HEI
Health Effects Institute

NUMBER 234
DECEMBER 2025

Walter A. Rosenblith New Investigator Award

RESEARCH REPORT

Robust Statistical Approaches to Understanding the Causal Effect of Air Pollution Mixtures

Joseph Antonelli, Heejun Shin, Suyeon Kang, Alexander Franks, Michelle Audirac, and Danielle Braun

INCLUDES A COMMENTARY BY THE INSTITUTE’S REVIEW COMMITTEE

www.healtheffects.org

Robust Statistical Approaches to Understanding the Causal Effect of Air Pollution Mixtures

Joseph Antonelli, Heejun Shin, Suyeon Kang, Alexander Franks,
Michelle Audirac, and Danielle Braun

with a Commentary by the HEI Review Committee

Research Report 234
Health Effects Institute
Boston, Massachusetts

Trusted Science for Cleaner Air and Better Health

Publishing history: This report was posted at *www.healtheffects.org* in December 2025.

Citation for report:

Antonelli J, Shin H, Kang S, Franks A, Audirac M, Braun D. 2025. Robust Statistical Approaches to Understanding the Causal Effect of Air Pollution Mixtures. Research Report 234. Boston, MA: Health Effects Institute.

© 2025 Health Effects Institute, Boston, MA, USA. Compositor: Progressive Publishing Services, York, PA.
Library of Congress Catalog Number for the HEI Report Series: WA 754 R432.

Contents of this report may not be used without prior permission from the Health Effects Institute. Please send requests to *pubs@healtheffects.org*.

Health Effects Institute and HEI are service marks registered in the US Patent and Trademark Office.

ISSN: 2688-6855 (online)

CONTENTS

About HEI ix

About This Report xi

Contributors xiii

HEI STATEMENT 1

INVESTIGATORS' REPORT *by Antonelli et al.* 3

ABSTRACT 3

CHAPTER 1: INTRODUCTION 4
  1.1 On the Use of a Causal Inference Framework for Multiple Exposures 5

CHAPTER 2: SPECIFIC AIMS AND OVERARCHING APPROACH 6
  2.1 Running Application to the Medicare Population 6
    2.1.1 Descriptive Information on Medicare Cohort and Exposure Data 7

CHAPTER 3: CAUSAL ANALYSIS OF AIR POLLUTION MIXTURES: ESTIMANDS, POSITIVITY, AND EXTRAPOLATION 11
  3.1 Introduction 11
  3.2 Causal Estimands for Environmental Mixtures 11
    3.2.1 Which Mixture Effect? Anchoring Causal Estimands to Interventions 11
    3.2.2 Key Assumptions and the Importance of Mixture Positivity 12
  3.3 Illustrating Positivity Violations and Model Extrapolation for Environmental Mixtures 12
  3.4 Potential Paths Forward: Diagnosing Mixture Positivity and Alternative Estimands 13
    3.4.1 Alternative Estimand One: Effects of Feasible Exposure Changes 14
    3.4.2 Alternative Estimand Two: Effects in Feasible Subpopulations 14
  3.5 Illustration of United States Air Pollution Mixtures 15
  3.6 Discussion and Conclusions 17

CHAPTER 4: HETEROGENEITY OF THE HEALTH EFFECTS OF AIR POLLUTION MIXTURES 18
  4.1 Introduction 18
  4.2 Summarizing Effect Modification for Air Pollution Mixtures 18
  4.3 Statistical Modeling Framework 19
    4.3.1 Variations of Bayesian Additive Regression Trees 20
  4.4 Data Analysis 20
    4.4.1 Results 21
  4.5 Discussion and Conclusions 23

CHAPTER 5: ACCOUNTING FOR MOBILITY WHEN ESTIMATING THE HEALTH EFFECTS OF AIR POLLUTION MIXTURES 25
  5.1 Introduction 25
  5.2 Formalizing the Effect of Mobility 25

5.2.1 Interference SUTVA Assumption 26
5.2.2 Choice of g(·) and Cell Phone Mobility Data 26
5.2.3 Estimands That Incorporate Mobility 27
5.3 What Happens When Mobility Is Ignored 27
5.4 Robustness to Errors in Mobility Estimates 28
5.5 Data Analysis 29
5.5.1 Statistical Modeling 29
5.5.2 Assessing Mobility and Its Relationship to Air Pollution 30
5.5.3 Analysis of $PM_{2.5}$ Only 30
5.5.4 Analysis of Air Pollution Mixtures 32
5.6 Discussion and Conclusions 33

CHAPTER 6: ADDRESSING UNMEASURED CONFOUNDING WITH MULTIPLE EXPOSURES AND MULTIPLE OUTCOMES 34
6.1 Introduction 34
6.2 Estimands and Corresponding Assumptions 34
6.3 Partial Identification Regions with Unmeasured Confounding Bias 35
6.3.1 Benefits of Multiple Treatments and Multiple Outcomes 35
6.3.2 Incorporating Negative Controls to Sharpen Partial Identification Regions 36
6.3.3 Joint Partial Identification Regions 37
6.4 Robustness of the Causal Effects of $PM_{2.5}$ 38
6.4.1 Partial Identification Regions Incorporating Negative Controls 38
6.4.2 Incorporating Additional Prior Information 39
6.5 Discussion and Conclusions 41

CHAPTER 7: SYNTHESIS, INTERPRETATION, AND IMPLICATIONS OF FINDINGS 42

DATA AVAILABILITY STATEMENT 43

ACKNOWLEDGMENTS 44

REFERENCES 44

HEI QUALITY ASSURANCE STATEMENT 48

ABOUT THE AUTHORS 49

OTHER PUBLICATIONS RESULTING FROM THIS RESEARCH 49

COMMENTARY 51

INTRODUCTION 51

SCIENTIFIC AND REGULATORY BACKGROUND 51

STUDY OBJECTIVES 52

SUMMARY OF ANALYSES 52
Aim 1: Causal Analysis of Air Pollution Mixtures: Estimands, Positivity, and Model Extrapolation 52
Aim 2: Addressing the Heterogeneity of the Health Effects of Air Pollution Mixtures 53

Aim 3: Accounting for Mobility When Estimating the Health Effects of Air Pollution Mixtures 54
Aim 4: Addressing Residual Confounding with Multiple Exposures and Multiple Outcomes 55

HEI REVIEW COMMITTEE'S EVALUATION 56
Conclusions 57

ACKNOWLEDGMENTS 57

REFERENCES 57

Abbreviations and Other Terms 59

Related HEI Publications 60

HEI Board, Committees, and Staff 61

ABOUT HEI

The Health Effects Institute is a nonprofit corporation chartered in 1980 as an independent research organization to provide high-quality, impartial, and relevant science on the effects of air pollution on health. To accomplish its mission, the Institute

- identifies the highest-priority areas for health effects research
- competitively funds and oversees research projects
- provides intensive independent review of HEI-supported studies and related research
- integrates HEI's research results with those of other institutions into broader evaluations
- communicates the results of HEI's research and analyses to public and private decision-makers.

HEI typically receives balanced funding from the US Environmental Protection Agency and the worldwide motor vehicle industry. Frequently, other public and private organizations in the United States and around the world also support major projects or research programs. HEI has funded more than 390 research projects in North America, Europe, Asia, and Latin America, the results of which have informed decisions regarding carbon monoxide, air toxics, nitrogen oxides, diesel exhaust, ozone, particulate matter, and other pollutants. These results have appeared in more than 275 comprehensive reports published by HEI, as well as in more than 2,500 articles in peer-reviewed literature.

HEI's independent Board of Directors consists of leaders in science and policy who are committed to fostering the public–private partnership that is central to the organization. The Research Committee solicits input from HEI sponsors and other stakeholders and works with scientific staff to develop a Five-Year Strategic Plan, select research projects for funding, and oversee their conduct. The Review Committee or Panel, which has no role in selecting or overseeing studies, works with staff to evaluate and interpret the results of funded studies and related research.

All project results and accompanying comments by the Review Committee or Panel are widely disseminated through HEI's website (*www.healtheffects.org*), reports, newsletters, annual conferences, and presentations to legislative bodies and public agencies.

ABOUT THIS REPORT

Research Report 234, *Robust Statistical Approaches to Understanding the Causal Effect of Air Pollution Mixtures,* presents a research project funded by the Health Effects Institute and conducted by Dr. Joseph Antonelli at the University of Florida and colleagues. The report contains three main sections:

The HEI Statement, prepared by staff at HEI, is a brief, nontechnical summary of the study and its findings; it also briefly describes the Review Panel's comments on the study.

The Investigators' Report, prepared by Antonelli and colleagues, describes the scientific background, aims, methods, results, and conclusions of the study.

The Commentary, prepared by members of the Review Committee with the assistance of HEI staff, places the study in a broader scientific context, points out its strengths and limitations, and discusses the remaining uncertainties and implications of the study's findings for public health and future research.

This report has gone through HEI's rigorous review process. When an HEI-funded study is completed, the investigators submit a draft final report presenting the background and results of the study. Outside technical reviewers first examine this draft report. The report and the reviewers' comments are then evaluated by members of the Review Panel, an independent panel of distinguished scientists who are not involved in selecting or overseeing HEI studies. During the review process, the investigators have an opportunity to exchange comments with the Review Panel and, as necessary, to revise their report. The Commentary reflects the information provided in the final version of the report.

Although this report was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR–83998101 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and may not necessarily reflect the views of the Agency; thus, no official endorsement by it should be inferred. The report has also not been reviewed by private party institutions, including those that support the Health Effects Institute, and may not reflect the views or policies of these parties; thus, no endorsement by them should be inferred.

CONTRIBUTORS

RESEARCH COMMITTEE

David A. Savitz, Chair *Professor of Epidemiology, School of Public Health, and Professor of Obstetrics and Gynecology and Pediatrics, Alpert Medical School, Brown University, USA*

Benjamin Barratt *Professor, Environmental Research Group, School of Public Health, Imperial College London, United Kingdom*

David C. Dorman *Professor, Department of Molecular Biomedical Sciences, College of Veterinary Medicine, North Carolina State University, USA*

Arlene Fiore* *Peter H. Stone and Paola Malanotte Stone Professor of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, USA*

Christina H. Fuller *Associate Professor, School of Environmental, Civil, Agricultural and Mechanical Engineering, University of Georgia College of Engineering, USA*

Marianne Hatzopoulou *Professor, Civil and Mineral Engineering, University of Toronto, Research Chair in Transport Decarbonization and Air Quality, Canada*

Amy H. Herring† *Sara & Charles Ayres Distinguished Professor of Statistical Science, Global Health, and Biostatistics and Bioinformatics, Duke University, USA*

Heather A. Holmes *Associate Professor, Department of Chemical Engineering, University of Utah, USA*

Kyle Messier† *Stadtman Investigator, National Institute of Environmental Health Sciences, USA*

Neil Pearce *Professor of Epidemiology and Biostatistics, London School of Hygiene and Tropical Medicine, United Kingdom*

Evangelia (Evi) Samoli *Professor of Epidemiology and Medical Statistics, Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Greece*

Alexandra M. Schmidt *Professor of Biostatistics, School of Population and Global Health, McGill University, Canada*

Neeta Thakur *Associate Professor of Medicine, University of California San Francisco, USA*

Gregory Wellenius *Beverly A. Brown Professor for the Improvement of Urban Health and Director, Center for Climate and Health, Boston University School of Public Health, USA*

REVIEW COMMITTEE

Sara D. Adar *Professor of Epidemiology and Global Public Health, Department of Epidemiology, University of Michigan School of Public Health, USA*

Katherine B. Ensor *Noah G. Harding Professor of Statistics, Rice University, USA*

Ulrike Gehring *Associate Professor, Institute for Risk Assessment Sciences, Utrecht University, the Netherlands*

Michael Jerrett *Professor, Department of Environmental Health Sciences, Fielding School of Public Health, University of California, Los Angeles, USA*

Eric J. Tchetgen Tchetgen‡ *University Professor and Professor of Biostatistics and Epidemiology, Perelman School of Medicine, and Professor of Statistics and Data Science, The Wharton School, University of Pennsylvania, USA*

John Volckens *Professor, Department of Mechanical Engineering, Walter Scott Jr. College of Engineering, Colorado State University, USA*

Scott Weichenthal *Professor, Department of Epidemiology, Biostatistics, and Occupational Health, School of Population and Global Health, McGill University, Canada*

HEI PROJECT STAFF

Eva Tanner *Staff Scientist (Study Oversight)*

Dan Crouse *Senior Scientist (Report Review)*

Kristin C. Eckles *Senior Editorial Manager*

Tara Hamilton *Consulting Editor*

---

*Consultant.

†Dr. Herring rotated off the Research Committee before this report's publication.

‡Dr. Tchetgen Tchetgen rotated off the Review Committee before this report's publication.

HEI STATEMENT

Synopsis of Research Report 234

New Statistical Approaches for Addressing Challenging Questions About Air Pollution and Health

BACKGROUND

Recent decades have seen an increase in the development of statistical methods to answer questions about whether exposure to outdoor air pollution is causally linked to the risk of death. The need for investigators to satisfy specific data and analytic conditions to estimate causal effects from observational data is a key challenge to conducting such analyses. Dr. Joseph Antonelli of the University of Florida and colleagues sought to develop new statistical approaches for addressing challenging questions about potential causal links between air pollution mixtures and risk of death. Their study was funded through HEI's Request for Applications *19-2*: Walter A. Rosenblith New Investigator Award.

