies by decomposing the exposure variable into two approximately orthogonal pieces of information - a local and a global exposure term - did not detect any association between $\mathrm{PM}_{2.5}$ and mortality at the local level in their studies. We conducted a number of simulation studies on the model proposed by Janes et al. (2007) to examine which, if any, of their modeling assumptions were violated and what the implications of those violations were. We identified that the local $\mathrm{PM}_{2.5}$ effect estimates were in fact still susceptible to residual spatial confounding, despite the underlying modeling assumption that they would not be due to the inclusion of location-specific random intercepts. We also identified that if the true exposure response relationship is given by a lagged relationship, then the association between $\mathrm{PM}_{2.5}$ and mortality could be severely biased, both globally and locally.

We did not directly test the modeling assumptions in the Greven framework, but rather, using the knowledge we gained through the simulation results pertaining to the Janes model, we were able to identify what particular scenarios could cause biased estimates in the Greven model. We saw that, generally, all of the modeling assumptions in the Greven paper did hold true, and its main advantage over the model described in Janes et al. (2007) is that the Greven model completely eliminates spatial variation (and thus the possibility of confounding due to spatially varying covariates) by estimating

location-specific hazard functions *separately* for each location. However, the estimates in the Greven model could still be biased if the true exposure response relationship is given by a lagged relationship. Additionally, while the decomposition in the Greven model does create two approximately orthogonal exposures - one local and one global - the local variable in the Greven model performed more poorly in our simulations with regards to bias and MSE than did the local exposure term specified by Janes and colleagues. It is also a more difficult quantity to interpret given the more complicated decomposition. The claim made by Greven et al. (2011) is that their decomposition is such that the approximate orthogonality of the local and global exposure terms is more closely orthogonal than the terms in the Janes decomposition. In our model, we found that this is in fact true, though not by as much as one would hope. The model-based correlation between the local and global terms in Greven model is -0.017, while for the Janes decomposition in the Greven model framework that correlation is -0.019.

However, this work should not be interpreted as a criticism of the Janes and Greven papers, but rather as a fusion of the two. Indeed, both approaches have their strengths and weaknesses; here, we ultimately combined those two approaches into one, attempting to keep only the strong points of each. In doing so, we constructed a model that: 1) effectively controls for spatial confounding by eliminating all spatial variation, and thus does not allow the parameter estimates to be informed by underlying differences among locations; 2) decomposes the exposure estimate into two orthogonal pieces of information - a "local" exposure and a "global" exposure - as a means to adjust for and identify residual confounding; 3) can identify an exposure-response relationship for lagged values of the exposure; and 4) is not sensitive to adjustment for local seasonality.

In this paper, we estimated that a $10\mu\mathrm{g}/\mathrm{m}^3$ increase in $\mathrm{PM}_{2.5}$ was statistically significantly associated (across all models) with anywhere between a 5.2% and 5.5% increase in the risk of mortality. The estimates were consistent across lagged and non-lagged models, as well as for different specifications of the lagged models. Generally speaking, these estimates are largely consistent with those estimated in previous studies examining the

impact of long-term $\mathrm{PM}_{2.5}$ exposure on mortality {Pope and Dockery (2006)}.

What is interesting to consider is why Greven et al. (2011) and Janes et al. (2007) did not observe any association between $\mathrm{PM}_{2.5}$ and mortality at the local level. There are many possible reasons. First, consider the Janes study. We showed via simulation that the modeling approach in Janes et al. (2007) was susceptible to spatial confounding. In previous work that exploited purely spatial information, we have seen that the $\mathrm{PM}_{2.5}$ effect on mortality is smaller in models that don't adjust for other covariates than in models adjusting for key socioeconomic and demographic variables over a similar time period {Correia et al. (2013)}. Thus, it is possible that the estimates from the Janes study were simply biased downward as a result of residual confounding. It's also possible that the span of the study 2000 - 2002 was not enough time to observe an association, and the study was simply underpowered due to the relatively little change in $\mathrm{PM}_{2.5}$ levels over such a short period of time.

In the Greven study, we know that confounding due to location-specific variables is not possible. While confounding due to local trends is possible, that same confounding would also affect the estimates in our model - yet, we observe a positive and statistically significant association in this study, whereas the Greven analysis, using the same data used here, does not. Thus, we will assume that local confounding is also not the reason why no association was observed in the Greven analysis. This means that the only other plausible explanation would be the different exposure decomposition used in our study compared to the original analysis by Greven et al. (2011). In our simulation, we showed that the Greven decomposition suffers from a larger MSE than the Janes decomposition, and on average - absent confounding - the local estimate under the Greven decomposition was not statistically significant. Thus, it seems that by further orthogonalizing the local and global exposure variables, Greven and colleagues were left with a variable that was no longer as strong a predictor of the outcome as the simpler decomposition, while not gaining a substantial amount with respect to further orthogonalizing the local and global terms. Here, we revert to the original decomposition from Janes et al. (2007), which

still achieves approximate orthogonality of the local and global terms for the purposes of Model 2.13, and also which has a local term that is a stronger predictor of the association between $\mathrm{PM}_{2.5}$ and mortality.

Exposure to Aircraft Noise and Hospital Admissions for Cardiovascular Diseases: A Large National Multi-Airport Population Cohort Study

Andrew W. Correia[1], Junenette L. Peters[2], Jonathan I. Levy[2], Steven Melly[3], and Francesca Dominici[1]

[1] Department of Biostatistics, Harvard University

[2] Department of Environmental Health, Boston University School of Public Health

[3] Department of Environmental Health, Harvard School of Public Health

3.1 Introduction

Aircraft noise has been associated with physiological responses and psychological reactions {Bluhm and Eriksson (2011); Hatfield et al. (2001)}, such as sleep disturbances, sleep-disordered breathing, nervousness, and annoyance {Hatfield et al. (2001); Rosenlund et al. (2001)}; however, the relationship of aircraft noise to direct health effects is less well established. Recent literature, primarily from one multicenter European study, has provided growing evidence for a relationship between aircraft noise and hypertension outcomes, including incidence of hypertension {Eriksson et al. (2010)}, self-reported hypertension {Rosenlund et al. (2001)}, increased blood pressure {Haralabidis et al. (2008); Jarup et al. (2008); Haralabidis et al. (2011)}, and antihypertensive medication use {Bluhm and Eriksson (2011); Greiser et al. (2007); Franssen et al. (2004); Floud et al. (2011)}. This is supported by a broader literature which evaluated the cardiovascular effects of noise and found substantial evidence for biological plausibility and positive associations between noise and hypertension, myocardial infarction (MI), and ischemic heart disease (IHD) {Babisch and Kim (2011)}. Potential biological mechanisms may include induced release of stress hormones {Ising and Kruppa (2004); Spreng (2000); Selander et al. (2009)} and indirect effects on sympathetic activity, which is associated with adverse metabolic outcomes {Selander et al. (2009); Grassi (2006b,a); Mancia et al. (2006)}.

However, few studies of the relationship between aircraft noise and cardiovascular disease (CVD) have been conducted to date {Bluhm and Eriksson (2011)}, in part because studies surrounding a small number of airports are not typically adequately powered. To our knowledge, only one study in Switzerland has examined CVD mortality, finding an association between airport noise and MI mortality {Huss et al. (2010)}. No major study has been conducted to date to estimate the association between long-term exposure to aircraft noise and hospital admissions for cardiovascular outcomes. Appropriately characterizing this association requires sufficient number of airports with large surrounding populations, applying statistical methods that can combine findings across airports and account for potential confounders both at the individual and the area level.

In this study, we leverage the large and nationally representative United States (US) population of Medicare enrollees to evaluate the CVD implications of airport-related noise in the US Specifically, this study aims to evaluate the relationship between average residential exposure to aircraft noise and hospital admission for cardiovascular-related diseases in the population ≥ 65 years of age residing near airports in the contiguous US. Understanding the link between aircraft noise and CVD outcomes is important in characterizing the potential benefits of intervention strategies {Stansfeld and Crombie (2011)}.

3.2 Methods

The cohort for this study was taken from 2009 US Medicare enrollees ≥ 65 years of age. Information obtained for enrollees includes date of death and hospitalization records, which contain date of hospitalization, length of hospital stay, the associated *International Classification of Diseases* (ICD) primary and secondary diagnostic and procedure codes, and the costs billed to Medicare. Additional individual-level data include age, gender, race, and zip code of residence.

Cause-specific hospitalizations for five cardiovascular outcomes were considered based on ICD, Ninth Revision (ICD-9) codes for primary diagnosis: heart failure (HF) (ICD-9 428); heart rhythm disturbances (ICD-9 426 to 427); cerebrovascular events (ICD-9 430 to 438); IHD (ICD-9 410 to 414, 429); and peripheral vascular disease (ICD-9 440 to 448). A variable for total CVD admissions was calculated as the sum of these causes.

3.2.1 Noise Exposure Estimates

We used 2009 noise contours developed for 89 airports in the contiguous US Noise contours were provided to us by the US Federal Aviation Administration (FAA) who, in turn, used the Integrated Noise Model (INM) version 7.0a {Federal Aviation Administration - Office of Environment and Energy (2007)}. Given our interest in characterizing noise on a continuous scale with high spatial resolution, the FAA ran the INM in a mode

that estimates noise exposure for each census block centroid surrounding each airport out to a minimum of 45 dB. The noise descriptor used was Day-Night Sound Level (DNL), which adds a 10 dB "penalty" to nighttime (i.e., 2200-0700 hr) {Miedema et al. (2000)}.

Medicare data provides residential information at the zip code level only. Therefore, we used the noise exposure measures in combination with 2010 US Census data on population counts, both at the census block level, to obtain aggregated measures of exposure to aviation-related noise at the zip code tabulation area (ZCTA) level. Specifically, to calculate the aggregated measures of noise exposure, we assumed that the study population is uniformly distributed within a census block. We then overlaid noise estimates by census block along with estimates of the population $\geq$ 65 years of age by census block, based on U.S. Census 2010 Summary File 1 (SF1) Table P12. For ZCTAs that included census blocks below 45 dB, we assumed that those blocks were exposed to 45 dB. The ZCTAs for which no census blocks had noise estimates above 45 dB were omitted from the analysis. We then constructed a number of candidate exposure metrics for each ZCTA, but focused on two in particular: 1) population-weighted average noise among the census blocks within each ZCTA, where each census block was weighted by the age $\geq$ 65 population, and 2) the 90^{th} percentile noise exposure among the census blocks within each ZCTA that contain non-zero population age $\geq$ 65.

More formally, for each ZCTA, we calculated the population-weighted noise exposure, x_z as: $x_z = \frac{1}{pop_z} \sum_j p_j \times x_j$, where j indexes the census blocks in ZCTA z, p_j is the number of individuals age $\geq$ 65 in census block j, x_j is the estimated noise exposure at the centroid of census block j, and pop_z is the total number of individuals age $\geq$ 65 in ZCTA z. For blocks that were split by the 45 dB contour line, noise exposure was estimated as the population exposed to over 45 dB - estimated as the fraction of p_j inside the contour line - multiplied by x_j, plus the estimated population outside of the contour line multiplied by 45 dB.

3.2.2 Statistical Methods

We estimate the effect of aircraft noise exposure on hospital admissions for all cardiovascular outcomes and for separate outcomes of cerebrovascular disease, IHD, and HF. Preliminary analyses indicated that heart rhythm disturbances and peripheral vascular disease were too infrequent to include as stand-alone outcomes.

Individual-level control variables included age (> 75 or ≤ 75), gender (male or female), and race (white [non-Hispanic] or non-white). We also used ZCTA-level Census 2000 data to characterize variables that might proxy for other potential confounders (% black, % Hispanic, % graduated high school, median household income), using a subset given high correlations among the variables. We also obtained ZCTA-level average yearly fine particulate matter ($\mathrm{PM}_{2.5}$) and ozone levels from the US Environmental Protection Agency's Air Quality System database where available.

For each ZCTA included into the analysis, we calculate the number of cause-specific admissions and the number of people at risk (Medicare enrollees) stratified by age/gender/race. Each ZCTA represented part of a cluster of other ZCTAs around one of the 89 airports. We then fit hierarchical Poisson models to data aggregated at the ZCTA level with airport-specific random intercepts and slopes. The models can be described in two stages. First, we specified a Poisson regression model for ZCTA-level counts to estimate risk of CVD admissions associated with exposure to aircraft noise separately for each airport, adjusted for individual-level and ZCTA-level confounders. Second, we combined information across airports and accounted for clustering by specifying airport-specific random effects. Our main analysis (combined CVD) was run in R (version 2.15.0) under a fully Bayesian approach, enabling us to estimate both airport-specific effects and overall fixed effects.

Specifically, we fit the following model:

$$\begin{aligned}\log(E[Y_{z,s}^A]) &= \log(N_{z,s}^A) + \beta_0^A + \beta_1^A \mathrm{I}(age > 75) + \beta_2^A \mathrm{I}(sex = M) \\ &+ \beta_3^A \mathrm{I}(race = nonwhite) + \beta_4^A (x_z^A - \bar{x}) + \boldsymbol{\gamma}^T W_z^A,\end{aligned} \tag{3.1}$$

where $Y_{z,s}^A$ and $N_{z,s}^A$ are the number of CVD hospitalizations and total population, respectively, for age/sex/race strata s in ZCTA z in airport A; x_z^A is the noise exposure variable for ZCTA z in airport A (either population weighted noise or the 90^{th} percentile of noise among the blocks in ZCTA z), and W_z^A is the vector of potentially confounding SES, demographic, or air quality variables for ZCTA z in airport A.

Let $\boldsymbol{\beta^A} = (\beta_0^A, \beta_1^A, \beta_2^A, \beta_3^A, \beta_4^A)^T$; we specify our fully Bayesian approach as follows:

$$\begin{aligned}\boldsymbol{\beta^A} &\sim \mathcal{N}_5(\boldsymbol{\theta}, \boldsymbol{\Sigma}), \\ \boldsymbol{\theta} &\sim \mathcal{N}_5(\boldsymbol{\mu_0}, \boldsymbol{\Lambda_0}), \\ \boldsymbol{\Sigma} &\sim \mathcal{IW}(\eta_0, \boldsymbol{S_0}), \\ \boldsymbol{\gamma} &\propto 1,\end{aligned} \tag{3.2}$$

where $\mathcal{N}_5(\alpha, V)$ denotes a 5-dimensional multivariate normal distribution with mean vector α and covariance matrix V, and $\mathcal{IW}(\eta, W)$ denotes an inverse Wishart distribution with η degrees of freedom and scale matrix W. In specifying our priors by Eq. 3.2, we are able to take advantage of the conjugacy of the normal and inverse Wishart distributions and compute our posteriors for $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ using a standard Gibbs sampling algorithm {Casella and George (1992)}. Indeed, it is straightforward to show that full conditional posterior distributions for $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ are given by:

$$\begin{aligned}P(\boldsymbol{\theta} \mid \cdot) &\sim \mathcal{N}_5\left((\boldsymbol{\Lambda_0^{-1}} + n_A \boldsymbol{\Sigma^{-1}})^{-1}(\boldsymbol{\Lambda_0^{-1}}\boldsymbol{\mu_0} + n_A \boldsymbol{\Sigma^{-1}}\bar{\boldsymbol{\beta}}^{\boldsymbol{A}}), (\boldsymbol{\Lambda_0^{-1}} + n_A \boldsymbol{\Sigma^{-1}})^{-1}\right), \\ P(\boldsymbol{\Sigma} \mid \cdot) &\sim \mathcal{IW}\left(\eta_0 + n_A, (\boldsymbol{S_0} + [\boldsymbol{\beta}_{\boldsymbol{mat}}^{\boldsymbol{A}} - \boldsymbol{\theta}_{\boldsymbol{mat}}]^T [\boldsymbol{\beta}_{\boldsymbol{mat}}^{\boldsymbol{A}} - \boldsymbol{\theta}_{\boldsymbol{mat}}])^{-1}\right),\end{aligned} \tag{3.3}$$

where n_A is the total number of airports in our dataset (89), $\boldsymbol{\beta^A_{mat}}$ is an $n_A \times p$ matrix, where each row corresponds to an airport-specific vector of parameter estimates $\boldsymbol{\beta^A}$, $\boldsymbol{\theta_{mat}}$ is also an $n_A \times p$ matrix, where each row is equal to $\boldsymbol{\theta}$, and $p = \text{Length}(\boldsymbol{\theta}) = 5$.

Because the location-specific effects, $\boldsymbol{\beta^A}$, and fixed effects, $\boldsymbol{\gamma}$, do not have conjugate priors relative to the Poisson likelihood implied by Eq. 3.1, however, we can not derive a closed form for the full conditional posterior distributions of $\boldsymbol{\beta^A}$ and $\boldsymbol{\gamma}$. Thus, we include in our Gibbs sampler a Metropolis step in order to estimate the posterior distributions of each $\boldsymbol{\beta^A}$ and the $\boldsymbol{\gamma}$.

Selecting appropriate values for $\boldsymbol{\mu_0}, \boldsymbol{\Lambda_0}$, and $\boldsymbol{S_0}$ can be a difficult problem, particularly in the absence of explicit prior data {Hoff (2009)}. However, using unit information priors {Kass and Wasserman (1995)}, a type of weakly informative prior, provides a way of obtaining reasonable estimates for these hyperparameters. To obtain estimates of $\boldsymbol{\mu_0}$ and $\boldsymbol{\Lambda_0}$, the population mean and covariance, respectively, of the $\boldsymbol{\beta^A}$'s, we follow the strategy outlined in Hoff (2009) and fit Eq. 3.1 separately for each airport with at least 20 surrounding ZCTAs. This gives us 42 different airport-specific MLE estimates, $\tilde{\boldsymbol{\beta}}^{\boldsymbol{a}}$. The unit information prior for $\boldsymbol{\theta}$, then, would be given by the multivariate normal distribution described in Eq. 3.2 where $\boldsymbol{\mu_0} = \frac{1}{42}\sum_a \tilde{\boldsymbol{\beta}}^{\boldsymbol{a}}$ and $\boldsymbol{\Lambda_0}$ set equal to the sample covariance of the $\tilde{\boldsymbol{\beta}}^{\boldsymbol{a}}$. We also set $\boldsymbol{S_0}$ equal to the sample covariance matrix, and we choose $\eta_0 = p + 2$.

For each cardiovascular outcome (combined CVD hospitalizations, and separate sub-analyses for cerebrovascular disease, ischemic heart disease and heart failure hospitalizations), we constructed three hierarchical models (as in Eq. 3.1) for each of the two noise metrics (population-weighted noise exposure and 90th percentile of noise exposure). These three different models are characterized by which covariates enter the model along with noise exposure to control for potential confounding. Models 1 controlled for individual-level variables (age, gender, and race) only. Models 2 additionally controlled for ZCTA-level SES and demographic variables (% Hispanic and median household income), and Models 3 added pollution variables ($\text{PM}_{2.5}$ and ozone) to Model 2. For Models 2 and 3, the ZCTA-level SES, demographic, and air pollution variables enter Eq. 3.1 as a

column the matrix W_z^A; for Model 1, $W_z^A = 0$. In the interest of computational time, the cause-specific analyses - that is, our sub-analyses for each of the single CVD hospitalization outcomes as opposed to our main outcome of interest, total CVD hospitalizations - were run using the `glmer()` function in the linear mixed effects models package (lme4) in R, which fits the hierarchical Poisson models (Eq. 3.1) in a frequentist framework by a Laplace approximation of the log-likelihood {Bates (2012)}. We note that overall effect-estimates, $\boldsymbol{\theta}$, obtained using the lme4 package in R are nearly identical to those obtained via our fully Bayesian approach, though we opt for the fully Bayesian approach in our main analysis of total CVD hospitalizations as this allowed us to estimate not only the overall effect of noise exposure on CVD hospitalizations, but also to estimate the airport-specific effect estimates as well as their standard errors and credible intervals.

3.3 Results

There were totals of 2,218 ZCTAs (779 with $\mathrm{PM}_{2.5}$ and ozone data) and 6,027,363 Medicare enrollees residing within the 45 dB contour level of the 89 airports. The number of ZCTAs within the 45 dB contour level ranged from seven to 107 across the airports. The number of Medicare enrollees in these ZCTAs ranged from 8,556 to 482,200 across the airports. Table 3.1 summarizes the population characteristics of this cohort, and Figure 1 provides a map presenting the 89 airports.

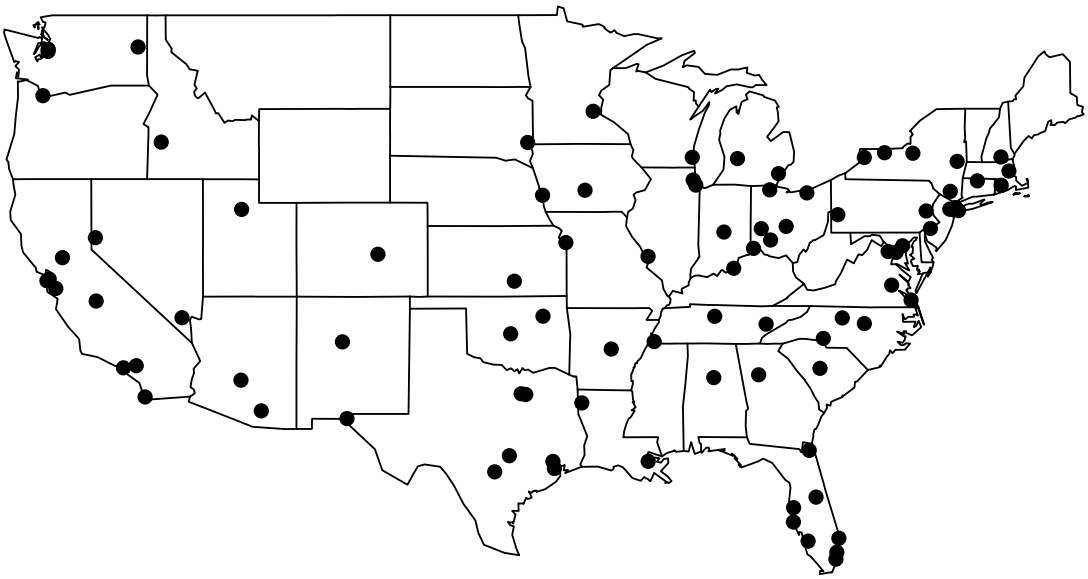
Results for the nationally aggregated relative risk of a CVD hospitalization for each one dB increase in population-weighted noise exposure and in 90^{th} percentile of noise exposure are displayed in Figure 3.2. For the 90^{th} percentile of noise exposure variable, controlling for age, gender, and race, an increase of one dB in the 90^{th} percentile of noise within a ZCTA was associated with an increase of 0.29% (95% CI 0.08 to 0.49) in the relative risk of having a CVD hospitalization (Model 1). In Model 2, which additionally controls for ZCTA-level SES and demographic variables, the log-relative risk of having a CVD hospitalization was only marginally significant (relative risk 0.16% (95% CI -0.02 to 0.34)). In Model 3, adding pollution variables to Model 2, an increase in the 90^{th} per-

Table 3.1: **Distribution of Zip Code Tabulation Area (ZCTA) Level Exposure for 2,218 ZCTAs and Risk Factor Data for Approximately Six Million Medicare Enrollees in 2009:** Values are the 25^{th}, 50^{th} (Median), and 75^{th} percentiles for each variable across all ZCTAs. (*) 1165 ZCTAs with data on $PM_{2.5}$. (†) 779 ZCTAs with data on ozone - a proper subset of the 1165 ZCTAs with data on $PM_{2.5}$.

Characteristics	$25^{th}\%$	Median	$75^{th}\%$
% > 75 years old (among population $\geq$ 65)	37.3	42.7	47.7
% Black	1.8	5.5	20.2
% Hispanic	2.1	6.2	19.8
Median household income (thou. $)	34.9	45.1	57.3
% graduated high school	72.8	82.9	90
$PM_{2.5}$ ($\mu g/m^3$)*	9.1	10.2	11.3
Ozone (ppm)†	0.022	0.025	0.028
Population-weighted noise (dB, DNL)	45.1	45.9	48.6
90th percentile of noise among populated census blocks (dB, DNL)	47.5	50.3	54.5
Hospital admission rate per 100,000 individuals			
Cerebrovascular events (stroke)	677.3	1093.4	2229.4
Ischemic heart disease	885.7	1429.8	2915.3
Heart failure	260.5	420.5	857.4
Heart rhythm disturbances	547.1	883.1	1800.6
Peripheral vascular disease	364.7	588.7	1200.4

Figure 3.1: **Map of the Locations of the 89 Airports in the Contiguous US Included in this Analysis.**

Figure 3.1 (Continued)



centile of noise of one dB was associated with an increase of 0.34% (95% CI 0.02 to 0.68) in the log-relative risk of having a CVD hospitalization. Airport-specific and aggregated relative risks (for Model 3) of having a CVD hospitalization per one dB increase in the 90^{th} percentile of noise exposure are displayed in Figure 3.3.

Figure 3.2: **Overall Estimates (averaged across the 89 airports) of the Relative Risk of Hospitalization for Cardiovascular Disease Associated with a One dB (DNL) Increase in Both Exposure Variables (population-weighted noise exposure and 90th percentile noise exposure) for Each of the Models:** Model 1 controls for individual demographics (age, gender, and race); Model 2 additionally controls for zip-code level socioeconomic status and demographics (% Hispanic and median household income); Model 3 adds to Model 2 by also controlling for fine particulate matter and ozone levels.