APPROACH

Antonelli and colleagues proposed developing approaches to address four specific challenges associated with conducting causal inference analyses related to air pollution mixtures and health:

1. Develop an approach to identify the effect of long-term exposure to a given air pollutant on health when multiple pollutants are present and when only limited data are available about their distributions.
2. Develop a method to estimate the effects of exposure to a mixture of pollutants on the risk of mortality and to identify subgroups of the population that are most susceptible to the effects of exposure to air pollution.
3. Evaluate the effect of accounting for daily mobility patterns on estimates of the health effects associated with long-term exposures to air pollution.
4. Develop a method to account for bias in cases for which potential confounding variables are not measured or accounted for perfectly in the statistical model.

These proposed aims were expected to advance the field by highlighting the challenges inherent in studying the health effects of multiple pollutant exposures and expanding the variety of research questions that can be answered regarding air pollution mixtures.

What This Study Adds

- This study is among the first to address a series of common challenges faced by researchers assessing the health effects of exposures to air pollution mixtures, from exposure assessment to causal inference.
- Antonelli and colleagues describe several new approaches to support causal inference, focusing on exposure to multiple pollutants, addressing bias resulting from confounding, and enhancing exposure assessment by incorporating people's mobility patterns.
- The approaches are applied to hypothetical scenarios and to data from a US Medicare cohort to demonstrate the effectiveness of the practice methods.
- Overall, the approaches presented here provide new solutions for overcoming challenges to assessing causal links between air pollution and health.

In some cases, the approaches and statistical models were described for theoretical applications, and in others, Antonelli and colleagues applied their methods to real-world data using a nationwide study of air pollution and health in the US Medicare population. Those analyses are based on the health information from over 30 million Medicare beneficiaries during the years 2000–2016 living in about 30,000 zip codes across the United States. For air pollution data, they used annual estimated concentrations of fine particles and several other pollutants from existing spatiotemporal models, all at a spatial resolution of about 1 km × 1 km, aggregated to zip codes.

KEY RESULTS

One key assumption for the validity of causal inference in environmental epidemiology is that every exposure value potentially of interest is, in fact, possible for all individuals in the dataset. That assumption would be violated if, for example, no

---

This Statement, prepared by the Health Effects Institute, summarizes a research project funded by HEI and conducted by Dr. Joseph Antonelli at the University of Florida and colleagues. Research Report 234 contains the detailed Investigators' Report and a Commentary on the study prepared by the HEI Review Committee.

men who are 40–45 years of age in the study dataset were exposed to the air pollution concentration of interest, as it would not be possible to learn about the risk of the health outcome associated with that exposure level among such men. As such, Antonelli and colleagues presented strategies to redefine causal models in a way that respects the limitations in the scope of the available exposure data.

The investigators also introduced a flexible statistical approach that allows one to better understand which population subgroups are most affected by exposure to air pollution or who might benefit most from an intervention to reduce exposures.

Antonelli and colleagues explored how estimates of health effects could be biased when mobility is ignored in estimating exposures to air pollution (e.g., by considering exposures only at study participants' residential addresses). Here, they presented a method that used location data derived from cell phones. They compared epidemiological models in which exposure was estimated for study participants, a) by using only the zip code in which each participant lived, and b) by modeling how much time each participant spent in their home and in other zip codes. Here, they found that although most participants spent about 22% of their time outside their home zip code, incorporating mobility information at that scale did not lead to appreciably different estimates of exposures, nor to appreciably different estimates of health effects. For example, the mean exposure to fine particles averaged across all study years using only the home zip code was 9.34 $\mu g/m^3$, and that which incorporated mobility was 9.64 $\mu g/m^3$.

INTERPRETATION AND CONCLUSIONS

In its independent evaluation of the Investigators' Report, the HEI Review Committee commended the investigators for developing original study aims and for tackling important issues for environmental epidemiology with which many statisticians continue to struggle. The Committee members agreed that the investigators did an excellent job setting the context and explaining the rationale for pursuing causal inference approaches to the study of air pollution exposures.

The innovative work presented by the investigators includes a rigorous assessment of the assumptions required to draw causal conclusions when studying complex mixtures of pollutants. They introduced several interesting theoretical approaches that advance the field. Their finding that incorporating mobility did not lead to appreciably different estimates of exposures corroborates recent findings from other studies.

Not surprisingly, the study team did not resolve all the important challenges described. For example, an overarching limitation of this study is that the various issues addressed by the investigators do not exist in isolation, yet the methods proposed here were presented independently. Their study is a valid starting point; however, addressing each challenge individually does not fully resolve the broader challenge of developing a comprehensive approach incorporating all of the important challenges in a single framework.

Another limitation of the study relates to the applied analyses with the Medicare data. The theoretical foundations presented here were evaluated in a very specific dataset with unique features (i.e., older adults and data available only at the zip code level) that might prevent valid inferences for reasons beyond the issues addressed by the investigators.

Although there is more work to be done to resolve these complex issues fully, this study provides a strong foundation for future work to extend the concepts described here to other settings.

INVESTIGATORS' REPORT

Robust Statistical Approaches to Understanding the Causal Effect of Air Pollution Mixtures

Joseph Antonelli[1], Heejun Shin[2], Suyeon Kang[3], Alexander Franks[4], Michelle Audirac[2], and Danielle Braun[2]

[1]University of Florida, Gainesville, Florida, USA; [2]Harvard University T.H. Chan School of Public Health, Boston, Massachusetts, USA; [3]University of Central Florida, Orlando, Florida, USA; [4]University of California, Santa Barbara, California, USA

ABSTRACT

Introduction Most existing epidemiological evidence on the health effects of air pollution has focused on single-pollutant analyses, although recent research has increasingly emphasized estimating the effects of multiple exposures simultaneously. In this report, we used causal inference methodology to highlight four impediments to analyses with multiple exposures: (1) there is little information in the data to estimate effects typically of interest, (2) the effects of air pollution mixtures can be heterogeneous, (3) exposure assessment using an individual's home location can be problematic when daily mobility takes them to areas of different exposure levels, and (4) bias due to unmeasured confounding. The objectives of this report were to address these four concerns through the development of rigorous statistical methodology and to provide a corresponding case study that examines the health effects of air pollution in the Medicare cohort in the United States.

Methods The statistical methodology developed in this report improves the analysis of environmental mixtures in two distinct ways. First, our results highlight inherent difficulties, which require careful consideration in any study of the health effects of multiple exposures. Second, we developed a statistical methodology that broadens the scope of questions that can be answered in analyses of air pollution mixtures and can increase the policy relevance of evidence obtained from epidemiological studies using multiple exposures. Additionally, we illustrated the aforementioned approaches in a nationwide study of the health effects of air pollution in the US Medicare population, extending the existing evidence on the health effects of air pollution within this cohort.

Results In specific aim 1, we found that quantities typically targeted in studies with multiple exposures are difficult to estimate from the observed data alone, as they frequently rely on model-based extrapolation, which can provide unreliable findings. We presented alternative strategies that provide policy-relevant evidence of health effects, while avoiding issues caused by extrapolation. In specific aim 2, we found that the adverse effects of particulate matter ≤2.5 µm in aerodynamic diameter ($PM_{2.5}$)* components are heterogeneous and that these effects are more pronounced in areas with lower socioeconomic status. Specific aim 3 studied the mobility of individuals and found that ignoring mobility can bias health effects, although typically toward the null of no exposure effect. Incorporating mobility in the Medicare cohort did not lead to substantially different findings; however, accounting for mobility tended to increase the magnitude of estimated health effects. In specific aim 4, we developed a methodology for assessing robustness of health effects to unmeasured confounding bias and found that there is robust evidence overall of a harmful effect of pollution on public health.

Conclusions Our studies provide strong evidence of air pollution effects on public health, and our methodology gives new insights into key issues about this effect. Specifically, the effects of air pollution are heterogeneous and affect certain subgroups of the population more than others, and these effects are moderately robust to unmeasured confounding bias. Future studies can incorporate the ideas and approaches developed in this report to address important questions in analyses with multiple exposures.

This Investigators' Report is one part of Health Effects Institute Research Report 234, which also includes a Commentary by the Review Committee and an HEI Statement about the research project. Correspondence concerning the Investigators' Report may be addressed to Dr. Joseph Antonelli, University of Florida, Department of Statistics, 206 Griffin-Floyd Hall, Gainesville, FL, 32601; email: *jantonelli@ufl.edu*. No potential conflict of interest was reported by the authors.

Although this report was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR–83998101 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and may not necessarily reflect the views of the Agency; thus, no official endorsement by it should be inferred. This report has also not been reviewed by private party institutions, including those that support the Health Effects Institute, and may not reflect the views or policies of these parties; thus, no endorsement by them should be inferred.

* A list of abbreviations and other terms appears at the end of this volume.

CHAPTER 1: INTRODUCTION

There is a well-established link between air pollution and adverse health outcomes that has been partially driven by large-scale observational studies that examined the health effects of pollutants such as particulate matter ≤2.5 µm ($PM_{2.5}$), particulate matter ≤10 µm (PM_{10}), or sulfates.[1–5] These studies, among many others, have been instrumental in understanding the adverse impacts of fine particulate matter on human health, although the majority of analyses in these studies focus on either a single pollutant at a time or a very small number of pollutants simultaneously. While there are undoubtedly still important questions being answered by single-pollutant analyses, the recent push in environmental epidemiology toward multiple pollutant approaches reflects the recognition that analyzing multiple pollutants simultaneously provides a more realistic assessment of human exposure to air pollution.[6,7] Many recent studies, such as those stemming from the recent HEI initiative to study health effects of low exposure levels, have included multiple pollutants in analyses studying health effects of ambient air pollution.[8–12] These studies have been immensely valuable for understanding additional aspects of the health effects of air pollution, particularly at levels currently observed in North America and Europe. However, because multiple pollutants are typically included as additional terms in a regression model, these analyses do not address the additional complexities of causal inference for multivariate exposures.

Key Terminology

Estimand: A scientific quantity of interest intended to be estimated using observed data.

Extrapolation: When a statistical model is used to predict outcomes for exposure or covariate values that are not within the range seen in the observed data.

Exposure effect heterogeneity: When the effects of environmental exposures vary across subgroups of the population.

Potential outcome: The outcome hypothetically observed if the exposure were fixed to a specific value.

Negative control exposure: An exposure known not to causally affect the outcome of interest.

Negative control outcome: An outcome known not to be causally affected by the exposure of interest.

Unmeasured confounding bias: When an omitted confounder biases the estimated effect of an exposure on an outcome.

Positivity assumption: An assumption necessary in many causal analyses that requires all subgroups of the population to have a positive probability of receiving each exposure level being examined.

Interference: When the outcomes of one observation are affected by the exposures of other observations.

Because of these additional complications, there has been rapid growth in statistical approaches tailored to the unique problem of estimating the health effects of mixtures. While some researchers have simplified the problem by categorizing continuous, multivariate exposures[13] or using weighted combinations of the exposures in the mixture,[14,15] the majority of methods are focused on complex regression models that relate the outcome to a large number of exposures in a manner that identifies important exposures and allows for nonlinear and interactive relationships between the exposures and the outcome.[16–23] Despite this surge of interest in statistical methodology for environmental mixtures, crucial gaps in the literature remain. Even if these statistical approaches are sufficient for modeling the complex relationship between the air pollution mixture and health, it is not clear how results from these analyses can be used to guide future air pollution regulations. Most approaches produce a set of important exposures that adversely affect health, but regulating single exposures is difficult, if not impossible. Air pollution regulations are implemented by reducing emissions from particular sources, and it is unlikely that reducing any one source of pollution will reduce only one environmental exposure. A critical question is how to use results from studies of environmental mixtures to target future air pollution policies.

Before proceeding further, it is important to clarify how our study relates to existing work on environmental mixtures and what we mean when we state we are studying mixtures. Typically, mixture analyses either attempt to identify the most harmful components of the mixture or investigate whether the joint effect of the mixture is different from the sum of the individual effects of each pollutant in the mixture. We did not have either of these goals in mind when studying air pollution mixtures. In this report, we documented our interest in studying the joint health effects of air pollution mixtures, which involved estimating how we would expect health outcomes to change if we were to change the level of some (or all) of the exposures. Such a change in air pollution exposure levels could be caused by a policy shift or intervention, such as the introduction of emission-reducing technology at coal-fired power plants. In principle, the ideas delineated in this report could be applied to questions such as identifying the most harmful components of a mixture or assessing whether the components have synergistic or antagonistic effects on health outcomes; however, pursuing these questions was not our primary goal. Instead, our focus was on answering a broader set of scientific questions related to air pollution mixtures and clarifying the significant and pervasive challenges they pose. Therefore, when we say we studied mixtures, we simply mean that we were interested in estimating the health effects of a moderate number of pollutants simultaneously.