Figure 3.2 (Continued)

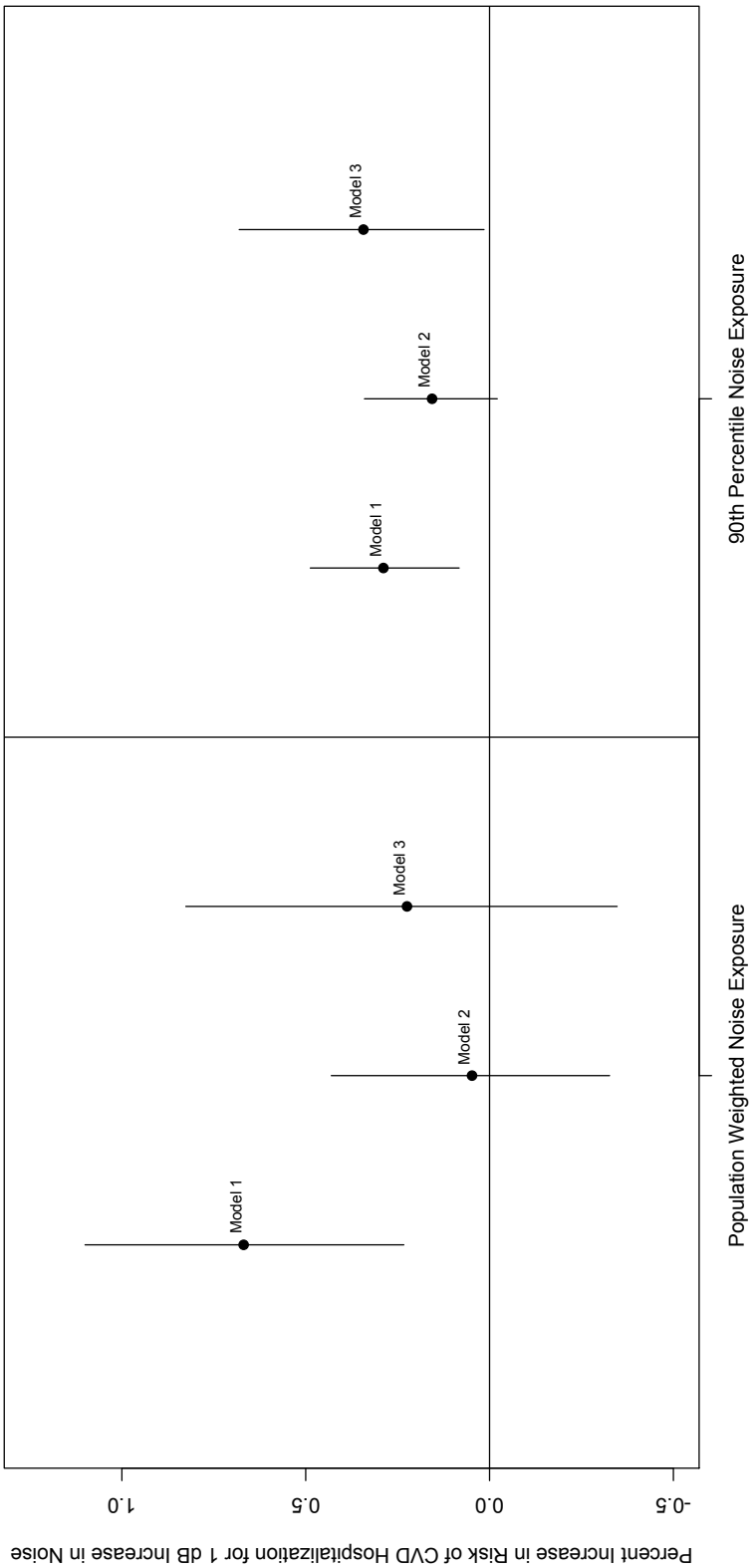


Figure 3.3: **Airport-specific and Overall Estimates of the Relative Risk of Hospitalization for Cardiovascular Disease Associated with a One dB (DNL) Increase in the 90th Percentile Noise Exposure among Census Blocks within ZCTAs:** This model controls for individual demographics (age, gender, and race), zip-code level socioeconomic status and demographics (% Hispanic and median household income), and average yearly fine particulate matter and ozone levels (Model 3). Airport-specific estimates are arranged from lowest to highest estimates.

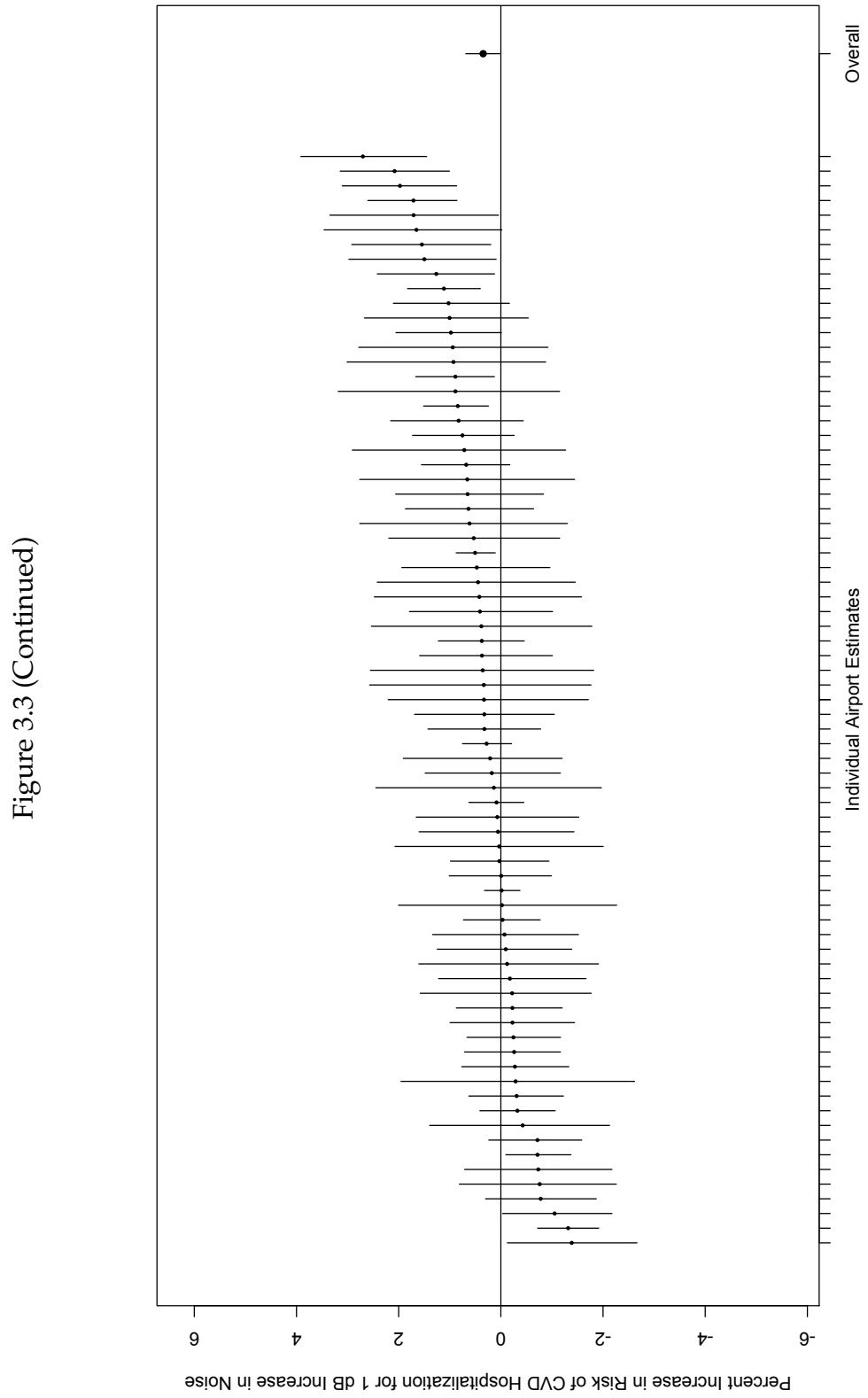


Figure 3.3 (Continued)

In contrast, for the population-weighted noise exposure, while there was an estimated 0.67% increase (95% CI 0.23% to 1.09%) in the overall relative risk of having a CVD hospitalization for a one dB increase in noise in Model 1, after controlling for SES, demographic, and pollution variables (Models 2 and 3), this association was no longer statistically significant. Figure 3.4 shows the airport-specific and aggregate results for Model 3 for population-weighted noise. The standard errors of the airport-specific estimates are generally larger than for the models using the 90^{th} percentile of noise within a ZCTA, potentially due in part to the relatively limited variability of population-weighted noise across ZCTAs within the dataset (Table 3.1).

Figure 3.4: **Airport-specific and Overall Estimates of the Relative Risk of Hospitalization for Cardiovascular Disease Associated with a One dB (DNL) Increase in Population-weighted Noise Exposure among Census Blocks within ZCTAs:** This model controls for individual demographics (age, gender, and race), zip-code level socioeconomic status and demographics (% Hispanic and median household income), and average yearly fine particulate matter and ozone levels (Model 3). Airport-specific estimates are arranged from lowest to highest estimates.

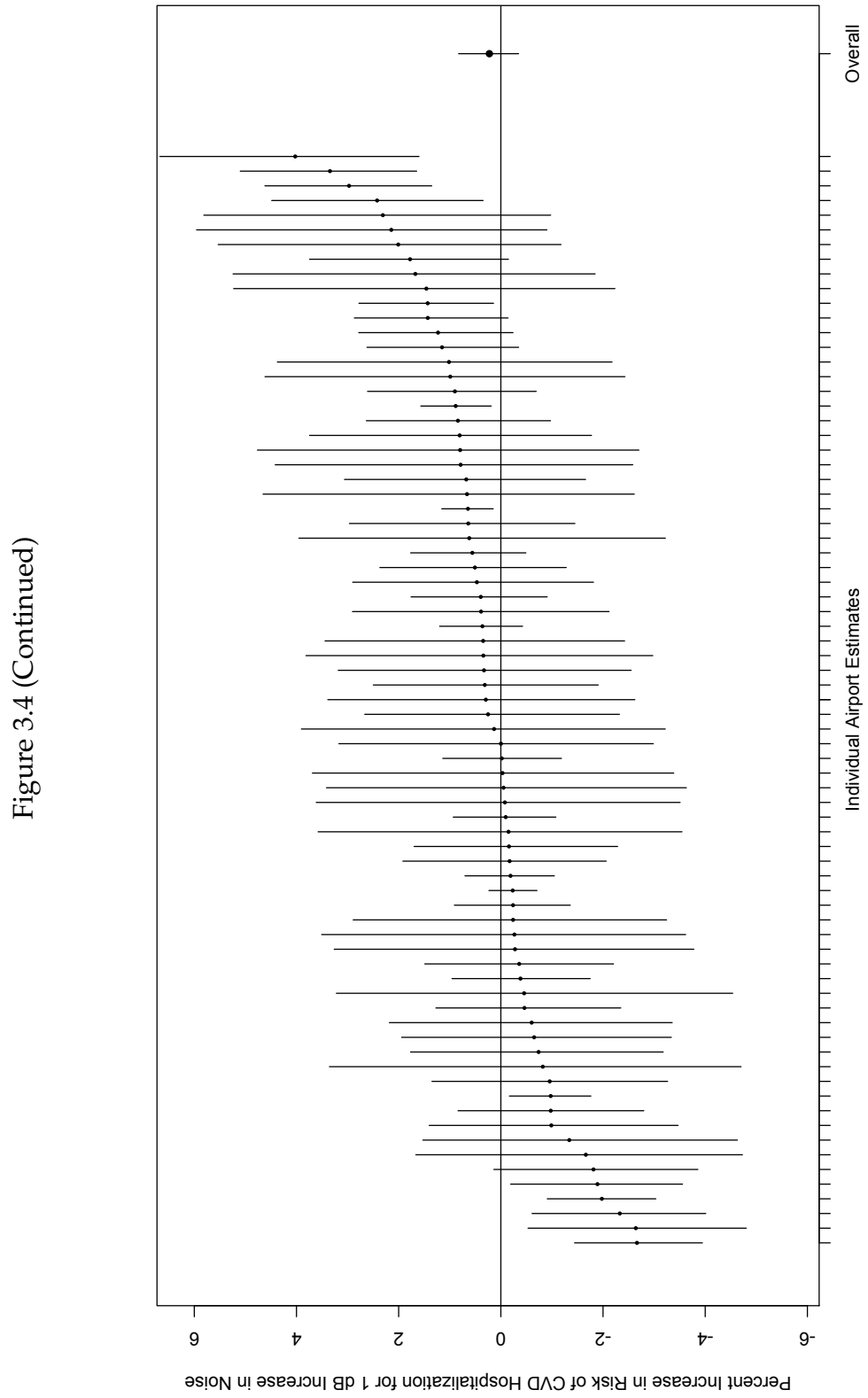
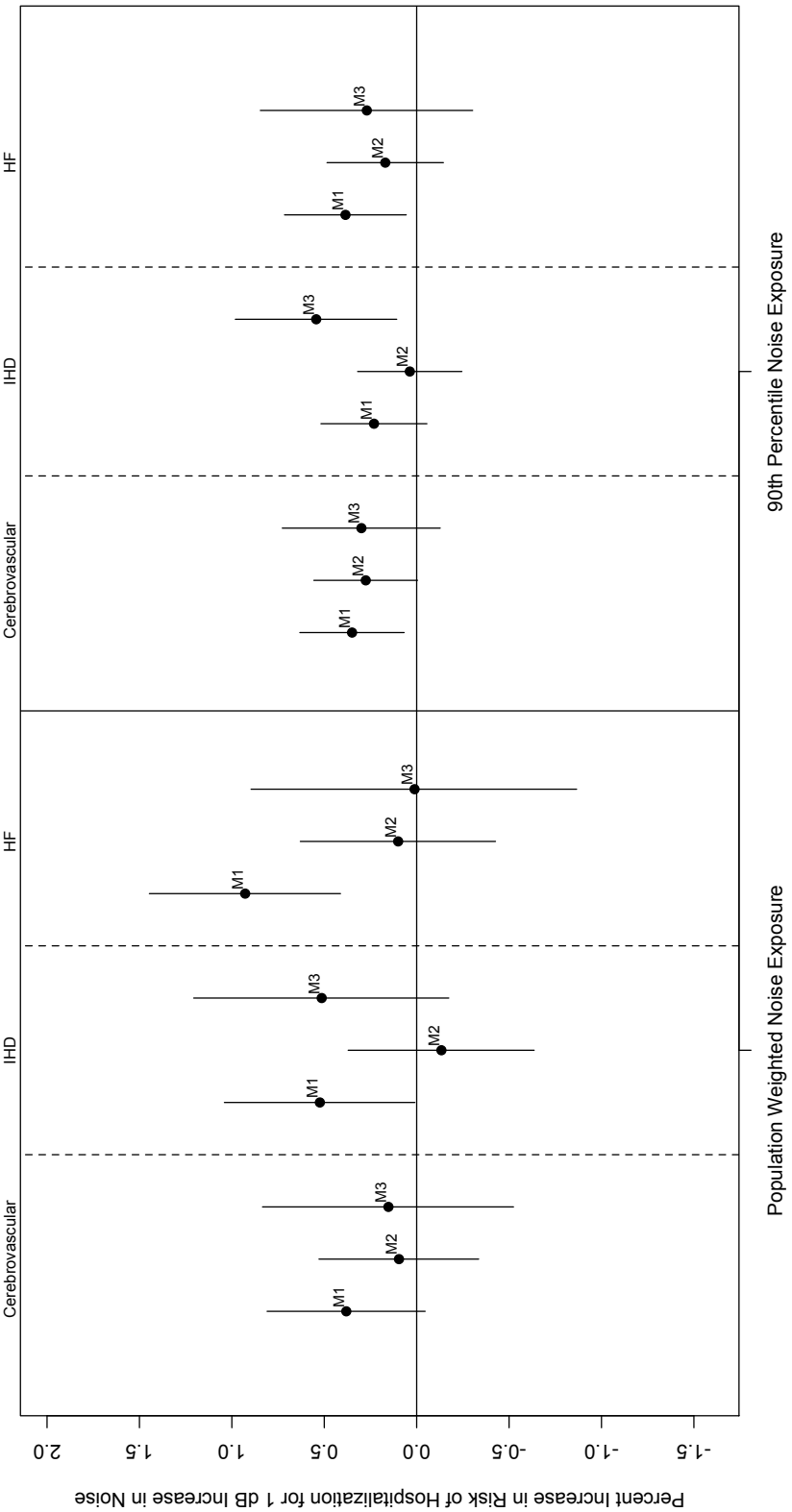


Figure 3.4 (Continued)

Considering sub-categories of CVD outcomes, we observed generally consistent patterns among models. For example, in Model 1, an increase in the 90^{th} percentile of noise of one dB was associated with cerebrovascular disease and HF, with a marginal association for IHD. Relative risk estimates were similar across outcomes (Figure 3.5). For Model 2, relative risk estimates for all three outcomes declined in magnitude and lost statistical significance. Inclusion of pollution variables (Model 3) led to stable or increased relative risk estimates for all three outcomes, relative to Model 2. These estimates lacked statistical significance other than for IHD but were similar in magnitude to the estimates from Model 1. For the population-weighted noise exposure, a similar pattern was observed (Figure 3.5).

Figure 3.5: **Airport-specific and Overall Estimates of the Relative Risk of Hospitalization for Cardiovascular Disease Associated with a One dB (DNL) Increase in Population-weighted Noise Exposure among Census Blocks within ZCTAs:** This model controls for individual demographics (age, gender, and race), zip-code level socioeconomic status and demographics (% Hispanic and median household income), and average yearly fine particulate matter and ozone levels (Model 3). Airport-specific estimates are arranged from lowest to highest estimates.

Figure 3.5 (Continued)



3.4 Sensitivity Analyses

The results outlined in the previous section suggesting a positive and statistically significant association between exposure to noise from airports and CVD hospitalizations among elderly US residents residing in ZCTAs surrounding airports could have a large impact on policy regarding noise mitigation strategies at and around airports in the US. Thus, we felt it was appropriate to perform a number of sensitivity analyses to be sure that the associations observed under Eq. 3.1 were robust to a number of modeling assumptions. Since we saw the strongest associations between noise and CVD hospitalizations using the 90^{th} percentile of noise variable, all models and sensitivity analyses discussed in the section will only consider the 90^{th} percentile noise variable as the exposure.

3.4.1 Residual Spatial Confounding

As always, a primary concern with any observational study is bias due to residual confounding {Christenfeld et al. (2004); Greenland and Morgenstern (2001)}. While we do our best to obtain information on potentially confounding SES and demographic variables from the US Census, the possibility for residual confounding due to some unmeasured covariate remains. Broadly speaking, there are two sources of variability in our dataset and thus two avenues for our effect estimates to be affected by residual confounding: 1) Within airport variation, and 2) Between airport variation. Within airport variability is characterized by variations in the covariates in Eq. 3.1 among the ZCTAs within a particular airport. Between airport variability is characterized by variations in the covariates across all of the airports in our study. Our first goal was to determine which of those two sources of variability contributed more to the estimated association between noise exposure and CVD hospitalizations in order to identify which sources of variation would most likely be responsible if our estimates were indeed biased due to confounding.

We began, as in the previous chapter, by decomposing our noise exposure variable in Eq. 3.1 into a "local" and "global" exposure term. Specifically, we recast Eq. 3.1 as:

$$\begin{aligned}\log(E[Y_{z,s}^A]) &= \log(N_{z,s}^A) + \beta_0^A + \beta_1^A \mathrm{I}(age > 75) + \beta_2^A \mathrm{I}(sex = M) \\ &+ \beta_3^A \mathrm{I}(race = nonwhite) + \beta_4^A (x_z^A - \bar{x}^A) + \beta_5 (\bar{x}^A - \bar{x}) + \boldsymbol{\gamma}^T W_z^A,\end{aligned} \tag{3.4}$$

where the β_4^A estimate the noise-CVD relationship using within airport variability and β_5 estimates the noise-CVD relationship using between airport variability. Under the covariate pattern specified by Model 3 (controlling for SES and air pollution variables), we estimated the overall effect, $\bar{\beta}_4^A = \theta_4$ of Eq. 3.1 to be 0.0034 with a p-value of 0.04. Under that same covariate pattern in Eq. 3.4, we estimate $\bar{\beta}_4^A = \theta_4 = 0.0033$ (p = 0.046) and $\beta_5 = -0.0020$ (p = 0.904), demonstrating that the overall noise effect estimate in Eq. 3.1, θ_4, is largely informed by "local" within airport variation. This result is somewhat intuitive; it tells us that, overall, populations surrounding airports tend to share similar characteristics across all airports, though within any particular airport the populations therein and their exposures to noise are quite varied. Because of this, we focus our attention on within-airport characteristics that could cause our estimates of the noise-CVD association (and the standard errors of those estimates) to be biased.

Given the nature of the data - ZCTAs nested in airports located throughout the US - a natural concern is that of unaccounted for spatial correlation among neighboring ZCTAs in the same airport. While the random effects defined in Eq. 3.1 impose a correlation structure among all ZCTAs that share a common airport, the imposed correlation structure is largely non-spatial as it essentially borrows information equally across *all* ZCTAs surrounding the airport, as opposed to a more localized approach that accounts for the relative closeness of the ZCTAs {Waller and Carlin (2010)}. Thus, an important question to answer first is whether, in fitting a standard Poisson GLM, we observe significant spatial correlation in the residuals. Note that residual spatial correlation would imply 1) that our model is missing some unmeasured covariates that, at least partially, explain how our outcome (total CVD hospitalizations) is correlated across locations {Wakefield (2007)}, and 2) that we are violating the basic GLM assumption of uncorrelated observations {Breslow (1995)}.

To examine these possibilities more closely, we focused individually on five of the larger airports in our dataset, chosen both because of their size (number of surrounding ZCTAs) and because their airport-specific estimates are varied - some positive, some negative. Additionally, because more than half of our ZCTAs are lost due to missingness when we control for $\mathrm{PM}_{2.5}$ and ozone, we focus only on models controlling for SES and demographic variables (Model 2) for the remainder of this section. We loosely proceed as in Wakefield, 2007, Section 3.

We begin by fitting Eq. 3.1 above separately for each of the 5 selected airports under a naive Poisson model, a quasi-Poisson likelihood allowing for overdispersion, and a ZCTA-level random intercept model. Table 3.2 summarizes the estimates and standard errors of the log relative risk of the noise-CVD association at each of the five airports under the Poisson, quasi-Poisson, and random intercept models. Table 3.2 also includes a column indicating the magnitude of spatial correlation among the residuals under the naive Poisson model and whether that correlation is statistically significant. To estimate residual spatial correlation, we use Moran's I test statistic {Moran (1950); Bivand et al. (2008)}, which, while not the only method of testing for statistically significant spatial correlation, is easily the most widely used. {Cliff and Ord (1972)}. For a vector of observations, x, across n different locations, Moran's I is given by:

$$\frac{n}{\sum_i \sum_j w_{ij}} \times \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2},$$

where w_{ij} is the ij^{th} entry in an $n \times n$ adjacency matrix W, and where $w_{ii} = 0$, and $w_{ij} = 1$ if locations i and j are neighbors - commonly defined as locations i and j sharing a common border. While there are other ways of defining W {Cliff and Ord (1972)}, this is a fairly standard approach {Besag et al. (1991); Waller and Carlin (2010)} and is the form we will adopt here as well. Note that Moran's I can take on any value between -1 and 1, and it can be loosely interpreted analogously to Pearson's correlation coefficient.

Two things are very obvious from Table 3.2: 1) There is a considerable amount of

Table 3.2: **Estimates and Standard Errors for β_4^A for Five Airports in our Study under Different Poisson Models:** Moran's I calculated under the naive Poisson model. (*) Indicates that Moran's I is statistically significant.

Airport	Poisson	Poisson Std. Error	Quasi-Poisson	Quasi-Poisson Std. Error	Rand. Int.	Rand. Int. Std. Error	Moran's I
A	0.0142	0.0040	0.0142	0.0054	0.0148	0.0056	0.033
B	0.0086	0.0027	0.0086	0.0042	0.0098	0.0051	0.014
C	-0.0096	0.0013	-0.0096	0.0026	-0.0071	0.0041	0.021
D	-0.0007	0.0015	-0.0007	0.0022	-0.0010	0.0022	-0.094
E	0.0018	0.0014	0.0018	0.0023	-0.0003	0.0034	0.006

overdispersion in each airport - indicated by the much larger standard errors in the quasi-Poisson and random intercept models, and 2) After including noise exposure, individual-level risk factors (age, sex, and race), and SES and demographic variables, there does not appear to be any residual spatial correlation to speak of - indicated by the consistently small (near zero) and non-significant Moran's I statistics. This lack of residual spatial correlation was also confirmed by examining variograms {Diggle et al. (1994); Bivand et al. (2008)} of the residuals for each of the airports, which also displayed no obvious trends. Across all airports, however, there is a high degree of spatial correlation in both the raw ZCTA-level mortality rates as well as the ZCTA-level noise exposures. Thus, at least with regards to local spatial variation, the results in Table 3.2 suggest that our models explain the local mortality rates relatively well.

While the effect estimates across all models for each of the airports are relatively consistent, there is one exception. Notice that the sign of the noise-CVD association flips at Airport E from the Poisson and quasi-Poisson models to the random intercept model. Why might this have happened? Recall that ZCTA-level random intercepts allow for information to be borrowed across all ZCTAs, and they also impose a correlation structure among the observations within a ZCTA. Also recall that there are 8 different age/sex/race strata for each ZCTA, meaning there are 8 different observations for each ZCTA. Though the overdispersion parameter in the quasi-Poisson model is able to broadly handle the extra variability created by having multiple observations in each ZCTA, it is not able to borrow strength globally as the random intercept model can {Waller and Carlin (2010)}.

This suggests that the naive and quasi-Poisson estimate of β_4^A for Airport E was likely inflated due to some number of outlying ZCTAs with a small number of observations that were pulled more closely to the global mean via the inclusion of the random intercept {Waller and Carlin (2010)}. Indeed, upon closer inspection of the data, we see that two of the 3 highest noise exposures at Airport E were found in ZCTAs with two of the highest observed age/sex/race specific mortality rates in Airport E; however, those mortality rates occurred in strata consisting of only 32 and 33 people, respectively. For reference, the median within-ZCTA population of an age/sex/race stratum in Airport E was 268. Because among the three models considered in Table 3.2, the random intercept model appears to be the most flexible in its ability to both account for extra variability and borrow strength globally, we opt to base our inference on the noise-CVD relationship for Airports A - E on the random intercept model.

3.4.2 Modeling Spatial Dependency

Based on the results displayed in Table 3.2, it does not appear as though modeling spatial correlation is necessary. However, it is an interesting exercise to investigate what would happen to our estimates, β_4^A, if we choose to model spatial dependency anyway. Consider the intrinsic conditional autoregressive (ICAR) approach to modeling spatial dependence in a Poisson regression like those we have considered throughout this section. An ICAR model would take the following general form:

$$\log(E[Y_i]) \quad = \quad \log(N_i) + X_i\boldsymbol{\beta} + u_i, \tag{3.5}$$

where u_i are spatially correlated random effects are conditionally defined as follows {Besag et al. (1991)}:

$$u_i \mid u_{j\neq i} \quad \sim \quad \mathcal{N}\left(\frac{\sum_{j\neq i} w_{ij}u_j}{\sum_{j\neq i} w_{ij}}, \frac{1}{\tau_u \sum_{j\neq i} w_{ij}}\right), \tag{3.6}$$

where w_{ij} have the same interpretation as before in the Moran's I statistic, and τ_u is a conditional precision parameter, the magnitude of which determines the amount of spatial variation; a "large" value of τ_u suggests that the estimated residual spatial dependency is small {Wakefield (2007); Hodges and Reich (2010)}. The model induces spatial dependence because it imposes that each u_i is a weighted average of its neighbors, j.