Much of the previous work on environmental mixtures does not address important questions about causal assumptions necessary for these results to imply causal associations. Recent work has shown that environmental mixtures, particularly in the common setting of highly correlated exposures, are particularly susceptible to issues from unmeasured confounding bias.[24] Other crucial assumptions have received little to no attention in the causal inference literature in the context of multivariate exposures, let alone in the literature on environmental mixtures. This report aimed to fill many of these gaps to better understand causal effects for air pollution mixtures. Throughout the report, we examined each of the core causal assumptions necessary to bridge statistical modeling and causal estimation, and we highlighted these issues in the context of multiple exposures, specifically, as certain issues and solutions to these problems are unique to this setting. In addition to studying crucial assumptions necessary for effects to be interpreted causally, we also improved on the aforementioned statistical methodology by developing nonparametric Bayesian models that allow mixture effects to depend on characteristics of the population being exposed to increased air pollution levels.

1.1 ON THE USE OF A CAUSAL INFERENCE FRAMEWORK FOR MULTIPLE EXPOSURES

Before we address specific objectives, it is important to first discuss the role of causal inference approaches that appear throughout this report. We are not the first to initiate such a discussion, as recent HEI reports have directly addressed this issue[12,25] along with a multitude of manuscripts that have advocated for causal inference approaches in environmental epidemiology.[26–29] Despite the multitude of recent pertinent discussions, we think it is important to include related ideas here for two reasons: (1) to provide perspective for this work for readers unfamiliar with these ideas, and (2) to emphasize these ideas in the specific context of our study, which differ somewhat from prior studies.

The first key point is that causal inference approaches are not an all-encompassing panacea for observational studies. They do not necessarily provide greater causal evidence of an effect of air pollution on public health than other epidemiological studies that do not use such approaches. The most important aspects to providing causal evidence from epidemiological studies are other core factors, such as using sound study design, addressing selection or confounding bias, and reducing measurement error, among a host of other possible issues. Additionally, many other approaches for establishing causality do not incorporate epidemiology, such as employing toxicology or animal studies. The question then becomes why should we develop causal inference approaches at all, and what value do they provide in the context of air pollution epidemiology? One benefit is the clear definition of the estimand of interest, which is the scientific quantity we are interested in estimating in our study. For example, an estimand could be the expected change in mortality if we were to decrease all exposure levels by one unit. Another commonly stated motivation for pursuing causal inference approaches is that they formalize the assumptions necessary to infer causality, encourage us to reason about or assess sensitivity to these assumptions, separate the estimand of interest from any one statistical model, and alleviate issues from model misspecification.

We agree with the aforementioned reasons to advocate for causal inference approaches and formality, but we also feel that it may not be abundantly obvious to researchers unfamiliar with these approaches what is really meant by each of these motivations and how they benefit analyses of the health effects of air pollution. We hope that these reasons will become increasingly apparent in this report, particularly in the difficult context of trying to estimate causal effects for air pollution mixtures. Each of the chapters of this report tackled a different critical issue for the causal analysis of air pollution mixtures. It is our view that the important issues addressed within each chapter either would have been difficult to identify as problems at all or would have been much more difficult to solve if not for the formalization of the problem within a causal framework. We emphasize again that this stance does not mean the strength of causal evidence from our analyses is greater, but rather that causal inference approaches helped illuminate many issues and their corresponding solutions, which can benefit environmental epidemiology moving forward. Additionally, it is worth noting that while we identified these issues and solved them within a causal framework, the implications of our findings go beyond the causal analysis of air pollution mixtures. We believe these findings can help more broadly in the analysis of air pollution mixtures, even when causal approaches are not utilized.

CHAPTER 2: SPECIFIC AIMS AND OVERARCHING APPROACH

In this chapter, we provide an overview of the distinct aims that are covered in this report. As mentioned above, each aim corresponds to a different crucial assumption necessary for inferring causal effects of air pollution mixtures or tackles an unanswered methodological issue for multiple pollutant analyses. Each ensuing chapter of the report corresponds to a different aim, and these are summarized below. Additionally, the location of the results from each chapter can be seen in **Table 1**.

Causal inference with multivariate air pollution mixtures is a difficult problem with multiple unique challenges. One key issue is that the observed data may be very limited in their ability to adequately estimate causal effects that are of scientific interest due to violations of an assumption known as positivity. We formalized this assumption for air pollution mixtures and highlighted the challenges it presents, provided researchers with new diagnostic metrics to identify issues from violations of positivity, and defined novel estimands that are still policy-relevant for air pollution mixtures but are more robust to positivity violations.

Table 1. Summary of the Main Aims of the Project and the Associated Locations Within the Report

Research Roadmap	
Aims and Research Conducted	**Methods Description**
Aim 1	
• Addressing model-extrapolation and mixture positivity	Chapter 3.1–3.4
• Evaluating implications for US air pollution mixtures	Chapter 3.5
Aim 2	
• Developing methodology to estimate heterogeneous mixture effects	Chapter 4.1–4.3
• Estimating heterogeneous effects in the Medicare cohort	Chapter 4.4
Aim 3	
• Formalizing mobility and associated biases	Chapter 5.1–5.4
• Assessing the impact of mobility on the health effects of air pollution	Chapter 5.5
Aim 4	
• Developing methodology to account for unmeasured confounding bias	Chapter 6.1–6.3
• Assessing unmeasured confounding in the Medicare cohort	Chapter 6.4

We developed a statistical methodology to address two key issues for the estimation of the effects of multivariate air pollution mixtures. We developed a flexible Bayesian methodology that reduces the detrimental impacts of model misspecification and allows the mixture effect to be heterogeneous and vary across subgroups of the population. We also proposed new estimands that describe the nature and magnitude of the heterogeneity of mixture effects, while also identifying characteristics that contribute most to heterogeneity.

In air pollution studies, exposure is commonly assigned according to an individual's home location, although many individuals travel to other geographic areas with different levels of air pollution. We formalized this issue as one of interference in causal inference studies, in which outcomes of one area can be affected by exposures from other areas. We derived bias formulas that show the direction and magnitude of bias that stems from ignoring mobility and developed a framework that can account for mobility if estimates of population-level mobility are available. We incorporated nationwide cell phone mobility data to account for this mobility and studied the potential effect of mobility on large-scale epidemiological studies of air pollution.

Unmeasured confounding is a ubiquitous threat to the validity of observational studies on the health effects of air pollution, and therefore, it is crucial to assess how robust our findings are to the presence of omitted confounders. We developed a framework for addressing unmeasured confounding when the observed data contain multiple exposures and multiple outcomes. We showed that this setting is particularly beneficial, as we are able to provide bounds on the magnitude of unmeasured confounding bias by assuming its nature — an assumption that is far weaker than presuming its absence. We also showed that incorporating negative control variables can provide additional information on both the magnitude and direction of unmeasured confounding bias, which can lead to substantially more informative results on the effects of air pollution exposures. These results have the potential to greatly strengthen the causal evidence of the health effects of air pollution by acknowledging uncertainty from the presence of unmeasured confounding.

2.1 RUNNING APPLICATION TO THE MEDICARE POPULATION

Each of these aims corresponds to a different methodological challenge for the causal analysis of air pollution mixtures, but throughout, we grounded our approaches through a running analysis of US air pollution data and their associated health effects in the Medicare population. In aim 1, we focused only on US exposure data for $PM_{2.5}$ components to highlight and address crucial difficulties that arise when estimating causal effects for this set of exposures. In aims 2–4, we used the methodology described above to provide new epidemiological insights into the effects of air pollution mixtures on public health outcomes, assess the robustness of key assumptions necessary for these findings to represent causal effects, and provide a general overview of the effects of air pollution mixtures in this population. We must emphasize that while we developed our approaches within a causal inference framework for the reasons mentioned in Chapter 1, there are still limitations with respect to the Medicare analyses that could prohibit a causal interpretation of our findings. For one, the Medicare analyses documented throughout the report are conducted at the zip code level, and therefore suffer from issues inherent to ecological studies, which limit our ability to make causal claims at the individual level. Additionally, Medicare is limited in the amount of individual-level covariate information available. While we alleviated these issues somewhat by incorporating many area-level covariates as potential confounders, unmeasured confounding can preclude a causal interpretation from our Medicare analyses in Chapters 4 and 5. In Chapter 6, we directly tackled the issue of unmeasured confounding and aimed to assess how robust our estimates are to this ever-present issue of observational studies.

Medicare is a federal health insurance plan for residents of the United States who are over the age of 65 or who are disabled, which provided us with health data on more than 30 million US residents per year, over the years 2000–2016. We had demographic, socioeconomic, and mortality information on all Medicare beneficiaries in the United States during this period. These data provided us with several characteristics for each enrollee, such as their sex, age, and race, and whether they have dual eligibility to Medicaid, which serves as a proxy for low socioeconomic status. Throughout this report, we focused on zip code–level analyses across the United States. This decision was made mostly because of the very high computational burden that running individual-level analyses would present for some of the methods summarized in the report, but we stress that all ideas discussed here would apply to individual-level analyses analogously. We discuss this issue further in Chapter 7. We observed some covariates unique to each zip code from the United States Census Bureau and the Centers for Disease Control's Behavioral Risk Factor Surveillance System. These covariates consisted of smoking rates, median household income, average body mass index (BMI), median house value, population density, education, and percentage of owner-occupied housing. We also measured climate variables such as summer temperature and summer relative humidity from the National Climatic Data Center. Summer relative humidity was calculated as the mean daily maximum relative humidity.

Estimates of air pollution exposure data were taken from one of two sources, depending on the specific pollutant. We incorporated exposure estimates of total $PM_{2.5}$, black carbon (BC), ammonium (NH_4), nitrates (NIT), organic matter (OM), and sulfate (SO_4) levels on a 0.01-by-0.01 degree monthly grid from the Atmospheric Composition Analysis Group.[30] We also incorporated estimates of ozone, elemental carbon (EC), and organic carbon on a 1-by-1 km daily grid from the Socioeconomic Data and Applications Center.[31–33] We did not have exact residential addresses of individuals in Medicare and only knew their residential zip code, and therefore all exposures were aggregated to the yearly level at each zip code. All other covariates were similarly aggregated up to the zip code level by taking their averages or proportions within each zip code. Lastly, all outcomes considered correspond to annual rates for that particular outcome. In the case of mortality, which is the focus of Chapters 4 and 5, this method implies that the outcome was the annual mortality rate within a zip code, which is defined as the number of deaths in that zip code for a particular year divided by the number of person-years in that zip code.

2.1.1 Descriptive Information on Medicare Cohort and Exposure Data

In this section, we have provided some basic, descriptive characteristics of the Medicare cohort as well as the exposure data utilized throughout all Medicare analyses. The same data were used in Chapters 4, 5, and 6; we therefore have provided this information here to give readers a general sense of the population for which we estimated health effects, and the pollution levels to which they were exposed. We observed both exposure and health outcomes data in each year from 2000 to 2016. To provide characteristics of our cohort and how these vary over time, **Table 2** presents the average and standard deviation of the study characteristics across zip codes for the years 2000 and 2016. The characteristics are fairly similar between the 2 years, with some expected differences that are due to trends over time, such as increasing income, housing value, and education rates. Overall, we had a sample size of $n = 34{,}004$ zip codes for the year 2016, with similar sample sizes across the studied years. The average zip code size in the contiguous United States is approximately 75 square miles, although this varies substantially across zip codes of varying population density. To provide context for the spatial variability of key exposure and covariate variables across the contiguous United States, **Figure 1** visualizes zip code–level values of four selected exposures and four selected covariates.

We also investigated the range of exposures for the pollutants considered throughout the report, as well as how these have changed over the period our study covers. We present the minimum, median, and maximum of all zip code–level exposures for the years 2000 and 2016 in **Table 3**. Note that all pollutants are measured in units of $\mu g/m^3$, while ozone is measured in parts per billion (ppb). As expected, there is a general reduction over time for all exposures, except for ozone, across the two periods considered. Additionally, the correlations between the exposures for the year 2016 are presented in **Figure 2**. Generally, there are moderate-to-high correlations between different exposures, although the extent varies by exposure. Additionally, ozone is not strongly correlated with any of the components of particulate matter.

Table 2. Study Characteristics of the Medicare Cohort in the Years 2000 and 2016[a]

	Study Year 2000	Study Year 2016
Mortality rate (deaths/person years)	0.053 (0.037)	0.043 (0.037)
Female (%)	56.0 (8.2)	53.0 (7.9)
Dual eligibility to Medicaid (%)	12.4 (11.1)	12.6 (11.2)
Average age	74.81 (1.730)	74.22 (2.010)
White (%)	89.6 (17.9)	86.0 (19.4)
Average BMI	26.92 (1.100)	27.85 (0.970)
Previous smoker (%)	48.0 (7.0)	46.0 (7.0)
Average income	40,782 (16,543)	55,329 (24,100)
Average house value	110,531 (87,218)	191,853 (173,307)
Graduated high school (%)	62.0 (18.0)	80.0 (15.0)
Population density (per square mile)	1,489 (4,796)	1,612 (5,367)
Owner-occupied housing (%)	73.0 (17.0)	71.0 (18.0)
Summer temperature (degrees Celsius)	28.81 (3.890)	29.93 (3.240)
Summer humidity (%)	91.1 (10.2)	88.2 (11.8)

[a] The numbers shown are the average (standard deviation) of the variables across zip codes.

Table 3. Distributions of Exposures in the Years 2000 and 2016

	Study Year 2000			Study Year 2016		
	Minimum	Median	Maximum	Minimum	Median	Maximum
Total $PM_{2.5}$	1.31	10.53	44.41	1.33	6.82	15.23
Black carbon	0.12	0.69	3.16	0.01	0.57	1.29
Ammonium	0.06	1.18	2.57	0.00	0.24	1.73
Nitrates	0.06	1.01	6.19	0.00	0.65	3.24
Organic matter	0.64	3.16	28.30	0.35	2.95	6.44
Sulfates	0.19	3.00	6.63	0.15	1.24	2.43
Ozone	18.31	38.98	60.17	26.14	39.06	58.97
Elemental carbon	0.09	0.62	2.57	0.03	0.31	1.83
Organic carbon	0.66	2.08	4.60	0.45	1.46	3.45

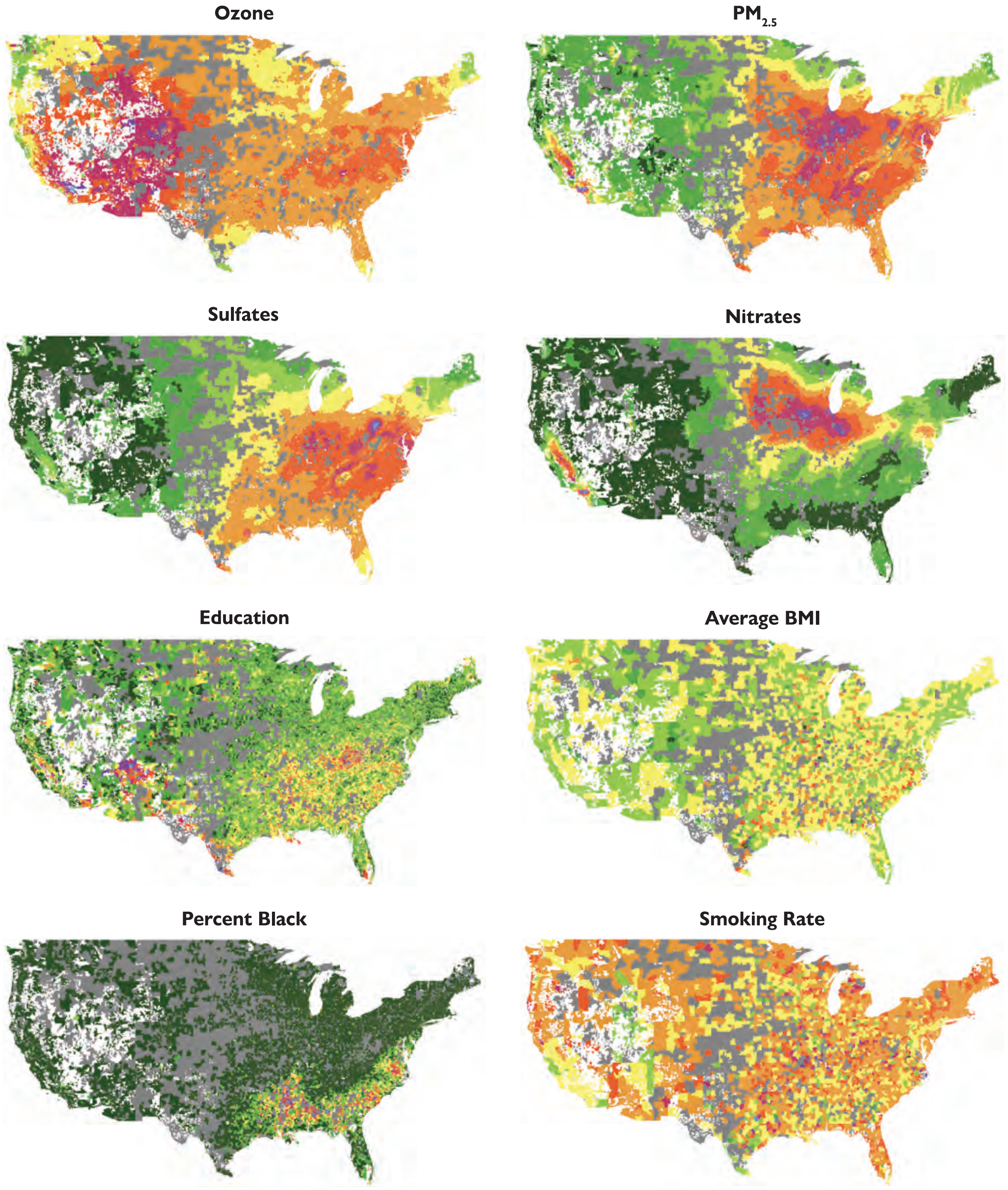
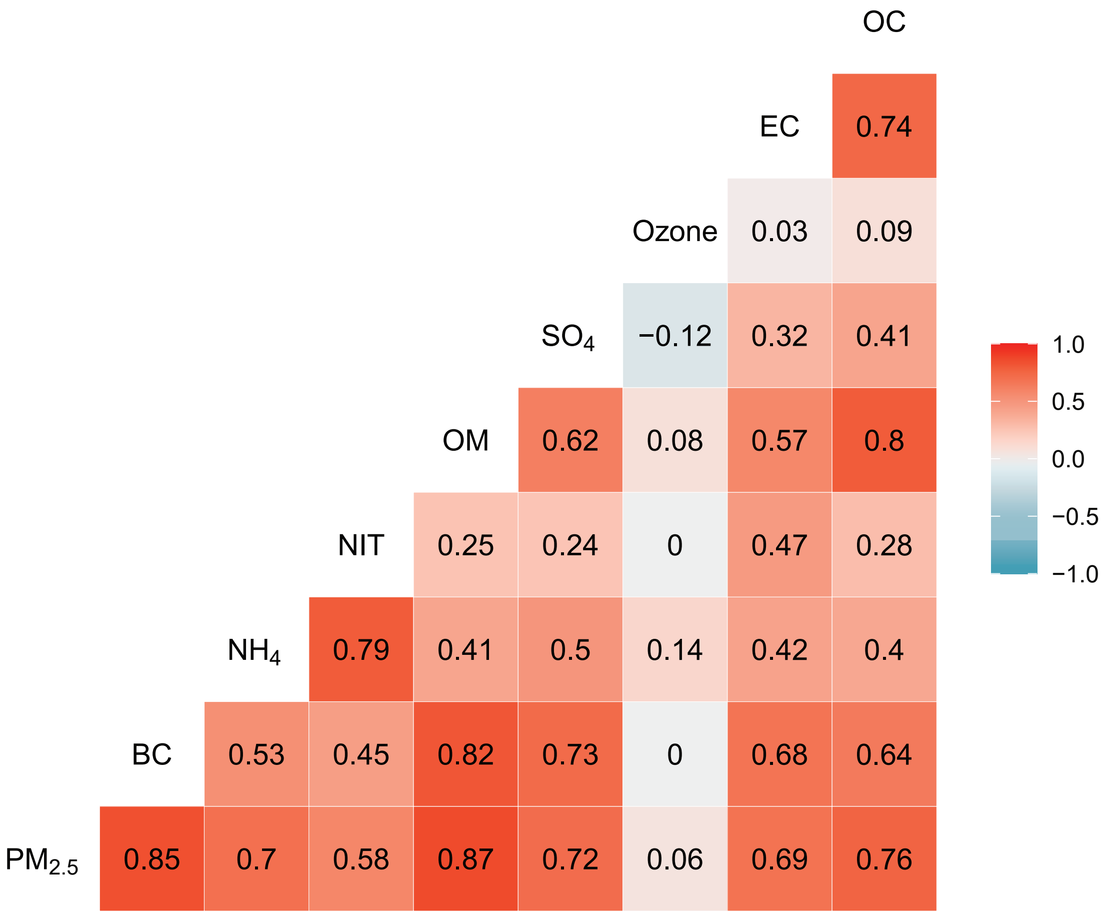

Ozone
$PM_{2.5}$
Sulfates
Nitrates
Education
Average BMI
Percent Black
Smoking Rate

Figure 1. Distribution of selected exposures and zip code level characteristics across the contiguous United States in the year 2010.



	BC	NH_4	NIT	OM	SO_4	Ozone	EC	OC
$PM_{2.5}$	0.85	0.7	0.58	0.87	0.72	0.06	0.69	0.76
BC		0.53	0.45	0.82	0.73	0	0.68	0.64
NH_4			0.79	0.41	0.5	0.14	0.42	0.4
NIT				0.25	0.24	0	0.47	0.28
OM					0.62	0.08	0.57	0.8
SO_4						−0.12	0.32	0.41
Ozone							0.03	0.09
EC								0.74

Figure 2. Correlations between zip code level exposures in the year 2016.

CHAPTER 3: CAUSAL ANALYSIS OF AIR POLLUTION MIXTURES: ESTIMANDS, POSITIVITY, AND EXTRAPOLATION*

3.1 INTRODUCTION

Studying the effect of air pollution on public health is but one epidemiological objective being actively enriched by the appreciation that humans are not exposed to individual contaminants in isolation, but rather to possibly complex mixtures of air pollutants.[6,34–36] Decades of evidence linking air pollution to increased mortality and morbidity[2,4,5,37–39] continue to drive interest in developing statistical methods to analyze health effects of exposure to air pollution (and other environmental) mixtures.[34–36] These methods set out to answer several questions concerning the effect of individual components of a mixture, interactions between mixture components, and the cumulative effect of the mixture.[40]

Increased emphasis on statistical methods relevant for complex mixtures is co-evolving with the continued adoption of explicit causal inference methods in environmental epidemiology.[28,29] Key intersections of these methodological perspectives appear in Wilson and colleagues (2018),[41] who tailored Bayesian model averaging to confounding uncertainty and adjustment in an exposure-wide association study, in Keil and colleagues (2020),[15] who generalized weighted quantile sum regression[14] using principles from g-computation to estimate causal effects of environmental mixtures, and in Traini and colleagues (2022),[42] who pursued a version of generalized propensity scores[43] for multipollutant causal effect estimation. While this and other work have offered important advances to causal inference with air pollution mixtures, the methods for effect estimation draw heavily from extending tools from the univariate exposure case.

This report takes aim at one salient challenge that is particularly pronounced in the context of air pollution mixtures — one that cannot be resolved through obvious extensions from the univariate setting. We focus on (a) defining causal estimands for air pollution mixtures and (b) offering a data-driven way to assess the extent to which observed data actually carry empirical support from which to estimate causal effects of interest. Linking such considerations to policy relevance is essential, with a clear definition of estimands, explicit acknowledgment and discussion of crucial assumptions, and the separation of the causal effect from a statistical model or parameter among the motivating drivers for causal inference.[27,28] A convincing acknowledgment of this argument appears in Keil and colleagues (2021),[44] who estimated the causal effect of a reduction in the air pollution mixture hypothesized to occur if a set of power plants were to be decommissioned. Discussion of that work in Zigler (2021)[45] pointed toward the tension between defining estimands of policy relevance and the threat that observed data contain little or no information to actually estimate them without model-based extrapolation.

We elaborate on how the construction of mixture estimands can quickly lead to violation of the fundamental assumption of positivity, that is, that people have a positive probability of being exposed to levels of a mixture that constitute a causal estimand. We highlight the ease with which positivity violations can arise in studies of air pollution mixtures, and how such violations can produce biased effect estimates that cannot be remedied with flexible statistical models. Our proposal consists of two related efforts: data-driven diagnostics are provided to assess positivity violations with multivariate mixtures, and these diagnostics are then used to provide alternative paths forward with redefined causal estimands for mixtures. We use ambient particulate matter chemical component data across the United States to highlight how easily these difficulties can arise in practice. Throughout, we focus on air pollution mixtures, but key considerations apply for general environmental mixtures of similar dimensions.

3.2 CAUSAL ESTIMANDS FOR ENVIRONMENTAL MIXTURES

3.2.1 Which Mixture Effect? Anchoring Causal Estimands to Interventions

Throughout, we consider data observed as $(Y_i, \boldsymbol{W}_i, \boldsymbol{X}_i)$ for $i = 1, \ldots, n$ observations. We let Y_i be a continuously scaled outcome of interest, but the key ideas hold for other outcome types. We let $\boldsymbol{W}_i$ represent a q-dimensional set of exposures that make up the air pollution mixture. In air pollution studies, q is typically of moderate size, so while the methods here are general, they are most practical for $q \leq 10$. We also observe a p-dimensional vector of covariates $\boldsymbol{X}_i$ that can be used to adjust for the confounding of the relationship between $\boldsymbol{W}_i$ and Y_i. Formulating causal estimands begins with defining $Y_i(\boldsymbol{w}_i)$ as the potential outcome that would be observed for unit i had exposure for that unit been $\boldsymbol{w}_i$. Implicit in this definition of potential outcomes is the stable unit treatment value assumption (SUTVA[46]), encoding that the value of $Y_i(\boldsymbol{w}_i)$ does not depend on the manner in which the value $\boldsymbol{w}_i$ is realized (no multiple versions of treatment), and that the potential outcome for unit i only depends on $\boldsymbol{w}_i$ and not exposures to other units.