It is recommended that a non-spatial random effect always be included along with the ICAR random effects because the ICAR model cannot take on a limiting form that allows non-spatial variability {Wakefield (2007)}; such a model is known as a *convolution model* {Besag et al. (1991); Waller and Carlin (2010)}, which takes the following general form:

$$\log(E[Y_i]) \quad = \quad \log(N_i) + X_i\boldsymbol{\beta} + u_i + v_i, \tag{3.7}$$

where v_i is a non-spatial random effect given by $v_i \sim \mathcal{N}(0, \tau_v)$, u_i is defined as in Eq. 3.6, and τ_v is precision parameter for the non-spatial random effects, v_i. Fitting these models is straightforward using the freely available `WinBUGS` software, which we use to fit Eq. 3.7 separately for each of Airports A - E. Table 3.3 below summarizes the ICAR results side-by-side with the random intercept results shown in Table 3.2.

Table 3.3: **Estimates and Standard Errors for $\boldsymbol{\beta_4^A}$ for Five Airports in our Study under an ICAR Model and a Random Intercept Model:** Moran's I calculated under the ICAR model. $(^*)$ Indicates that Moran's I is statistically significant.

Airport	ICAR	ICAR Std. Error	Rand. Int.	Rand. Int. Std. Error	τ_u	τ_v	Moran's I
A	0.0151	0.0083	0.0148	0.0056	158.67	64.71	-0.247*
B	0.0087	0.0057	0.0098	0.0051	172.18	91.70	0.023
C	-0.0028	0.0045	-0.0071	0.0041	64.75	97.61	0.004
D	-0.0019	0.0034	-0.0010	0.0022	262.92	150.50	0.096
E	-0.0027	0.0041	-0.0003	0.0034	43.58	116.80	0.081

It is interesting to see which effect estimates and standard errors change the most under the ICAR model. Effect estimates in Airports A, B, and D are the most consistent between the ICAR and random intercept models. This is expected, as Hodges and Reich (2010) show that if the ratio τ_u/τ_v is large, then it is unlikely that the spatially correlated random effects will bias the parameter estimates. On the other hand, for Airports C and E, we see that the noise-CVD effect estimates change quite a bit. This, however, should not be unexpected as τ_u/τ_v is relatively small compared to the other airports {Hodges and Reich (2010)}.

Moran's I statistic calculated for the residuals of the ICAR model at each airport actually *increased* in magnitude across 4 of the 5 airports. Most notably, we now observe a large, statistically significant negative correlation in the residuals at Airport A. Why this happens is not clear, particularly since τ_u is quite large, suggesting little spatial dependence in the data. Interestingly, though, the standard error of the ICAR estimate for Airport A was inflated more than that of any other airport, which - together with the residual spatial correlation - perhaps suggests some sort of over-adjustment for spatial dependence.

For completeness, we also fit the convolution model and the random intercept model to the entire dataset for all 89 airports at once. Specifically, we compared the following:

$$\begin{aligned} \log(E[Y_{z,s}^A]) &= \log(N_{z,s}^A) + \beta_0 + \beta_1 \mathrm{I}(age > 75) + \beta_2 \mathrm{I}(sex = M) \\ &+ \beta_3 \mathrm{I}(race = nonwhite) + \beta_4 (x_z^A - \bar{x}) + \boldsymbol{\gamma}^T W_z^A + v_z + u_z, \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} \log(E[Y_{z,s}^A]) &= \log(N_{z,s}^A) + \beta_0^A + \beta_1^A \mathrm{I}(age > 75) + \beta_2^A \mathrm{I}(sex = M) \\ &+ \beta_3^A \mathrm{I}(race = nonwhite) + \beta_4^A (x_z^A - \bar{x}) + \boldsymbol{\gamma}^T W_z^A + v_z, \end{aligned} \tag{3.9}$$

where v_z and u_z are defined as they were previously. Note that we do not specify airport-specific terms for $\beta_0, \ldots, \beta_4$ in the convolution model (Eq. 3.8), as attempts to do so yielded unreasonably inflated standard errors, likely as a result of overfitting with regards to modeling the spatial dependence within airports. We do, however, specify the airport-specific effects in the random intercept model (Eq. 3.9), as this allows us to impose a broad correlation structure on all ZCTAs surrounding the same airport.

In estimating the overall effect of noise on CVD hospital admissions, the results for Eqs. 3.8 and 3.9 were quite similar. In the convolution model, we estimated log relative risk of 0.0011 (se = 0.0011, CI = [-0.0010, 0.0031]) for a 1 dB increase in noise on CVD hospitalization rates. Similarly, in the ZCTA-level random intercept model, we estimated a log relative risk of 0.0010 (se = 0.0009, CI = [-0.0007, 0.0027]). These are actually relatively consistent with the overall estimate from Model 2 in our main analysis, which yielded a log relative risk of 0.0016 (CI = [-0.0002, 0.0034]). This is encouraging, as it suggests that the original models specified by Eq. 3.1 were not severely biased due to some unaccounted for residual spatial correlation. However, because we observed significant overdispersion locally at the five airports in our sensitivity analysis, it appears as though Eq. 3.9 is preferable to Eq. 3.1. Further, because of the inconsistencies that can arise locally when specifying the ICAR model, we believe that Eq. 3.9 is preferable Eq. 3.8 as well.

Generally speaking, the results from this section are relatively consistent with previous research done in this area. Despite the substantial amount of work required to set up and fit a model with spatially correlated random effects, the results from such a model are not vastly different than a well thought-out and executed non-spatial random-effects model that place an emphasis on obtaining the appropriate variables to control for confounding. Indeed, Wakefield (2003) argues that in spatial regression studies, more effort should be placed on confounding/within-area modeling than spatial dependence, as the latter will be of secondary importance. Our results are also consistent with Hodges and Reich (2010) and Reich et al. (2006), who show that it is a fallacy that including spatially correlated random effects will "control for unmeasured spatial confounding," and that the

inclusion of spatially correlated random effects will "remove bias from the estimate of the parameter of interest, β, and generally be conservative." We saw this to be true in our sub analyses of the five airports, which confirmed the notion that the spatially correlated random effects only haphazardly adjust the estimate of β - particularly when τ_u/τ_v is "small" - and are not necessarily conservative {Hodges and Reich (2010)}. We conclude this section as did Wakefield (2007): "For spatial regression fitting, an appropriate mean model is more important than the choice of any particular inferential paradigm."

3.4.3 Threshold Effects and Non-linearity in the Exposure Response Curve

We additionally performed basic analyses to determine whether there might be a threshold for the effect of noise on CVD hospitalizations. Briefly, to examine the existence of a threshold, we fit Eq. 3.9 above, but with noise exposure categorized as being in the lower 33^{rd} percentile, the middle 33^{rd} percentile, or the upper 33^{rd} percentile. Specifically, the model takes the form:

$$\begin{aligned} \log(E[Y^A_{z,s}]) &= \log(N^A_{z,s}) + \beta^A_0 + \beta^A_1 \mathrm{I}(age > 75) + \beta^A_2 \mathrm{I}(sex = M) + \beta^A_3 \mathrm{I}(race = nonwhite) \\ &+ \beta^A_4 I(x^A_z \in [x^{(33)}, x^{(66)}]) + \beta^A_5 I(x^A_z > x^{(66)}) + \boldsymbol{\gamma}^T W^A_z + v_z, \end{aligned} \tag{3.10}$$

where $x^{(33)}$ and $x^{(66)}$ denote the 33^{rd} and 66^{th} percentiles, respectively, of the noise exposure distribution; the observations for which $x^A_z < x^{(33)}$ act as the reference group. Denote the reference group as "low exposure," the middle 33^{rd} percentile as "medium exposure," and the upper 33^{rd} percentile as "high exposure." We estimated β_5, the overall log relative risk of CVD hospitalization associated with being in a high exposure ZCTA compared to a low exposure ZCTA to be 0.0199 (se = 0.0097, CI = [0.0009, 0.0389]), and β_4, the log relative risk of CVD hospitalization associated with being in a medium exposure ZCTA compared to a low exposure ZCTA, was estimated to be 0.0066 (se = 0.0104, CI = [-0.0138, 0.0270]). This suggests that there is likely a threshold below which

there is no effect of noise on CVD hospitalizations, as we are unable to detect a significant difference in the relative risk of CVD hospitalizations between low and medium exposure ZCTAs, though we observe a large difference between low exposure ZCTAs and high exposure ZCTAs.

The fact that a threshold likely exists also suggests that the exposure response curve is probably not linear, as it is less likely to be true that after a certain point the exposure response curve switches from zero to some linear relationship with a constant slope than it is that the true relationship gradually and smoothly changes as noise exposure increases. More work is needed to investigate both of these issues.

3.5 Discussion

We evaluated the relationship between residential exposure to aircraft noise and hospitalization for CVD in older US adults. In models only controlling for individual demographics, we found that this association is positive and statistically significant using both of our noise exposure metrics. Furthermore, the positive association generally persisted when accounting for ZCTA-level SES and demographic variables, and regional air pollution - particularly for the 90^{th} percentile of noise exposure variable. Positive associations were also observed for individual cardiovascular hospitalization outcomes, but statistical power was reduced and relationships were often attenuated after controlling for demographic and air pollution variables.

We observed heterogeneity in the relationship between ambient aircraft noise and cardiovascular hospitalization across airports, consistent with prior research {Floud et al. (2011)}. As proposed elsewhere {Floud et al. (2011)}, our observed heterogeneity may reflect differences across the country in sound transmission from outdoors to indoors (where most exposure would be anticipated to occur). This could include structural attributes of the housing stock or the frequency of open windows. Airport-specific differences may also reflect different degrees of soundproofing that have occurred around

various airports in terms of the type of soundproofing program, the area (radii) covered, and the time since soundproofing, as these could influence personal exposures and spatial patterns of risk. In addition, Job and Hatfield (2001) proposed that the "soundscape" (presence of other sources of noise), "enviroscape" (physical environment in which noise occurs) and "psychscape" (individual psychology) can influence reaction to noise, leading to differences in health effects {Hatfield et al. (2001)}.

Our findings add to previous literature in several key ways. First, we investigated the noise-cardiovascular hospitalization relationship across gradients of airport noise exposure levels across the largest number of airports studied to date, using statistical techniques that allowed us to estimate airport-specific effects while maximizing information from each airport for a pooled estimate. Second, we leveraged administrative data capturing the majority of older US adults, who represent an age group at greater risk for CVD. To our knowledge, this is also the largest studied population of US elderly living near airports to date. We thus had a large number of events, increasing our power to detect relationships. Third, we evaluated the relation of noise with cardiovascular hospitalization as the outcome, which, to our knowledge, has been rarely considered in previous noise studies. An ecological study of 62 municipalities around an airport in Amsterdam found no clustering of cardiovascular hospitalizations in areas close to the airport {Babisch (2006)}, but we improve on this study by assessing the relation for individual at-risk subjects. Fourth, we accounted for the potential confounding of air pollution.

The estimated associations of similar magnitude across several CVD-specific outcomes are broadly consistent with the literature. For example, in areas with more aircraft noise, there were more subjects under medical treatment for heart trouble and with "pathological heart shape" {Knipschild (1977)}. A 2009 review of epidemiological studies found sufficient evidence of positive relationships among aircraft noise and high blood pressure and cardiovascular medication use {Babisch and Kamp (2009)}. One study included in this review investigated the relationship between aircraft noise and incidence of hypertension and found a positive association, particularly in older subjects {Eriks-

son et al. (2007)}. Hypertension is not typically a primary reason for hospital admission, so it was not specifically included in our analyses, but hypertension is associated with multiple cardiovascular sequelae that would contribute to hospitalizations.

In addition, although aircraft-related noise has a different profile from traffic-related noise, our findings are consistent with the noise-CVD health literature. For example, in models controlling for individual characteristics, ZCTA-level SES and demographics, and air pollution, we found the strongest association (positive and statistically significant) with IHD hospitalizations, consistent with conclusions of an expert report regarding likely mechanisms of noise-related health effects {Babisch and Kim (2011)}. Our findings were also consistent with studies looking jointly at noise and air pollution. For example, Beelen et al. (2009) found excess cardiovascular mortality in the highest category of noise, which was reduced slightly after controlling for air pollution. Huss et al. (2010) found that the association between aircraft noise and mortality from MI was not attenuated with adjustment for air pollution. de Kluizenaar et al. (2007) found that after controlling for PM_{10}, the relationship between road traffic noise and hypertension became marginally significant. We found that controlling for air pollution increased the relative risk for both noise exposure metrics (relative to Model 2).