Causal estimands can be defined as any comparison between potential outcomes under two different values of $\boldsymbol{w}_i$. However, the nature of multivariate exposures introduces the key question: Which differences in $\boldsymbol{w}$ are actually of interest? While earlier development focused on estimating effects of changes in one mixture component at a time, there is growing appreciation that such estimands may not represent the full implications of changes in the mixture or have any bearing on how a mixture is expected to change in response to any practicable intervention.[44,47] Throughout, we formulate causal estimands as comparisons between observed levels, $\boldsymbol{W}_i$, and levels that might plausibly arise under some specified intervention of interest, $\boldsymbol{W}_{i,int}$. Importantly, the action needn't have actually occurred, such as in the case study of Keil and colleagues (2021)[44] that assumed a uniform proportional reduction in six airborne metals that was expected under the decommissioning of nearby power plants. Alternatively, $\boldsymbol{W}_{i,int}$ may be linked to an actual change in mixture observed following an intervention, such as in Nethery and colleagues (2021)[48] who targeted the impact of the Clean Air Act amendments. Using a specific intervention has the dual purpose of anchoring inference to a practicable action and isolating from the infinitely many hypothetical changes in $\boldsymbol{W}$. An additional benefit, which we prove important in subsequent discussion, is that this helps to ensure that interventions correspond to mixture levels that do not stray too far from those naturally observed. Note that there may be other scientific questions of interest in the analysis of mixtures, such as those studied in Gibson and colleagues (2019)[35] that may not require formalization as effects of interventions.

---

*This chapter was originally published as Antonelli J, Zigler CM. 2024. Causal analysis of air pollution mixtures: estimands, positivity, and extrapolation, Am J Epidemiol 193:1392–1398, https://doi.org/10.1093/aje/kwae115. Adapted with permission from Antonelli et al. 2025.

Let $\Delta_i = \boldsymbol{W}_i - \boldsymbol{W}_{i,int}$ denote the change in the environmental mixture for unit *i* in response to the intervention. Note that this shift is unit-specific to accommodate the case for which an intervention changes the mixture differently across the population being studied. For instance, an intervention to reduce pollution emissions from a point source may differentially affect locations at various distances from the source. The causal estimand defining the effect of interest can be expressed as

$$\frac{1}{n}\sum_{i=1}^{n}\left\{Y_i\left(\boldsymbol{W}_i\right)-Y_i\left(\boldsymbol{W}_{i,int}\right)\right\}, \tag{1}$$

which represents the average impact of the change in mixture following the specified action across the *n* observations under study. Although we focus on (1), the considerations below apply to any estimand comprising a prespecified contrast in $\boldsymbol{W}$. Note that this estimand can be viewed as a sample-level analog to the population-level estimands discussed in Haneuse and Rotnitzky (2013),[49] who examined shifts in a univariate, continuous exposure from the observed exposure levels. One of the main benefits of their estimand, as well as our estimand in (1), is that focus is on exposure levels that could actually be observed for each unit in the study, which is not always the case when looking at prespecified levels of the exposure that are the same for all units. In other contexts, researchers have used similar ideas of grounding estimands in the observed exposure distribution, such as the stochastic estimands in Papadogeorgou and colleagues (2019),[50] which modified exposure assignment depending on the observed propensity scores in their study.

3.2.2 Key Assumptions and the Importance of Mixture Positivity

Estimating (1) with observed data critically rests on several assumptions. Chief among them in observational studies is the assumption of no unmeasured confounding, which states that there are no unmeasured common causes of the mixture and the outcome. The importance of this assumption means that it must be carefully evaluated within the context of any study. Despite this importance, strategies for confounding adjustment are not the focus of this work, in part because confounding considerations are not unique to studies of environmental mixtures, with ample methodologies available from the context of a single exposure. Additionally, we explore the assumption of no unmeasured confounding in detail in Chapter 6.

The final foundational assumption required for the causal inference — and the one on which we focus most here — is that of *positivity*. Letting $f_{W|X}(\boldsymbol{w} \mid \boldsymbol{X} = \boldsymbol{x})$ denote the density of the exposures conditional on the covariates taking value $\boldsymbol{x}$, positivity in our setting can be defined as

$$f_{W|X}(\boldsymbol{W}_{i,int} \mid \boldsymbol{X} = \boldsymbol{X}_i) > 0 \text{ for all } \boldsymbol{W}_{i,int}, \boldsymbol{X}_i.$$

Note that our positivity assumption is unique to the sample-level estimand in (1), and other estimands would require a modified assumption. Many of the ideas presented here, however, extend to other notions of positivity. Positivity violations can be categorized into either structural positivity violations that occur when certain exposure values are not possible for certain covariate values, or finite sample positivity violations that come from not observing certain exposures for a particular covariate value in the sample, even if these are hypothetically possible.[51] We focus on the latter of these two throughout, as it is more common in air pollution epidemiology. The essence of the positivity assumption states that the observed data contain empirical support for estimating values of $Y_i(\boldsymbol{W}_{i,int})$ for all units: inferring the unobserved potential outcome under $\boldsymbol{W}_{i,int}$ for unit *i* requires observations with similar covariate values having been observed with that value of the mixture. Violations of positivity correspond to the absence of such information, in which case inference for causal effects must rely on extrapolation, typically using a parametric model. This assumption introduces threats to causal validity for air pollution mixtures that can be more salient than in studies of univariate exposures.

3.3 ILLUSTRATING POSITIVITY VIOLATIONS AND MODEL EXTRAPOLATION FOR ENVIRONMENTAL MIXTURES

In the case of a univariate exposure, there is ample literature on the positivity assumption and on the threats to validity that can arise amid its violations.[52–54] To illustrate how these issues are exacerbated amid multidimensional consideration of whether a value of $\boldsymbol{W}_{i,int}$ lies within the observed joint distribution of a multivariate mixture, we simulate data with no covariates, $q = 2$ exposures, and a moderately nonlinear exposure-response curve. We use a large dataset of $n = 1{,}000{,}000$ to minimize sampling variability. **Figure 3** shows both the marginal distributions of the two exposures, as well as the joint distribution of the two exposures for both the observed data and a hypothetical intervention distribution. Judging only from the marginal distributions, which would be analogous to positivity assessment in the univariate case, there is substantial overlap in the observed and intervention distribution for exposure 1, and a moderate degree of overlap for exposure 2. Thus, marginal investigation of each mixture component indicates that values of $\boldsymbol{W}_{i,int}$ lie within the observed data. However, investigating the joint mixture distribution clearly indicates otherwise, with values of $\boldsymbol{W}_{i,int}$ falling completely outside of the observed distribution.

To illustrate the consequences of this lack of mixture overlap, we estimate the treatment effect using both a linear model and nonlinear models with 3 and 5 degrees of freedom natural cubic splines, each allowing for interactions between the mixture components. For full details of the estimation strategies and data-generating mechanism for this simulation study, see the appendix in Antonelli and Zigler (2024).[55] The true causal effect defined by (1) in this scenario is 1.45, the linear model estimates it to be 1.22, the 3-degree-of-freedom model estimates it to be 1.17, and the 5-degree-of-freedom model estimates it to be −0.37. The lack of overlap leads to extrapolation, which amplifies bias from model misspecification and results in poor estimates of the causal effect, particularly with more flexible models, which tend to perform worse. This example highlights three relevant issues for environmental mixtures. First, notions of overlap borrowed from the univariate exposure case do not imply overlap in the mixture. Second, overlap for mixtures with $q > 2$ will be increasingly difficult to visualize and diagnose. Lastly, although we have omitted them for illustration, overlap and positivity must be evaluated conditionally on covariates $\boldsymbol{X}$, a requirement that is even more restrictive and less likely to hold in practice.

3.4 POTENTIAL PATHS FORWARD: DIAGNOSING MIXTURE POSITIVITY AND ALTERNATIVE ESTIMANDS

The preceding section supports the need for formal diagnostics to identify and address problems stemming from positivity violations for environmental mixtures. We focus on situations without covariates $\boldsymbol{X}$, although we emphasize that these same issues are likely exacerbated in the presence of confounders and discuss extensions in Section 3.6. Following similar ideas to diagnose model extrapolation presented in King and Zeng (2006),[56] we use the concept of the convex hull of multivariate exposures. Intuitively, the convex hull of the exposures is the smallest polygon that contains all of the observed exposure values, and therefore, it can be thought of as the region where exposure values are actually observed. See **Figure 4** for an illustration of the case of $q = 2$. If $\boldsymbol{W}_{i,int}$ lies outside the convex hull, then extrapolation is necessarily required to estimate potential outcomes under that value of $\boldsymbol{W}_{i,int}$.

For each observation, we let $\boldsymbol{W}_{i,hull}$ denote the point in the convex hull that is closest to the intervention point $\boldsymbol{W}_{i,int}$. Then, a metric quantifying how far the intervention point lies from the convex hull can be calculated as

$$R_i = \frac{\text{distance}\left(\boldsymbol{W}_{i,hull}, \boldsymbol{W}_{i,int}\right)}{\text{distance}\left(\boldsymbol{W}_i, \boldsymbol{W}_{i,int}\right)}$$

where distance$(\boldsymbol{w}_1, \boldsymbol{w}_2)$ is the Euclidean distance between points $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$. The value of R_i is necessarily between 0 and 1, with larger values implying a heavier reliance on model extrapolation, because the distance from the convex hull is large. The best value, $R_i = 0$, implies no model extrapolation because the intervention point already lies in the convex hull. The distribution of R_i across the sample can indicate the extent to which estimates of the quantity in (1) rely on model extrapolation. If R_i is small, such as less than 0.1, for all obser-

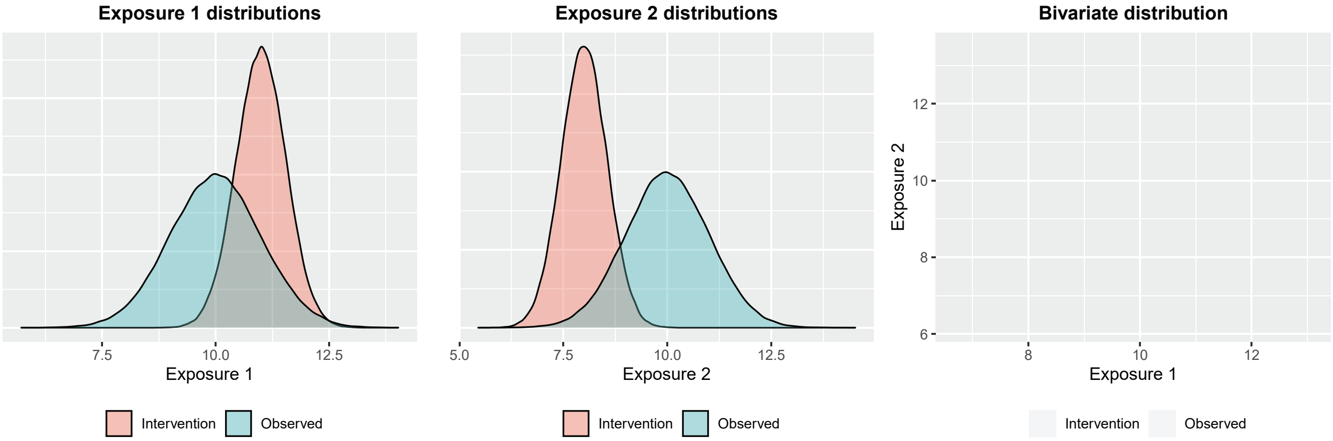


Figure 3. The left panel shows the distribution of exposure 1 in the observed data and in the intervention distribution; the middle panel shows the same distributions, but for exposure 2; the right panel shows the bivariate distribution for exposures 1 and 2. Source: Adapted with permission from *Antonelli et al.* 2024.

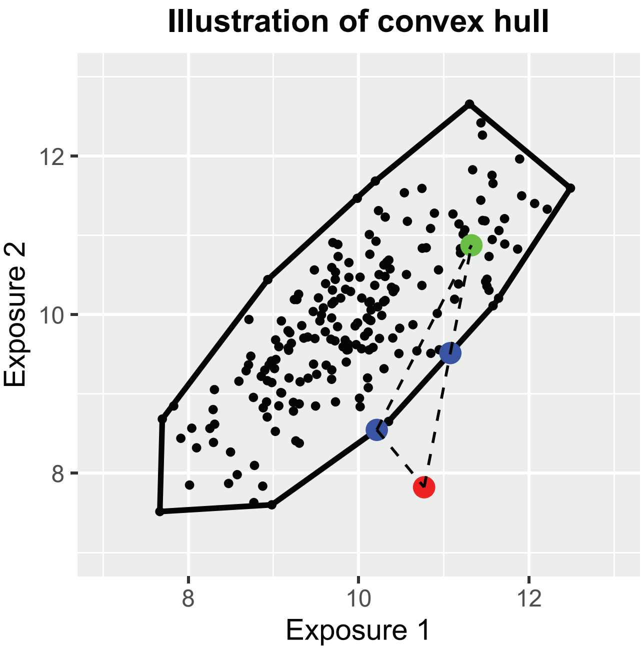


Figure 4. Illustration of the convex hull on a subset of the observed exposures from Section 3.3. The green point is an observed exposure of interest, $\boldsymbol{W}_i$. The red point is the corresponding interventional point of interest, $\boldsymbol{W}_{i,int}$. The blue points represent two different feasible values of the exposure that fall in the convex hull as described in Section 3.4. Source: Adapted with permission from *Antonelli et al.* 2024.

vations, then causal estimates will not be greatly affected by extrapolation. In contrast, high values of R_i in the sample indicate a threat of extrapolation and motivate alternative estimands such as those proposed in the following sections.