There are multiple limitations with our analysis, related in part to limitations with the population database. Although using the Medicare data allowed us to cover nearly the entire US elderly population, this database was developed for administrative purposes and has been shown to be subject to misclassification {Losina et al. (2003); Kiyota et al. (2004)} and geographic variability in evaluation and management {Havranek et al. (2004); Baicker et al. (2004)}. Geographic variation has been linked to characteristics of providers and hospital variation {Havranek et al. (2004)}. It is conceivable (although unlikely) that these characteristics may also relate to physical and social factors such as housing stock and community organization and involvement around noise mitigation and thus, they might contribute to heterogeneity in the noise-cardiovascular hospitalization relationship. We only used primary diagnosis, which should reduce misclassification

of outcomes {Dominici et al. (2006)}, and our analyses of combined CVD outcomes are unlikely to have significant misclassification. Previous studies investigating the relation of particulate pollution and cardiovascular hospitalization found no evidence of effect modification by underlying diagnosis rates; however, this test was not performed specific to our noise-related risk estimates.

Other limitations of the Medicare data include limited individual data on risk factors. For example, we were not able to control for smoking and diet, strong risk factors for CVD. These variables would only confound the noise-CVD hospitalization association if there were significant correlations between noise exposures and these risk factors. Noise contours, however, display fairly sharp gradients and skew as a function of prevailing wind directions, given runway orientation and arrival/departure patterns. Thus, we believe behavioral characteristics would unlikely be similarly patterned in space, unless property values were strongly tied to airport noise and socioeconomic patterns therefore followed noise contours. We are, however, able to control for a number of ZCTA-level SES and demographic variables - among which was median household income - allowing us to broadly control for potential confounding due to socioeconomic patterns. Our estimates were generally robust to area-level SES covariates, but we lacked the individual-level addresses and SES characteristics to formally address this question. In addition, our ZCTA-level SES and demographic variables were taken from Census 2000 data because only limited SES information from Census 2010 was available at the ZCTA-level at the time of our analysis. We thus assumed that patterns of ZCTA-level SES typically remained similar over that time. More generally, the availability of only ZCTA-level address information can lead to exposure misclassification. Noise gradients are substantial at close proximity to airports, and we were unable to differentiate among individuals' noise exposure within zip codes. However, the use of a study population closely aligned with census data (given near-universal enrollment in Medicare) allowed us to reasonably estimate a representative ZCTA-resolution population exposure, with error most likely to be Berksonian with unbiased regression coefficients and inflated standard errors.

Using INM to predict noise exposure also has limitations. INM uses average annual input conditions. Therefore, values may lack precision because certain local acoustical variables, such as humidity effects, ground absorption, individual aircraft directivity patterns and sound diffraction around terrain or buildings, are not averaged or may not be explicitly modeled {Federal Aviation Administration - Office of Environment and Energy (2007)}. That said, INM is well-established internationally {Eriksson et al. (2010)}, and is the required noise assessment tool in the US for airport noise compatibility planning and environmental assessments and impact statements. Each of our derived exposure metrics had its own inherent limitations, with the population-weighted average potentially reducing the contrast between ZCTAs and the 90th percentile of noise exposure not capturing the exposure profile of the entire ZCTA.

We were, however, able to conduct a number of sensitivity analyses to test the robustness of our results to various statistical assumptions. In particular, we showed that 1) the majority of the information responsible for estimating the relationship between exposure to noise and CVD hospitalization rates comes from variability within airports, with little contribution from between-airport variability; 2) despite the fact that most of the variability comes from within-airport variation, our estimates do not appear to biased as a result of residual spatial confounding, as our airport-specific residuals did not exhibit significant spatial correlation; 3) even if we do choose to include spatially correlated errors in our model, though airport-specific estimates could potentially be biased, the overall estimate of the association between noise exposure and CVD hospitalization rate was relatively stable across various modeling assumptions; and 4) initial exploratory work into the possibility of a threshold below which there is no significant association between CVD hospitalizations and noise exposure suggests that a threshold may exist.

In summary, despite some data-related limitations, we found that airport noise, particularly characterized by the 90th percentile of noise exposure among census blocks within ZCTAs, is statistically significantly associated with higher relative risk of cardiovascular hospitalization among older subjects. This relationship remained after control-

ling for individual data, ZCTA-level SES and demographics, and air pollution variables. Our results provide evidence of a potential adverse effect of airport noise on cardiovascular health, and further research should refine these associations and strengthen causal interpretation by investigating modifying factors at the airport or individual level.

Appendix

Appendix A: Life Expectancy Calculation

County-level life expectancies in Dataset 1 were calculated by applying a mixed-effects spatial Poisson model to mortality data from the National Center for Health Statistics (NCHS) and population data from the U.S. Census to obtain robust estimates of the number of deaths in each county {Kulkarni et al. (2011)}. These estimated counts are then used to calculate county life expectancies using standard life table techniques, which we discuss in more detail below. Specifically, the model is given by:

$$\log E(y_{rjt}) = \beta_0 + \beta_1 income_{jt} + \beta_2 education_j + \beta_3 \sigma_j^{post} + \beta_4 race + \gamma_j t + \mu_j \tag{4.1}$$

where y_{rjt} is the death count for race r within county j in year t; $income_{jt}$ is county per-capita income for year t; $education_j$ is the percent of adults within county j having completed high school in the 2000 census data; and race is a dummy variable for three race groups (white, black, and other). σ_j^{post} is a geospatial component, calculated as the average of the posterior mode of the county random intercept for counties adjacent to county j to account for residual spatial patterns, the values of which were derived from first running as a prior step the same model above without the geospatial component to derive the posterior values of the county random effect. Similarly, μ_j is the posterior value of the county random intercept. Lastly, γ_j is a random slope on time, t, for each county {Kulkarni et al. (2011)}.

These estimated counts, which are more robust due to the borrowing of information across space and time, are then used to calculate county life expectancies using standard life table techniques. Briefly, the estimated death counts within each age stratum, or interval, are divided by the mid-year population in that interval, providing us with an age-specific death rate for each age interval. This age-specific death rate is then used

together with a term that estimates the average number of years lived by persons who die in each particular age interval - a term which, when expressed as a fraction, is often estimated as $\frac{1}{2}$ - to estimate the probability of dying within each age interval. Using this age interval-specific probability of dying, one can then project the mortality experience of a hypothetical cohort that experiences the same age-specific probabilities of dying as our observed population. This is done for all counties. More details on calculating life expectancy in the life table setting are available in various texts {Chiang (1968); Keyfitz and Flieger (1990); Preston et al. (2001)}. Note then, that the hierarchical model only provides more robust estimates of the death counts within each age interval for each county, which often times can be small and unstable for smaller counties. The actual process of calculating life expectancy in a life table setting does not change.

Appendix B: Datasets 2 and 3 - Variables and Data Sources

The variables in Dataset 2 (211 counties, 1980 - 2000) were: life expectancy, $\mathrm{PM}_{2.5}$, per capita income, population, and proportions of the population who were high school graduates, who had not lived in that county 5 years earlier (5-year in-migration), who had an urban residence, and who reported they were white, black, or Hispanic. Age-standardized death rates for lung cancer and chronic obstructive pulmonary disease (COPD) were included in the dataset to account for smoking prevalence in the population. Each variable had a value for both 1980 and 2000. This data and its sources are described in more detail elsewhere {Pope et al. (2009)}.

The variables in Dataset 3 (211 counties, 2007) were the same as those in Dataset 2, and the data sources for these data were identical to those of the 545 county dataset with two exceptions: 1) proportion of the population that did not live in the county 5 years earlier, and 2) proportion of the population with an urban residence; these two variables are only available from the decennial Census, so we used year 2000 values as a proxy for 2007. Additionally, as in Dataset 1, due to the availability of NCHS data, 2005 death rates were used as a proxy for 2007.

Yearly average $\mathrm{PM}_{2.5}$ for 2007 was calculated at the MSA level by averaging the yearly county-level $\mathrm{PM}_{2.5}$ readings for all counties in a given MSA. We calculated both population-weighted and non-weighted averages. Combining Datasets 2 and 3 enabled us to extend the analysis in Pope et al5 to the periods 1980 - 2007, and 2000 - 2007. When we exclusively analyzed the 211 counties in Pope et al. (2009) regardless of the time period, we did so with $\mathrm{PM}_{2.5}$ calculated at the MSA level for all counties, consistent with the original analysis. We also note that per capita income in Dataset 2 was obtained from the U.S. Census, while per capita income in Dataset 3 was obtained from the Bureau of Economic Analysis (BEA). BEA per capita income estimates were consistently higher than Census estimates, thus, for consistency we also obtained BEA per capita income estimates for 1980 and 2000, and results for the 211 counties from Pope et al. (2009) for the periods 1980 - 2007 and 2000 - 2007 were obtained using BEA per capita income estimates. When re-analyzing Dataset 2 (1980 - 2000), we obtained results using Census per capita income estimates as in Pope et al. (2009) and also using BEA per capita income estimates. When adjusting for changes in per capita income, the effect of $\mathrm{PM}_{2.5}$ on life expectancy was not sensitive to the choice of the income variable.

We additionally note that the estimated counts used to calculate life expectancy in Dataset 3 for the year 2007 (described above in Section A) were calculated using a slightly different method than the one used to calculate the estimated death counts used to calculate life expectancy for the 211 counties in Pope et al. (2009) for the periods 1980 and 2000 (Dataset 2) {Kulkarni et al. (2011); Ezzati et al. (2008)}. However, the two methods are only substantially different in locations with very small populations (pop < 7000) {Kulkarni et al. (2011)}, which is not the case here as all of these counties are in metropolitan areas, and no counties had a year 2007 population less than 22,000. For the year 2000, where we have life expectancy estimates for the 211 counties using both methods, the correlation between the two was greater than 0.98.

Appendix C: Results from Cross-sectional Analyses

For Dataset 1 (545 counties, 2000 - 2007), simple models including only $\mathrm{PM}_{2.5}$ as a predictor estimated an increase in life expectancy of 2.09 ± 0.19 and 2.63 ± 0.28 years for a $10\mu\mathrm{g/m}^3$ decrease in $\mathrm{PM}_{2.5}$ in 2000 and 2007, respectively (p < 0.001 for both). Models controlling for population, per capita income, proportion of the population that is black or Hispanic, and death rates for lung cancer and COPD showed markedly smaller associations, with $\mathrm{PM}_{2.5}$ estimates of 0.33 ± 0.11 (p = 0.005) and 0.39 ± 0.17 (p = 0.021) years for 2000 and 2007, respectively.

Similarly, a cross-sectional analysis of the year 2007 of the 211 counties in Pope et al. (2009) gave a simple estimate of 2.80 ± 0.64 (p < 0.001). $\mathrm{PM}_{2.5}$ effects were attenuated when controlling for population, per capita income, proportion of the population that was black or Hispanic, and death rates for lung cancer and COPD (estimate = 0.30 ± 0.38; p = 0.44). Cross-sectional analyses for the 211 counties for the years 1980 and 2000 were no different than originally reported {Pope et al. (2009)}. In all datasets, however, additionally controlling for proportion of the population who are high school graduates shrank estimates of the effect of $\mathrm{PM}_{2.5}$ towards zero, and yielded much higher p-values (0.200 < p < 0.946).

Appendix D: Additional Tables

Table 4.1: **Summary Characteristics of the 211 Counties Analyzed for the Years 1980 to 2007:** $(^*)$, 2005 death rates are used as a proxy for 2007 death rates. † Indicates values from the 2000 Census are used as a proxy for 2007 values.

Variable	Mean(SD)
Life Expectancy (yr.)	
1980	74.32 (1.52)
2007	78.12 (1.86)
Change	3.80 (1.21)
$PM_{2.5}$ ($\mu g/m^3$)	
1980	20.62 (4.36)
2007	12.44 (2.17)
Reduction	8.18 (3.00)
Per Capita Income (in thousands of $)	
1980	20.39 (3.65)
2007	33.64 (8.58)
Change	13.25 (5.68)
Population (in hundreds of thousands)	
1980	3.83 (8.47)
2007	5.17 (10.49)
Change	1.34 (2.91)
5-year In-migration (% of pop.)	
1980	0.25 (0.10)
2007†	0.24 (0.08)
Change	-0.01 (0.06)
Urban Residences (% of pop.)	
1980	0.58 (0.33)
2007†	0.78 (0.22)
Change	0.20 (0.18)
HS Graduates (% of pop.)	
1980	0.68 (0.11)
2007	0.87 (0.05)
Change	0.19 (0.14)
Black Population (% of pop.)	
1980	0.097 (0.118)
2007	0.116 (0.128)
Change	0.019 (0.069)
Hispanic Population (% of pop.)	
1980	0.035 (0.072)
2007	0.088 (0.101)
Change	0.053 (0.053)
Deaths from Lung Cancer (no./ 10,000 pop.)*	
1980	14.38 (2.95)
2007	15.25 (3.37)
Change	0.87 (3.27)
Deaths from COPD (no./ 10,000 pop.)*	
1980	7.92 (1.85)
2007	11.99 (3.24)
Change	4.07 (3.13)

Table 4.2: **Summary Characteristics of the 211 Counties Analyzed for the Years 2000 to 2007:** $(^*)$, 2005 death rates are used as a proxy for 2007 death rates.