3.4.1 Alternative Estimand One: Effects of Feasible Exposure Changes

Upon diagnosing that estimation of (1) would rely on model extrapolation, estimands alternatively defined based on the convex hull may be of interest. One such estimand is to find a "feasible" level of exposure, $\boldsymbol{W}_{i,feas}$, that is similar to the intervention value of interest but not subject to model extrapolation. Similar ideas were explored in Haneuse and Rotnitzky (2013),[49] in which shifts in exposures were only considered if they did not violate the positivity assumption. We explore two options for feasible values:

1. $\boldsymbol{W}_{i,feas} = \boldsymbol{W}_{i,hull}$ so that the feasible value is the point in the convex hull that is closest to $\boldsymbol{W}_{i,int}$
2. Find the largest value of ϕ between 0 and 1 such that $\boldsymbol{W}_{i,feas} = \boldsymbol{W}_i + \phi(\boldsymbol{W}_{i,int} - \boldsymbol{W}_i)$ is in the convex hull.

An illustration of these two values can be found in Figure 4. The first specifies a mixture within the hull of the observed mixture values that is closest to the hypothesized mixture under the intervention. The second value represents a shift in the exposure mixture in the same direction as the hypothesized value under the intervention, but smaller in magnitude to remain within the hull of the observed data. In either case, the estimand in (1) can be decomposed as

$$\frac{1}{n}\sum_{i=1}^{n}\left\{Y_i\left(\boldsymbol{W}_i\right)-Y_i\left(\boldsymbol{W}_{i,int}\right)\right\}=\underbrace{\frac{1}{n}\sum_{i=1}^{n}\left\{Y_i\left(\boldsymbol{W}_i\right)-Y_i\left(\boldsymbol{W}_{i,feas}\right)\right\}}_{feasible\ estimand}+\underbrace{\frac{1}{n}\sum_{i=1}^{n}\left\{Y_i\left(\boldsymbol{W}_{i,feas}\right)-Y_i\left(\boldsymbol{W}_{i,int}\right)\right\}}_{extrapolation\ component}$$

This decomposition explicitly delineates the part of the causal effect that can be estimated with empirical support from the observed data, from the part that is only available through extrapolation.

Table 4 displays estimates of these effect components using the simulated data from Section 3.3. We see that the linear model underestimates both the overall effect and the feasible component, which is expected owing to the nonlinearity of the simulated outcomes. The nonlinear models provide estimates closer to the true value of the feasible component, but get increasingly worse at estimating the extrapolation component as the complexity of the model grows. This illustration highlights the importance of decomposing (1). The feasible estimand is clearly less sensitive to model specification, while the extrapolation component is highly sensitive, with different models providing estimates very far from the truth as misspecification bias becomes amplified in the presence of extrapolation. The median value of R_i in these data is 0.34, suggesting a high degree of model extrapolation that ultimately yields sensitivity to model choice.

3.4.2 ALTERNATIVE ESTIMAND TWO: EFFECTS IN FEASIBLE SUBPOPULATIONS

An alternative option is to focus on a subpopulation with feasible values of $\boldsymbol{W}_{i,int}$. We can define estimands as

Table 4. Estimates of Treatment Effects Under Different Statistical Models from the Simulated Dataset of Section 3.3 When Looking at Feasible Estimands[a]

	True Effect	Linear Model	3 df Model	5 df Model
Overall effect	1.45	1.22	1.17	–0.37
Feasible component	0.81	0.68	0.75	0.74
Extrapolation component	0.64	0.54	0.42	–1.11

[a] Source: Adapted with permission from *Antonelli et al.* 2024.

$$\sum_{i=1}^{n} \gamma_i \left\{ Y_i \left(\boldsymbol{W}_i \right) - Y_i \left(\boldsymbol{W}_{i,int} \right) \right\},$$

where the γ_i are positive and sum to 1 (i.e., $\Sigma_i \gamma_i = 1$). The effect in (1) is a special case where $\gamma_i = 1/n$ for all i, implying that all data points are assigned equal weight. However, when confronting a lack of overlap, weights can be incorporated so that observations requiring more extrapolation receive less weight in the definition of the causal effect. Similar estimands that downweight observations requiring more extrapolation have been used previously, such as those in Vansteelandt and Dukes (2022)[57] or in Li and colleagues (2018). One option is to specify weights $\gamma_i = 0$ for all units with R_i above some threshold, excluding units from the treatment effect that exhibit exposure values requiring large amounts of extrapolation. This approach is commonly done for binary treatments, wherein units with extreme propensity score estimates are trimmed from the sample before estimating treatment effects.[58–62] A continuous alternative would assign values of γ_i inversely proportional to the amount of extrapolation required for each data point. For example, setting

$$\gamma_i = \frac{\left(1 - R_i\right)}{\sum_{j=1}^{n} \left(1 - R_j\right)}$$

would assign the largest weights to points requiring no extrapolation, and decreasing weight as the amount of extrapolation increases. These weighted estimands present a trade-off between interpretability and bias with respect to the overall effect in (1). While less susceptible to bias amplified by model extrapolation, they estimate an effect in some weighted subpopulation that may be difficult to understand or describe. **Table 5** shows estimates from the simulated data from Section 3.3 when using $\gamma_i = 1/n$ and when using weights that are trimmed so that only observations with $R_i < 0.05$ are included. We see that the true value of the estimand changes, but importantly, the lack of required extrapolation leads all models to provide similar estimates of this alternatively defined effect.

Table 5. Estimates of Treatment Effects Under Different Statistical Models from the Simulated Dataset of Section 3.3 When Looking at Trimmed Estimands[a]

	True Value	Linear Model	3 df Model	5 df Model
Equal weights	1.45	1.22	1.17	–0.37
Trimmed weights	0.79	0.66	0.73	0.71

[a]Source: Adapted with permission from *Antonelli et al.* 2024.

3.5 ILLUSTRATION OF UNITED STATES AIR POLLUTION MIXTURES

We illustrate the ideas above in the context of estimating causal effects of ambient $PM_{2.5}$ component mixtures. We only consider exposure data, as the approach would hold direct relevance in estimating the effect of these mixture components on any health outcome. Extensions to adjust for observed confounding are discussed in Section 3.6.

Air pollution exposure data for the year 2015 are obtained from the Atmospheric Composition Analysis Group.[30] We use data on annual average levels of BC, OM, NH_4, NIT, and SO_4 within the contiguous United States. We examine commonly specified causal mixture effects with values of $\boldsymbol{W}_{i,int}$ corresponding to (a) reducing a single mixture component by a specified proportion, or (b) simultaneously reducing all mixture components by a specified proportion. In both cases, the proportion specified is varied from 10% to 90% to investigate increasingly pronounced reductions and use the metrics defined in previous sections to explore the degree of empirical support for estimating the health effects of such shifts in exposures. Note that while we only focus here on settings where all components of the mixture are shifted in the same direction, the same ideas would hold for any shift in exposures caused by the intervention.

Figure 5 shows the percentage of $\boldsymbol{W}_{i,int}$ values that fall inside the convex hull of the observed exposure data as the exposure reduction becomes more pronounced. As expected, more extreme reductions result in fewer values of $\boldsymbol{W}_{i,int}$ falling within the convex hull. The degree to which empirical support suffers varies according to the specific estimand: Specifying a reduction in only OM or BC yields many values of $\boldsymbol{W}_{i,int}$ outside the convex hull, even under relatively small reductions. Other components, such as NIT or NH_4, could be specified with large exposure reductions and still contain a majority of the implied values of $\boldsymbol{W}_{i,int}$ within the convex hull. The estimand specifying simultaneous reduction of all mixture components maintains nearly all observations within the convex hull up to percent reductions of 50%, but larger reductions lose empirical support quickly.

An examination based on the distance from the hull encoded by R_i is depicted in **Figure 6**. Results echo those of Figure 5. A 50% reduction in all components shows that most values of R_i at or near 0, but assuming a 90% reduction yields more values of R_i further from zero. The estimand, considering a 50% reduction in only OM, has many values of R_i far from zero, with a 90% reduction in OM corresponding to an extreme case where extrapolation is required to estimate the effect of such a reduction.

This illustration presents three key points related to estimating causal effects of $PM_{2.5}$ component mixtures. First, the degree to which a particular mixture estimand can be supported with empirical observations versus model extrapolation is dependent on which component(s) are of interest. Second, larger reductions in exposure will be harder to estimate without extrapolation, with most estimands exhibiting very little empirical support for reductions greater than 50%. Finally, note that the simultaneous estimand indicates more robustness to model extrapolation than some of the estimands

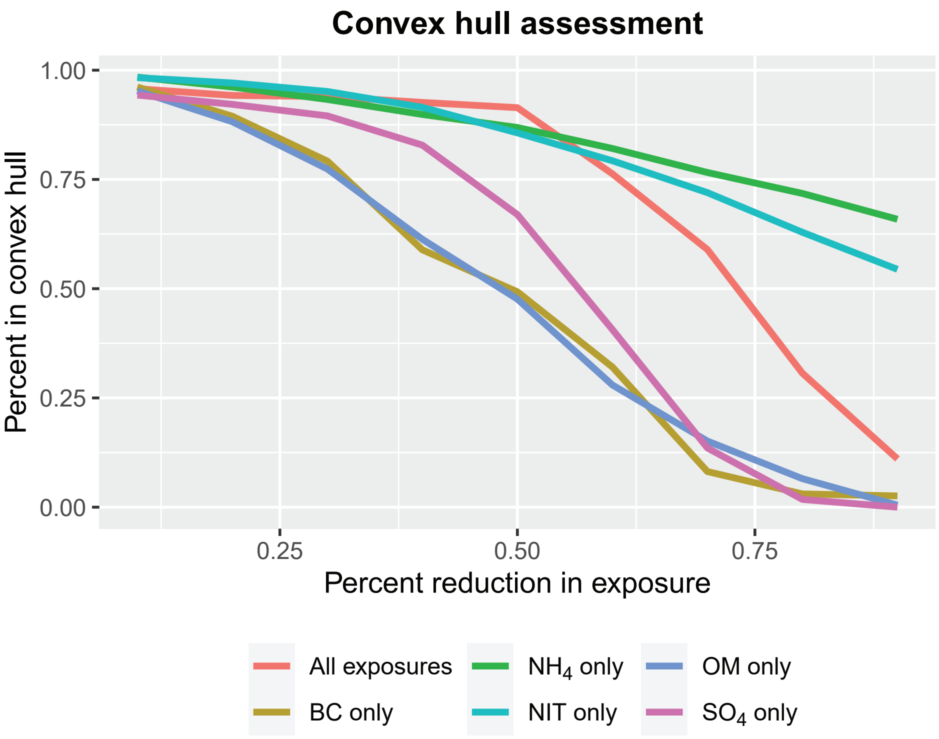


Figure 5. Percentage of $W_{i,int}$ values that fall within the convex hull of the observed mixture data. Source: Adapted with permission from *Antonelli et al.* 2024.

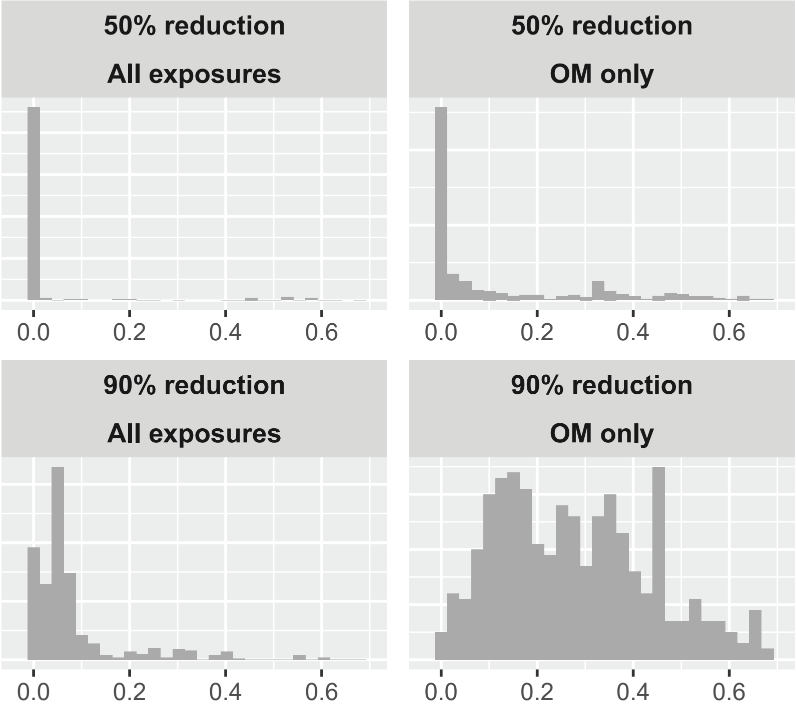


Figure 6. Distribution of R_i when targeting an intervention that reduces exposure to either organic matter only or all components of the mixture simultaneously. Source: Adapted with permission from *Antonelli et al.* 2024.

corresponding to single-component reductions, highlighting the importance of considering the entire joint mixture distribution when diagnosing positivity violations and empirical support for causal effect estimation. Ultimately, the usefulness of the alternatively defined estimands in Section 3.4 must be judged in light of the degree of extrapolation and the implied differences with the overall effect in (1).

3.6 DISCUSSION AND CONCLUSIONS

We have emphasized the importance of defining relevant causal estimands for air pollution mixtures and formalized how positivity violations and reliance on model extrapolation can challenge estimation. Relative to other important methodological considerations for causal inference that can follow development in the case of a univariate exposure, these issues present a particularly pressing challenge when attempting causal inference with mixtures.

Using the convex hull of observed exposures, we quantified the extent to which a specified mixture estimand can be supported by available data. This work has conceptual links to previous work assessing the dangers of model extrapolation when drawing causal inferences across regions of covariate space with limited representation for each of two treatment groups.[56] In the context of air pollution mixtures, when estimands of initial interest correspond to mixture values that are not represented in the observed data, we proposed novel alternative estimands as useful quantities for understanding the effect of environmental mixtures that are less reliant on model extrapolation. Even when some degree of extrapolation is necessary — which is a common setting in the analysis of mixtures — it is essential to assess sensitivity to model specification and evaluate the degree to which ultimate inferences rely on modeling assumptions versus empirical support from the observed data.

Whenever dealing with positivity violations, researchers must thoroughly consider the trade-off between interpretability and feasibility that comes with the modified estimands considered here. The proposed estimands are more feasible in terms of being able to estimate them from the observed data, but they lose interpretability relative to the original estimand of interest, which is driven by scientific expertise or a particular policy choice. There are two ways in which the proposed estimands lose interpretability. For one, we focused on sample-level estimands that are unique to the observed data and may not generalize well to other populations. This approach is somewhat less problematic as it can be addressed by looking at population-level estimands and exploring treatment effect heterogeneity, which can inform how the exposure effects may differ in populations with different characteristics. For this reason, we explore this issue thoroughly in the following chapter by reviewing approaches we developed for exposure effect heterogeneity with air pollution mixtures. A potentially more problematic issue that affects the interpretability of estimands arises when we explore feasible levels of exposure or weighted estimands, which were considered in Section 3.4. These necessarily change the target of inference, and researchers must take great care in ensuring that the proposed estimands maintain public health relevance and still address the scientific question of interest that motivated the study.

The theoretical underpinnings of our discussion of air pollution mixtures extend to more general forms of environmental mixtures. The dominant practical consideration relates to the number of mixture components, which could far exceed that typical of air pollution mixtures in, for example, exposome-wide association studies. While the diagnostics proposed in this report could still be performed, they will likely indicate very little information in the data to estimate causal contrasts of high-dimensional exposures, and the proposed alternative estimands may be of limited use. Additionally, the present work focuses on positivity violations that require extrapolation outside the convex hull of the exposure data, but violations can also arise within the convex hull when exposure values in the hull still differ from observed values. This issue can give rise to the need for *interpolation* to support estimation of causal effects, which, while distinct from extrapolation, might prove equally salient in high-dimensional settings where sparsity within the convex hull of exposures persists. Additional metrics to quantify the degree of sparsity and interpolation are warranted.

For simplicity, the exposition reviewed here did not explicitly consider simultaneous confounding adjustment, which would always be required in observational settings. In technical terms, our discussion focused on marginal positivity with respect to the marginal exposure distribution, whereas notions of conditional positivity (conditional on covariate values, $\boldsymbol{X}$) are analogous and essential. Conditional positivity represents a more restrictive assumption, and therefore any issues with model extrapolation identified in marginal distributions of the exposures are expected to be exacerbated when additionally conditioning on covariates. Because of the importance of this extension, we briefly discuss how one could potentially apply the same ideas once covariates are included. Both of the alternative approaches to inference in the presence of positivity violations discussed in Section 3.4 can be extended to address issues of conditional positivity violations. When defining a feasible level of exposure $W_{i,feas}$, we could find the value of the exposures that is closest to $W_{i,int}$, but also satisfies the positivity assumption in the sense that $\boldsymbol{f}_{W|X}(W_{i,feas} \mid \boldsymbol{X} = \boldsymbol{X}_i) > \boldsymbol{t}$. Here $\boldsymbol{f}_{W|X}$ is the conditional density of the exposures given the covariates, and $\boldsymbol{t} > \boldsymbol{0}$ is a prespecified threshold value that ensures positivity holds. This approach would require an estimate of the distribution of the exposures given the covariates, also known as the generalized propensity score,[43] but would ensure that the exposure effects being examined are those that the data can empirically support. The decomposition of the original estimand into feasible and extrapolation components would still hold and would allow researchers to target exposure effects less susceptible to extrapolation. The second alternative approach involved weighting the units in our sample, where weights γ_i were close to zero for units that required more model-based extrapolation. Similar ideas would hold for conditional positivity, where the weights would depend on the magnitude of $\boldsymbol{f}_{W|X}(W_{i,int} \mid \boldsymbol{X} = \boldsymbol{X}_i)$. Observations with small values of this density are those that would require a high degree of extrapolation, and therefore they should receive small values of γ_i. As seen in Section 3.4, we can even set $\gamma_i = \mathbf{0}$ for observations with overly small values of $\boldsymbol{f}_{W|X}(W_{i,int} \mid \boldsymbol{X} = \boldsymbol{X}_i)$, which represents a trimming estimator related to that seen in Branson and colleagues (2023).[63] A more thorough investigation of these extensions is a topic for future research, which would be required for broader use of this methodology due to the importance of confounding adjustment. Nonetheless, we hope that the discussion of the marginal case can help focus attention on positivity and extrapolation with exposure mixtures and that the provided tools can anchor continued progress in causal effect estimation in the context of environmental mixtures.

CHAPTER 4: HETEROGENEITY OF THE HEALTH EFFECTS OF AIR POLLUTION MIXTURES

4.1 INTRODUCTION

Interest has been growing in the environmental statistics literature about the best ways to model the health effects of environmental mixtures.[35,64,65] In light of this increased interest, several different approaches have been developed to fill this gap. One popular approach is known as Bayesian kernel machine regression (BKMR[17]), which utilizes Gaussian processes combined with model selection to identify important components of a mixture, while allowing for nonlinear relationships between the exposures and the outcome. A number of other approaches have been developed as well that have aimed to address any number of concerns that are commonly found in the analysis of environmental mixtures, such as identifying interactions between exposures, allowing for nonlinear exposure-response curves, and addressing high degrees of correlation between exposures, among others.[14,16,18–23,66,67] These approaches are widely used and have been immensely useful for researchers trying to disentangle complex associations between environmental exposures and health outcomes.

While useful, these approaches all assume that the effect of the mixture is homogeneous with respect to other covariates observed in the data, which is partially driven by the fact that modeling a flexible relationship between a large set of exposures and an outcome is already a difficult problem, even without the additional complication that this effect may vary across people with different characteristics. A natural next step, however, is to allow the mixture effect to vary across subgroups of the population, which is particularly important for air pollution research, as recent analyses of $PM_{2.5}$ have found associations that vary by effect modifiers such as race, age, or socioeconomic status.[68,69] In this work, we develop a flexible, nonparametric Bayesian methodology to estimate the health effects of air pollution mixtures that allows this association to be modified by characteristics of the population being exposed to increased levels of pollution.

Before discussing our approach in more detail, it is important to note that while this methodology was developed for the problem of air pollution mixtures specifically, it has implications in the broader causal inference literature. This problem is typically referred to as treatment effect heterogeneity in this body of literature and has received a lot of attention recently.[70–73] Nearly all of this focus, however, is on the simpler problem for which exposure is a single, binary variable, such as in clinical trials wherein a patient is assigned to one of two treatment arms. Clearly, statistical modeling of effect modification is more complicated for multivariate, continuous exposures than for univariate, binary ones; however, the problem is also more complex from a conceptual perspective, which is because it is not apparent what quantities we wish to estimate from our models. Each exposure in a mixture can have effects that are modified by characteristics of the population, but they might vary in distinct ways, and it is not clear how to summarize this effect modification in a scientifically meaningful manner. In the binary exposure setting, some work has been done to summarize the nature of the heterogeneity in exposure effects by identifying covariates that modify the exposure effect the most,[74–76] but these do not immediately apply to the setting of air pollution mixtures. We address these issues by first defining meaningful quantities that summarize effect modification for air pollution mixture effects and then by developing statistical methodology to estimate such effects. Note that for the remainder of this chapter, we refer to effect modification and exposure effect heterogeneity interchangeably, as both terms are used to describe settings in which the effect of the air pollution mixture is modified by other covariates.

4.2 SUMMARIZING EFFECT MODIFICATION FOR AIR POLLUTION MIXTURES

As in Chapter 3, we assume that we observe $(Y_i, \boldsymbol{X}_i, \boldsymbol{W}_i)$ for $i = 1, \ldots, n$ observations in the data. The outcome of interest is Y_i, which will represent the mortality rates for zip code i. The exposures for zip code i are given by $\boldsymbol{W}_i$, and the covariates that represent characteristics of zip code i are given by $\boldsymbol{X}_i$. As before, $Y_i(\boldsymbol{w})$ denotes the potential outcome that would have been observed if exposure had been set to $\boldsymbol{w}$ for zip code i. We make the following causal assumptions throughout this chapter that are used to identify causal effects from the observed data:

i. *SUTVA*[77]*:* Exposure for one zip code does not affect potential outcomes for other zip codes, and exposure is well defined in the sense that $Y_i = Y_i(\boldsymbol{W}_i)$.

ii. *Positivity:* $0 < f_{W|X}(\boldsymbol{w} \mid \boldsymbol{X} = \boldsymbol{x})$ for all $\boldsymbol{x}$ and $\boldsymbol{w}$ where $f_{W|X}$ is the conditional density of the exposures given the covariates.

iii. *No unmeasured confounding:* $Y(\boldsymbol{w}) \perp \boldsymbol{W} \mid \boldsymbol{X}$ for all $\boldsymbol{w}$.

Positivity was discussed in detail in Chapter 3. SUTVA allows us to link our observed outcomes to the potential outcomes, and the no unmeasured confounding assumption ensures there are no unmeasured variables that confound the exposure–outcome relationship. One issue in this setting is that both exposures and covariates are multivariate, and the exposures are necessarily continuous, which means there are infinitely many estimands one could target in this setting. As a starting point, we can assume that we have two sets of exposure values given by $\boldsymbol{w}_1$ and $\boldsymbol{w}_0$, under which we want to compare the outcomes. These could be two hypothetical levels of the air pollution mixture before and after an intervention is applied to reduce air pollution. They could also represent levels above and below the current National Ambient Air Quality Standards levels. Then we could examine

$$\tau_{w_1,w_0}(\boldsymbol{x}) = E\{Y(\boldsymbol{w}_1) - Y(\boldsymbol{w}_0) \mid \boldsymbol{X} = \boldsymbol{x}\},$$

which represents the expected effect on mortality that this shift in the air pollution mixture would cause, for a zip code with characteristics $\boldsymbol{x}$. We could examine this function across different zip code characteristics to see what type of individuals are most impacted by such a change in the air pollution mixture.

In other scenarios, however, there may not be two distinct exposure levels of interest, and we care about heterogeneity of the overall mixture effect. Suppose then that we have some baseline level of the mixture denoted by $\boldsymbol{w}_0$. Importantly, this choice of $\boldsymbol{w}_0$ should not greatly affect the results that follow. Given this choice, we can define

$$\tau_{w_0}(\boldsymbol{x}, \boldsymbol{w}) = E\{Y(\boldsymbol{w}) - Y(\boldsymbol{w}_0) \mid \boldsymbol{X} = \boldsymbol{x}\},$$

which is a function of both covariates and exposures and describes how the outcome of interest varies by these two quantities. This quantity is not particularly interpretable as it depends on two multivariate quantities and therefore cannot be easily visualized or summarized. To address this issue, we develop what we subsequently refer to as multivariate treatment effect variable importance measures (MTE-VIM) that summarize the extent to which each covariate modifies the effect of the exposures. To construct these importance metrics, we first define the following two quantities:

$$\phi = E_W\left[Var_X\left\{\tau_{w_0}(\boldsymbol{x}, \boldsymbol{w})\right\}\right]$$

$$\phi_j = E_W\left(Var_{X_{-j}}\left[E_{X_j|X_{-j}}\left\{\tau_{w_0}(\boldsymbol{x}, \boldsymbol{w})\right\}\right]\right)$$

Here we let $\boldsymbol{X}_{-j}$ represent all covariates excluding the j-th covariate. These quantities are closely related to a variety of variable importance metrics that have been proposed in recent literature, such as leave-one-covariate-out,[78] or the minimum mean squared error gap and the corresponding Floodgate procedure.[79] These approaches all aim to infer how much prediction accuracy is lost without the inclusion of a particular covariate. One key difference here is that we are applying these ideas to the treatment effect function $\tau_{w_0}(\boldsymbol{x}, \boldsymbol{w})$, and therefore are aiming to infer how much effect modification exists, and how much of this effect modification is driven by particular covariates. Specifically, ϕ is a measure of the overall amount of heterogeneity in the mixture effect due to covariates $\boldsymbol{X}$, which can also be described as the overall amount of effect modification. The second quantity ϕ_j is interpreted as the overall amount of heterogeneity in the mixture effect due to covariates $\boldsymbol{X}$, which can be explained without covariate *j*. ϕ_j is necessarily less than ϕ, and small values of ϕ_j relative to ϕ indicate that covariate *j* is an important modifier of the effect of the air pollution mixture. Our final quantity that we focus on is given by

$$\psi_j = 1 - \frac{\phi_j}{\phi},$$

which is the proportion of the treatment effect heterogeneity of $\tau_{w_0}(\boldsymbol{x}, \boldsymbol{w})$ that cannot be explained without covariate *j*. This quantity is useful as it is scale-free and lies between 0 and 1, with values close to 0 indicating that covariate *j* does not greatly modify the exposure effect, with larger values indicating that the particular covariate is a very important driver of heterogeneity of the mixture effect.