Variable	Mean(SD)
Life Expectancy (yr.)	
2000	77.04 (1.82)
2007	78.12 (1.86)
Change	1.08 (0.64)
$PM_{2.5}$ ($\mu g/m^3$)	
2000	14.10 (2.86)
2007	12.44 (2.17)
Reduction	1.67 (1.25)
Per Capita Income (in thousands of $)	
2000	31.69 (8.01)
2007	33.64 (8.58)
Change	1.95 (2.70)
Population (in hundreds of thousands)	
2000	4.82 (10.13)
2007	5.17 (10.49)
Change	0.35 (0.79)
HS Graduates (% of pop.)	
2000	0.869 (0.050)
2007	0.875 (0.046)
Change	0.006 (0.015)
Black Population (% of pop.)	
2000	0.115 (0.130)
2007	0.116 (0.128)
Change	0.001 (0.028)
Hispanic Population (% of pop.)	
2000	0.068 (0.093)
2007	0.088 (0.101)
Change	0.019 (0.016)
Deaths from Lung Cancer (no./ 10,000 pop.)*	
2000	16.73 (3.27)
2007	15.25 (3.37)
Change	-1.48 (1.96)
Deaths from COPD (no./ 10,000 pop.)*	
2000	12.37 (2.71)
2007	11.99 (3.24)
Change	-0.38 (2.15)

Table 4.3: **Results of Selected Regression Models for MSA-Level Analysis, 1980 - 2007:** (a), Included only counties with a 1986 population $\geq$ 100,000; (b), Included only counties with the largest 1986 population in the MSA. Changes in LC ASDR and COPD ASDR are changes in the age standardized death rate for lung cancer and chronic obstructive pulmonary disease, respectively.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6[a]	Model 7[b]	Model 8[b]
Intercept	3.01±0.32	1.38±0.23	2.58±0.34	2.56±0.35	2.58±0.35	2.62±0.42	1.69±0.70	1.27±0.68
Reduction in $PM_{2.5}$	0.99±0.39	0.83±0.20	0.57±0.19	0.64±0.18	0.56±0.19	0.60±0.22	1.02±0.36	1.13±0.36
Change in income	–	0.10±0.02	0.08±0.01	0.09±0.01	0.09±0.01	0.08±0.01	0.15±0.03	0.13±0.02
Change in pop.	–	0.09±0.04	0.06±0.03	0.07±0.02	0.06±0.02	0.06±0.03	0.04±0.03	0.03±0.03
Change in mig. rate	–	3.27±1.13	3.66±1.00	–	3.57±0.89	5.45±1.80	3.95±2.35	4.97±2.27
Change in urban rate	–	-0.31±0.30	-0.10±0.29	–	–	–	-2.80±2.64	–
Change in HS%	–	1.25±0.83	0.19±0.65	–	–	–	-1.79±1.30	–
Change in black%	–	-1.97±0.76	-3.05±0.66	-3.29±0.74	-3.02±0.62	-3.71±1.21	-8.79±2.74	-8.80±2.79
Change in Hisp%	–	2.66±1.48	1.86±1.12	–	1.85±1.11	1.97±1.28	2.32±2.57	4.13±2.13
Change in LC ASDR	–	–	-0.07±0.03	-0.06±0.03	-0.07±0.03	-0.11±0.03	-0.09±0.05	-0.10±0.05
Change in COPD ASDR	–	–	-0.10±0.04	-0.10±0.03	-0.10±0.04	-0.06±0.04	-0.01±0.07	0.01±0.07
No. of county units	211	211	211	211	211	127	51	51

Table 4.4: **Results of Selected Regression Models for MSA-Level Analysis, 2000 - 2007:** (a), Included only counties with a 1986 population $\geq$ 100,000; (b), Included only counties with the largest 1986 population in the MSA. Changes in LC ASDR and COPD ASDR are changes in the age standardized death rate for lung cancer and chronic obstructive pulmonary disease, respectively.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6[a]	Model 7[b]	Model 8[b]
Intercept	0.98±0.05	0.84±0.10	0.83±0.09	0.79±0.07	0.83±0.05	0.83±0.08	0.94±0.19	0.81±0.12
Reduction in $PM_{2.5}$	0.66±0.34	1.28±0.36	1.09±0.35	1.05±0.36	1.00±0.32	1.58±0.48	1.55±0.54	1.60±0.53
Change in income	–	0.02±0.02	0.02±0.02	0.02±0.02	–	–	-0.03±0.03	–
Change in pop.	–	0.14±0.09	0.13±0.08	0.13±0.07	0.13±0.07	0.08±0.06	0.05±0.06	0.08±0.05
Change in HS%	–	-4.90±3.15	-3.47±3.03	–	–	–	-12.35±7.05	–
Change in black%	–	-10.06±2.28	-9.69±2.14	-9.67±2.07	-10.28±2.43	-20.35±3.90	-33.15±5.46	-29.32±5.13
Change in Hisp%	–	0.19±2.75	-1.26±3.03	–	–	–	-3.09±5.88	–
Change in LC ASDR	–	–	-0.03±0.02	-0.03±0.02	-0.03±0.02	-0.00±0.05	0.04±0.05	0.03±0.05
Change in COPD ASDR	–	–	-0.05±0.02	-0.05±0.02	-0.05±0.02	-0.07±0.05	-0.03±0.05	-0.03±0.05
No. of county units	211	211	211	211	211	127	51	51

References

ALMON, S. (1965). The distributed lag between capital appropriations and expenditures. *Econometrica* **33** 178–196.

BABISCH, W. (2006). Transportation noise and cardiovascular risk: updated review and synthesis of epidemiological studies indicate that the evidence has increased. *Noise & Health* **8** 1 – 29.

BABISCH, W. and KAMP, I. (2009). Exposure-response relationship of the association between aircraft noise and the risk of hypertension. *Noise & Health* **11** 161 – 168.

BABISCH, W. and KIM, R. (2011). Environmental noise and cardiovascular disease. In *WHO European Centre for Environmental Health*. Burden of disease from environmental noise: Quantification of healthy life years lost in Europe, Copenhagen: World Health Organization, 15 – 44.

BAICKER, K., CHANDRA, A., SKINNER, J. S. and WENNBERG, J. E. (2004). Who you are and where you live: How race and geography affect the treatment of medicare beneficiaries. *Health Affairs* **23** 33 – 44.

BATES, D. (2012). Linear mixed model implementation in lme4. Tech. rep., University of Wisconsin - Madison.

BEELEN, R., HOEK, G., HOUTHUIJS, D., VAN DEN BRANDT, P. A., GOLDBOHM, R. A. and FISCHER, P. (2009). The joint association of air pollution and noise from road traffic with cardiovascular mortality in a cohort study. *Occupational and Environmental Medicine* **66** 243 – 250.

BESAG, J., YORK, J. and MOLLIÉ, A. (1991). Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistics and Mathematics* **43** 1 – 59.

BIVAND, R. S., PEBESMA, E. J. and GOMÉZ-RUBIO, V. (2008). *Applied Spatial Data Analysis with R*. Springer.

BLUHM, G. and ERIKSSON, C. (2011). Cardiovascular effects of environmental noise: research in Sweden. *Noise & Health* **13** 212 – 216.

BRESLOW, N. E. (1995). Generalized linear models: Checking assumptions and strengthening conclusions. Tech. rep., Congresso Nazionale, Societa' Italiana di Biometria.

BRUNEKREEF, B. (1997). Air pollution and life expectancy: is there a relation? *British Medical Journal* **54** 781 – 784.

BUJA, A., HASTIE, T. and TIBSHIRANI, R. (1989). Linear smoothers and additive models. *The Annals of Statistics* **17** 453–510.

BURNETT, R. T., MA, R., JERRETT, M., GOLDBERG, M., CAKMAK, S., POPE, C. A. and KREWSKI, D. (2001). The spatial association between community air pollution and mortality: a new method of analyzing correlated geographic cohort data. *Environmental health perspectives* **109** 375 – 380.

CASELLA, G. and GEORGE, E. I. (1992). Explaining the gibbs sampler. *The American Statistician* **46** 167 – 174.

CHEN, L. H., KNUTSEN, S. F., SHAVLIK, D., BEESON, W. L., PETERSON, F., GHAMSARY, M. and ABBEY, D. (2005). The association between fatal coronary heart disease and ambient particulate air pollution: Are females at greater risk? *Environmental Health Perspectives* **113** 1723–1729.

CHIANG, C. L. (1968). *Introduction to Stochastic Processes in Biostatistics*. Wiley, New York.

CHRISTENFELD, N. J., SLOAN, R. P., CARROLL, D. and GREENLAND, S. (2004). Risk factors, confounding, and the illusion of statistical control. *Psychosomatic Medicine* **66** 868–875.

CLIFF, A. and ORD, J. K. (1972). Testing for spatial autocorrelation among regression residuals. *Geographical Analysis* **4** 267 – 284.

CORRADI, C. and GAMBETTA, G. (1976). The estimation of distributed lags by spline functions. *Empirical Economics* **1** 41–51.

CORREIA, A. W., POPE, C. A., DOCKERY, D. W., WANG, Y., EZZATI, M. and DOMINICI, F. (2013). Effect of air pollution control on life expectancy in the United States: An analysis of 545 US counties for the period from 2000 to 2007. *Epidemiology* **24** 23–31.

COSSMAN, J. S., JAMES, W. L., COSBY, A. G. and COSSMAN, R. E. (2010). Underlying causes of the emerging nonmetropolitan mortality penalty. *American Journal of Public Health* **100** 1417–1419.

DE KLUIZENAAR, Y., GANSEVOORT, R. T., MIEDEMA, H. M. and DE JONG, P. E. (2007). Hypertension and road traffic noise exposure. *Occupational and Environmental Medicine* **49** 484 – 492.

DIGGLE, P., LIANG, K. and ZEGER, S. L. (1994). *Analysis of Longitudinal Data*. Oxford: Clarendon Press.

DOCKERY, D. W., POPE, C. A. and XU, X. (1993). An association between air pollution and mortality in six US cities. *New England Journal of Medicine* **329** 1753–1759.

DOMINICI, F., DANIELS, M., ZEGER, S. L. and SAMET, J. M. (2002). Air Pollution and Mortality: Estimating Regional and National Dose-Response Relationships. *J Am Stat Assoc* **97** 100–111.

DOMINICI, F., MCDERMOTT, A., ZEGER, S. L. and SAMET, J. M. (2003). Airborne particulate matter and mortality: Timescale effects in four U.S. cities. *American journal of Epidemiology* **157** 1055–1065.

DOMINICI, F., PENG, R. D., ZEGER, S. L. and SAMET, J. M. (2006). Recent developments of the national morbidity mortality air pollution study: 1987-2000. *Epidemiology* **17** S19 – S19.

EFTIM, S. E., SAMET, J. M., JANES, H., MCDERMOTT, A. and DOMINICI, F. (2008). Fine particulate matter and mortality: a comparison of the six cities and American Cancer Society cohorts with a medicare cohort. *Epidemiology* **19** 209 – 216.

ERIKSSON, C., BLUHM, G., HILDING, A., OSTENSON, C. G. and PERSHAGEN, G. (2010). Aircraft noise and incidence of hypertension - gender specific effects. *Environmental Research* **110** 764 – 772.

ERIKSSON, C., ROSENLUND, M., PERSHAGEN, G., HILDING, A., OSTENSON, C. G. and BLUHM, G. (2007). Aircraft noise and incidence of hypertension. *Epidemiology* **18** 716 – 721.

EVANS, J., TOSTESON, T. and KINNEY, P. (1984). Cross-sectional mortality studies and air pollution risk assessment. *Environment International* **10** 55–83.

EZZATI, M., FRIEDMAN, A., KULKARNI, S. and MURRAY, C. J. (2008). The reversal of fortunes: trends in county mortality and cross-county mortality disparities in the united states. *PLoS Medicine* **5** e66.

FEDERAL AVIATION ADMINISTRATION - OFFICE OF ENVIRONMENT AND ENERGY (2007). Integrated Noise Model (INM) Version 7.0 User's Guide.

FLOUD, S., VIGNA-TAGLIANTI, F., HANSELL, A., BLANGIARDO, M., HOUTHUIJS, D. and BREUGELMANS, O. (2011). Medication use in relation to noise from aircraft and road traffic in six European countries: results of the HYENA study. *Occupational and Environmental Medicine* **68** 518 – 524.

FRANSSEN, E. A., VAN WIECHEN, C. M., NAGELKERKE, N. J. and LEBRET, E. (2004). Aircraft noise around a large international airport and its impact on general health and medication use. *Occupational and Environmental Medicine* **61** 405 – 413.

GASPARRINI, A., ARMSTRONG, B. and KENWARD, M. G. (2010). Distributed lag non-linear models. *Statis Med* **29** 2224–2234.

GRASSI, G. (2006a). Sympathetic and reflex cardiovascular abnormalities in diabetes and metabolic syndrome. *Journal of Hypertension* **24** 22 – 22.

GRASSI, G. (2006b). Sympathetic overdrive and cardiovascular risk in the metabolic syndrome. *Hypertension Research* **29** 839 – 847.

GREENLAND, S. and MORGENSTERN, H. (2001). Confounding in health research. *Annual Review of Public Health* **22** 189–212.

GREISER, E., JANHSEN, K. and GREISER, C. (2007). Air traffic noise increases prescriptions of cardiovascular drugs in the vicinity of a major airport. *Epidemiology* **18** S33.

GREVEN, S., DOMINICI, F. and ZEGER, S. L. (2011). An approach to the estimation of chronic air pollution effects using spatio-temporal information. *Journal of the American Statistical Association* **106** 396–406.

HARALABIDIS, A. S., DIMAKOPOULOU, K., VELONAKI, V., BARBAGLIA, G., MUSSIN, M. and GIAMPAOLO, M. (2011). Can exposure to noise affect the 24 h blood pressure profile? Results from the HYENA study. *Journal of Epidemiology and Community Health* **65** 535 – 541.

HARALABIDIS, A. S., DIMAKOPOULOU, K., VIGNA-TAGLIANTI, F., GIAMPAOLO, M., BORGINI, A. and DUDLEY, M. L. (2008). Acute effects of night-time noise exposure on blood pressure in populations living near airports. *European Heart Journal* **29** 658 – 664.

HATFIELD, J., JOB, R., CARTER, N. L., PEPLOE, P., TAYLOR, R. and MORRELL, S. (2001). The influence of psychological factors on self-reported physiological effects of noise. *Noise & Health* **3** 1 – 13.

HAVRANEK, E. P., WOLFE, P., MASOUDI, F. A., RATHORE, S. S., KRUMHOLZ, H. M. and ORDIN, D. L. (2004). Provider and hospital characteristics associated with ge-

ographic variation in the evaluation and management of elderly patients with heart failure. *Archives of Internal Medicine* **164** 1186 – 1191.

HODGES, J. S. and REICH, B. J. (2010). Adding spatially-correlated errors can mess up the fixed effects you love. *The American Statistician* **64** 325 – 334.

HOFF, P. D. (2009). *A First Course in Bayesian Statistical Methods*. Springer.

HOLFORD, T. (1980). The analysis of rates and of survivorship using log-linear models. *Biometrics* **36** 299–305.

HUSS, A., SPOERRI, A., EGGER, M. and ROOSLI, M. (2010). Aircraft noise, air pollution, and mortality from myocardial infarction. *Epidemiology* **21** 829 – 836.

ISING, H. and KRUPPA, B. (2004). Health effects caused by noise: Evidence in the literature from the past 25 years. *Noise & Health* **6** 5 – 13.

JANES, H., DOMINICI, F. and ZEGER, S. L. (2007). Trends in air pollution and mortality. *Epidemiology* **18** 416–423.

JARUP, L., BABISCH, W., HOUTHUIJS, D., PERSHAGEN, G., KATSOUYANNI, K. and CADUM, E. (2008). Hypertension and exposure to noise near airports: the HYENA study. *Environmental Health Perspectives* **116** 329 – 333.

JERRETT, M., BURNETT, R. T., MA, R., POPE, C. A., KREWSKI, D., NEWBOLD, K. B., THURSTON, G. D., SHI, Y., FINKLESTEIN, N., CALLE, E. E. and THUN, M. J. (2005). Spatial analysis of air pollution and mortality in los angeles. *Epidemiology* **16** 727 – 736.

JOB, R. F. S. and HATFIELD, J. (2001). The impact of soundscape, enviroscape, and psychscape on reaction to noise: Implications for evaluation and regulation of noise effects. *Noise Control Engineering* **49** 120 – 124.

KASS, R. E. and WASSERMAN, L. (1995). A reference Bayesian test for nested hypotheses and its relationship to the Schwartz Criterion. *Journal of the American Statistical Association* **90** 928 – 934.

KEYFITZ, N. and FLIEGER, W. (1990). *World Population Growth and Aging: Demographic Trends in the Late Twentieth Century*. University of Chicago, Chicago.

KIYOTA, Y., SCHNEEWEISS, S., GLYNN, R. J., CANNUSCIO, C. C., AVORN, J. and SOLOMON, D. H. (2004). Accuracy of Medicare claims-based diagnosis of acute myocardial infarction: Estimating positive predictive value on the basis of review of hospital records. *American Heart Journal* **148** 99 – 104.

KNIPSCHILD, P. (1977). Medical effects of aircraft noise - community cardiovascular survey 5. *International Archives of Occupational and Environmental Health* **40** 185 – 190.

KREWSKI, D., BURNETT, R. T., GOLDBERG, M., HOOVER, K., SIEMEATYCKI, J., ABRAHAMOWICZ, M., VILLENEUVE, P. and WHITE, W. (2005a). Reanalysis of the Harvard Six Cities Study, part II: sensitivity analysis. *Inhalation Toxicology* **17** 343 – 353.

KREWSKI, D., BURNETT, R. T., GOLDBERG, M., HOOVER, K., SIEMIATYCKI, J., ABRAHAMOWICZ, M. and WHITE, W. (2000). Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of particulate air pollution and mortality: special report. Health Effects Institute.

KREWSKI, D., BURNETT, R. T., GOLDBERG, M., HOOVER, K., SIEMIATYCKI, J., ABRAHAMOWICZ, M. and WHITE, W. (2005b). Reanalysis of the Harvard Six Cities Study, part I: validation and replication. *Inhalation Toxicology* **17** 335 – 342.

KULKARNI, S., LEVIN-RECTOR, A., EZZATI, M. and MURRAY, C. J. (2011). Falling behind: life expectancy in US counties from 2000 to 2007 in an international context. *Population Health Metrics* **9** 16.

LADEN, F. D., SCHWARTZ, J., SPEIZER, F. E. and DOCKERY, D. W. (2006). Reduction in fine particulate air pollution and mortality: extended follow-up of the harvard six cities study. *American Journal of Respiratory and Critical Care Medicine* **173** 667 – 672.

LAIRD, N. and OLIVIER, D. (1981). Covariance analysis of censored survival data using log-linear analysis techniques. *J Am Stat Assoc* **76** 231–240.

LOSINA, E., BARRETT, J., BARON, J. A. and KATZ, J. N. (2003). Accuracy of Medicare claims data for rheumatologic diagnoses in total hip replacement recipients. *Journal of Clinical Epidemiology* **56** 515 – 519.

LOUIE, P. K., CHOW, J. C., CHEN, L. W., WATSON, J. G., LEUNG, G. and SIN, D. W. (2005). PM2.5 chemical composition in Hong Kong: urban and regional variations. *Science of the Total Environment* **338** 267–281.

MANCIA, G., DELL'ORO, R., QUARTI-TREVANO, F., SCOPELLITI, F. and GRASSI, G. (2006). Angiotensin-sympathetic system interactions in cardiovascular and metabolic disease. *Journal of Hypertension* **24** S51 – S56.

MCMICHAEL, A., ANDERSON, H., BRUNEKREEF, B. and COHEN, A. (1998). Inappropriate use of daily mortality analyses to estimate longer-term mortality effects of air pollution. *International Journal of Epidemiology* **27** 450 – 453.

MIEDEMA, H. M., VOS, H. and DE JONG, R. G. (2000). Community reaction to aircraft noise: time-of-day penalty and tradeoff between levels of overflights. *Journal of the Acoustical Society of America* **107** 3245 – 3253.

MOOLGAVKAR, S. (1994). Air pollution and mortality. *New England Journal of Mediciine* **330** 1237–1238.

MOOLGAVKAR, S. (2005). A review and critique of the EPAs rationale for a fine particle standard. *Regulatory Toxicology and Pharmacology* **42** 123–144.

MORAN, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika* **35** 255 – 260.

OSTRO, B., LIPSETT, M., REYNOLDS, P., GOLDBERG, D., HERTZ, A., GARCIA, C., HENDERSON, K. D. and BERNSTEIN, L. (2010). Long-term exposure to constituents of fine particulate air pollution and mortality: Results from the California Teachers Study. *Environmental Health Perspectives* **118** 363–369.

ÖZKAYNAK, H. and THURSTON, G. (1987). Associations between 1980 US mortality rates and alternative measures of airborne particle concentration. *Risk Analysis* **7** 449 – 461.

PENG, R. and DOMINICI, F. (2008). *Statistical Methods for Environmental Epidemiology in R: A Case Study in air Pollution and Health*. New York: Springer.

PETO, R., LOPEZ, A. D., BOREHAM, J., THUN, M. J. and HEATH, C. (1992). Mortality from tobacco in developed countries: indirect estimation from national vital statistics. *Lancet* **339** 1268 – 1278.

POPE, C. A. (2007). Mortality effects of longer term exposures to fine particulate air pollution: Review of recent epidemiological evidence. *Inhalation Toxicology* **19** 33 – 38.

POPE, C. A., BURNETT, R. T., THUN, M. J., CALLE, E. E., KREWSKI, D., ITO, K. and THURSTON, G. D. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Journal of the American Medical Association* **287** 1132 – 1141.

POPE, C. A., BURNETT, R. T., THURSTON, G. D., THUN, M. J., CALLE, E. E., KREWSKI, D. and GODLESKI, J. (2004). Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. *Circulation* **109** 71 – 77.

POPE, C. A. and DOCKERY, D. W. (2006). Health effects of fine-particulate air pollution: lines that connect. *Journal of the Air and Waste Management Association* **56** 709–742.

POPE, C. A., .EZZATI, M. and DOCKERY, D. W. (2009). Fine-particulate air pollution and life expectancy in the United States. *New England Journal of Medicine* **360** 376–386.

POPE, C. A., THUN, M. J. and NAMBOODIRI, N. M. (1995). Particulate air pollution as a predictor of mortality in a prospective study of US adults. *American Journal of Respiratory and Critical Care Medicine* **151** 669–674.

PRESTON, S., HEUVELINE, P. and GUILLOT, M. (2001). *Demography: Measuring and Modeling Population Processes*. Wiley-Blackwell, Malden (Massachusetts).

PUETT, R. C., HART, J. E., SUH, H., MITTLEMAN, M. and LADEN, F. (2011). Particulate matter exposures, mortality, and cardiovascular disease in the health professionals follow-up study. *Environmental Health Perspectives* **119** 1130–1135.

RABL, A. (2003). Interpretation of air pollution mortality: number of deaths or years of life lost? *Journal of the Air and Waste Management Association* **53** 41 – 50.

REICH, B. J., HODGES, J. S. and ZADNIK, V. (2006). Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models. *Biometrics* **62** 1197 – 1206.

ROSENLUND, M., BERGLIND, N., PERSHAGEN, G., JARUP, L. and BLUHM, G. (2001). Increased prevalence of hypertension in a population exposed to aircraft noise. *Occupational and Environmental Medicine* **58** 769 – 773.

SAMET, J. M., DOMINICI, F., CURRIERO, F. C., COURSAC, I. and ZEGER, S. L. (2000). Fine particulate air pollution and mortality in 20 US cities, 1987 - 1994. *New England Journal of Medicine* **343** 1742 – 1749.

SCHMIDT, M., HASSETT-SIPPLE, B. and RAJAN, P. (2010). PM2.5 air quality analyses - memorandum to PM NAAQS review. United States Environmental Protection Agency.

SCHWARTZ, J. (1991). Particulate air pollution and daily mortality in Detroit. *Environmental Research* **56** 204 – 213.

SCHWARTZ, J. (1992). Particulate air pollution and daily mortality in Steubenville, Ohio. *American Journal of Epidemiology* **135** 12 – 19.

SCHWARTZ, J. (2000). The distributed lag between air pollution and daily deaths. *Epidemiology* **11** 320–326.

SCHWARTZ, J. (2001). Is there harvesting in the association of airborne particles with daily deaths and hospital admissions? *Epidemiology* **12** 55–61.

SCHWARTZ, J., COULL, B. A., LADEN, F. D. and RYAN, L. M. (2008). The effect of dose and timing of dose on the association between airborne particles and survival. *Environmental Health Perspectives* **116** 64 – 69.

SELANDER, J., BLUHM, G., THEORELL, T., PERSHAGEN, G., BABISCH, W. and SEIFFERT, I. (2009). Saliva cortisol and exposure to aircraft noise in six european countries. *Environmental Health Perspectives* **117** 1713 – 1717.

SHRESTHA, L. B. (2005). Life expectancy in the united states. Tech. rep., CRS Report for Congress.

SPRENG, M. (2000). Possible health effects of noise induced cortisol increase. *Noise & Health* **2** 59 – 63.

STANSFELD, S. and CROMBIE, R. (2011). Cardiovascular effects of environmental noise: Research in the United Kingdom. *Noise & Health* **13** 229 – 233.

TAINIO, M., TUOMISTO, J., HÄNNINEN, O., RUUSKANEN, J., JANTUNEN, M. and PEKKANEN, J. (2007). Parameter and model uncertainty in a life-table model for fine particles (PM2.5): a statistical modeling study. *Environmental Health* **6** 24.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY (1997). The benefits and costs of the Clean Air Act, 1970 - 1990.

VEDAL, S. (1997). Ambient particles and health: Lines that divide. *Journal of the Air and Waste Management Association* **47** 551–581.

WAKEFIELD, J. (2003). Sensitivity analyses for ecological regression. *Biometrics* **59** 9 – 17.

WAKEFIELD, J. (2007). Disease mapping and spatial regression with count data. *Biostatistics* **8** 158 – 183.

WALLER, L. and CARLIN, B. (2010). Disease Mapping. In *Handbook of Spatial Statistics* (A. E. Gelfand, P. J. Diggle, M. Fuentes and P. Guttorp, eds.). Chapman & Hall, 217 – 244.

YEH, R. W., NORMAND, S. L., BARR, C. D. and DOMINICI, F. (2011). Geographic disparities in the incidence and outcomes of hospitalized myocardial infarction: Does a rising tide lift all boats? *Circulation: Cardiovascular Quality and Outcomes* **6** 9–16.

ZANOBETTI, A., WAND, M. P., SCHWARTZ, J. and RYAN, L. M. (2000). Generalized additive distributed lag models: quantifying mortality displacement. *Biostatistics* **1** 279–292.

ZEGER, S. L., DOMINICI, F., MCDERMOTT, A. and SAMET, J. M. (2008). Mortality in the Medicare population and chronic exposure to fine particulate air pollution in urban centers (2000 - 2005). *Environmental Health Perspectives* **116** 1614–1618.