One difficulty with this estimand is that statistical inference, particularly the construction of confidence or credible intervals, becomes challenging near the null value of $\psi_j = 0$ when covariate *j* is not an effect modifier. The main consequence of this issue is that confidence intervals constructed in a standard manner will contain the true parameter ψ_j less often than the stated coverage rate. This point was discussed in detail in Verdinelli and Wasserman (2023),[80] who proposed ad hoc adjustments to address this issue, which could potentially be applied in our setting, although we leave this as a topic for future research. Despite this issue, the proposed importance measures still provide users with a relative measure of the importance of each covariate in the overall exposure effect, which helps identify subgroups of the population most affected by the exposures.

4.3 STATISTICAL MODELING FRAMEWORK

Given that the positivity, SUTVA, and no unmeasured confounding assumptions all hold, we are able to write estimands of interest in terms of quantities that can be estimated from our observed data. Specifically, under these assumptions, we have

$$E\{Y(\boldsymbol{w}) \mid \boldsymbol{X} = \boldsymbol{x}\} = E(Y \mid \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{W} = \boldsymbol{w}),$$

which is simply a conditional expectation that we can estimate using regression approaches. Much of the statistical literature on environmental mixtures focuses on estimating this model in a way that allows for the nonlinear effects of $\boldsymbol{W}$ on $\boldsymbol{Y}$, but does not focus on the effects of $\boldsymbol{X}$. Most models take the form

$$E(Y \mid \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{W} = \boldsymbol{w}) = \boldsymbol{x}\boldsymbol{\beta} + g(\boldsymbol{w}),$$

and they incorporate flexible models for $g(\boldsymbol{w})$ that allow for interactions between mixture components and nonlinear exposure-response curves. These models all make an additivity assumption by separating the effects of $\boldsymbol{X}$ and $\boldsymbol{W}$ in a way that ensures the effect of $\boldsymbol{w}$ on Y does not depend on covariates $\boldsymbol{X}$. To drop this assumption, and to allow the effect of the mixture to change with $\boldsymbol{x}$, we fit a model of the form

$$E(Y \mid \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{W} = \boldsymbol{w}) = c + f(\boldsymbol{x}) + g(\boldsymbol{w}) + \sum_{j=1}^{p} h_j(x_j, \boldsymbol{w}),$$

where $\boldsymbol{x}_j$ is the j-th component of $\boldsymbol{x}$. This allows for interactions between exposures and covariates, which will induce heterogeneity in the mixture effect. We want to ensure that $h_j(x_j, \boldsymbol{w})$ truly captures interactions and does not simply capture main effects of x_j or $\boldsymbol{w}$, which should be captured by $f(\boldsymbol{x})$ and $g(\boldsymbol{w})$, respectively. For this reason, we impose that $h_j(x_j, \boldsymbol{w}) = h_j^{cov}(x_j) h_j^{exp}(\boldsymbol{w})$ to ensure this function only captures heterogeneity. Even with this restriction, the individual functions are not currently identifiable from the data, and additional constraints must be imposed on the model. We detail these constraints and other technical details of this model in Shin and colleagues (2024).[81] The main issue we focus on in this report is how we choose to model $f(\boldsymbol{x})$, $g(\boldsymbol{w})$, and $h_j(x_j, \boldsymbol{w})$.

4.3.1 Variations of Bayesian Additive Regression Trees

For each of the individual functions, we incorporate recent extensions to Bayesian additive regression trees (BART),[82] which have been shown to work exceedingly well in a wide range of contexts, including recent causal inference estimation competitions.[83] We will not discuss many of the technical details of different BART-based prior distributions, but we point readers to recent reviews in Linero (2017) and Hill and colleagues (2020) for approachable backgrounds to these models.[84,85] The standard BART model, if applied to $f(\boldsymbol{x})$, takes the form

$$f(x) = \sum_{m=1}^{M} Tree(\boldsymbol{x}, \mathcal{T}_{\mathrm{m}}, \mathcal{M}_m),$$

where $Tree(\boldsymbol{x}, \mathcal{T}_{\mathrm{m}}, \mathcal{M}_m)$ corresponds to a regression tree function of $\boldsymbol{x}$ with tree structure $\mathcal{T}_{\mathrm{m}}$ and predictions in terminal nodes given by $\mathcal{M}_m$. To provide intuition (conceptual insight) for these values, **Figure 7** shows a single regression tree as a function of age and smoking status. The tree structure $\mathcal{T}_{\mathrm{m}}$ represents the fact that we first split the covariates on whether age is above 70, and then split on smoking status conditional on age being below 70. The terminal node values $\mathcal{M}_m$ are given by (5.20, 6.14, 6.45), and they are the output of the function in each node of the tree. BART fits M of these regression trees and averages the results across all trees. The prior distribution for BART encourages shallow trees, or trees without many splits, and places shrinkage priors on the terminal node values. The default prior distribution for BART has been shown to work well in general settings, and it requires very little tuning by the user. The only potential issue is that the discrete nature of decision trees may not approximate smooth exposure-response curves between air pollutants and health outcomes. For this reason, we adopt a recent extension, referred to as SoftBART,[86] which should approximate smooth functions better than standard BART. **Figure 8** shows this for a simulated exposure-response curve, as we can see that the SoftBART model is much smoother and generally more accurate than the standard BART model.

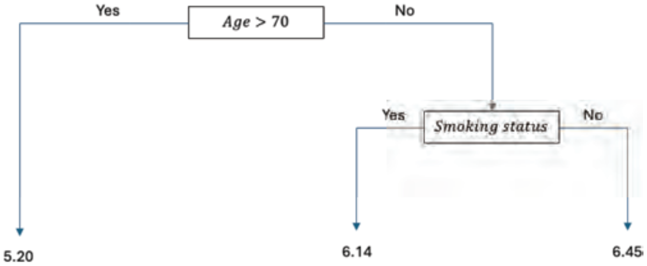


Figure 7. Illustration of a single regression tree.

As a result, we use the SoftBART model for both $f(\boldsymbol{x})$ and $g(\boldsymbol{w})$, but this does not immediately apply to modeling $h_j(x_j, \boldsymbol{w})$, which captures interactions between exposures and covariates. Following Li and colleagues (2022),[87] we assume that $h_j(x_j, \boldsymbol{w}) = \sum_{m=1}^{M} \mathcal{B}_m(x_j) Tree(\boldsymbol{w}, \mathcal{T}_{\mathrm{m}}, \mathcal{M}_m)$, where $\mathcal{B}_m(x_j) = \sqrt{2}\cos(\omega_m x_j + b_m)$ for some constants ω_m and b_m, which we also estimate. We also use SoftBART for the trees in these interaction functions to obtain the same benefits as for the main effect functions. This model is quite flexible, but importantly, we can also place shrinkage priors on the parameters of the $h_j(x_j, \boldsymbol{w})$ functions so that we can shrink these functions toward zero if the effect of the mixture is in fact homogeneous. Simulation studies in Shin and colleagues (2024)[81] showed that this modeling approach outperforms popular models, such as BKMR, even when there is no heterogeneity of the mixture effect, and is well suited to capturing heterogeneity of mixture effects.

4.4 DATA ANALYSIS

In this section, we incorporate the aforementioned statistical methodology into our analysis of the US Medicare cohort.

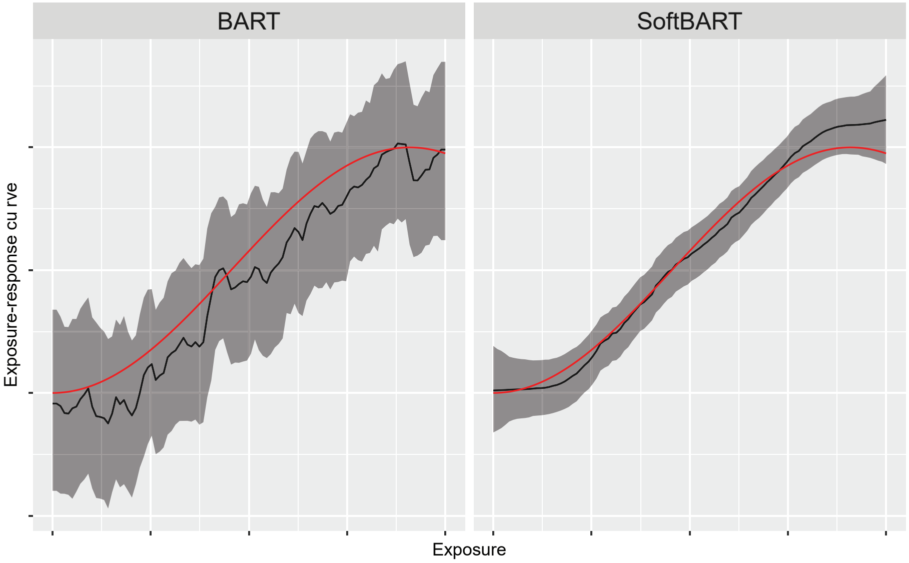


Figure 8. Comparison between BART and SoftBART on a simulated example. The true exposure-response curve is given by the red line.

For this analysis, we look at the following list of exposures: $PM_{2.5}$, NH_4, NIT, SO_4, EC, organic carbon, and ozone. The set of covariates included in the model as both confounders and modifiers of the mixture effect are percentage dual eligibility to Medicaid, average age, percentage female, percentage White individuals, average BMI, smoking rates, median household income, population density, percentage of owner-occupied housing, and median house value. For more details on these covariates and exposures, as well as where they were obtained, see Chapter 2. The outcome of interest is the mortality rate within a particular zip code and year, and our analysis was conducted separately by year. This approach was taken for computational reasons, as it was too large a computational burden to fit the aforementioned models on one dataset that contains all zip code and year pairs. Our approach was a Bayesian model, and we ran the Markov chain Monte Carlo (MCMC) algorithm for estimating the posterior distribution of all model parameters for 5,000 iterations, in which we removed the first 1,000 and only kept every fourth iteration after this burn-in period. Convergence diagnostics, such as the potential scale reduction factor,[88] were examined along with visual inspection of trace plots to confirm MCMC convergence.

4.4.1 Results

Before examining the heterogeneity of the health effects of the air pollution mixture, we examined estimates of marginal effects that average over the covariate distribution in the population. Specifically, we estimated $E[Y(\boldsymbol{w}_1) - Y(\boldsymbol{w}_0)]$, where we let $\boldsymbol{w}_1$ represent the third quartile of all exposures, and $\boldsymbol{w}_0$ represents the first quartile, which implies that a positive estimate would denote a detrimental effect of air pollution on mortality. These results can be found in **Figure 9**, which shows these results broken down by year, along with an overall estimate. The overall point estimate was taken as the average of the yearly estimates, and the corresponding interval was derived by assuming a normal approximation. To obtain the variance of this normal distribution, the covariance matrix of all the point estimates was assumed to follow an autoregressive structure with lag 1 and a high correlation of 0.8 between adjacent years. While statistical significance varies across years, in each year the model estimates that increasing pollution leads to increases in mortality, with the mortality rate estimated to increase by 0.5 deaths per 100,000 person-years when averaged over all years in the study.

These results are indicative of a harmful effect of pollution on mortality, but we then investigated whether this effect is larger in certain subgroups of the Medicare population than in others. To understand this question better, we first examined the MTE-VIMs, which quantify how much each covariate modifies the exposure effect. **Table 6** shows estimates of the MTE-VIM for each covariate, averaging over the estimates for each individual year. Dual eligibility to Medicaid has the largest variable importance metric, followed closely by the percentage of White individuals in a zip code. The interpretation of these quantities is such that approximately 22% of the effect modification can only be captured by dual eligibility to Medicaid. Other variables such as age, population density, and gender contribute somewhat to effect modification, but much less so than the aforementioned characteristics. We now examine these two key contributors to effect modification in greater detail, along with average age, which is both a well-established modifier of air pollution effects and the variable with the third-highest MTE-VIM value.

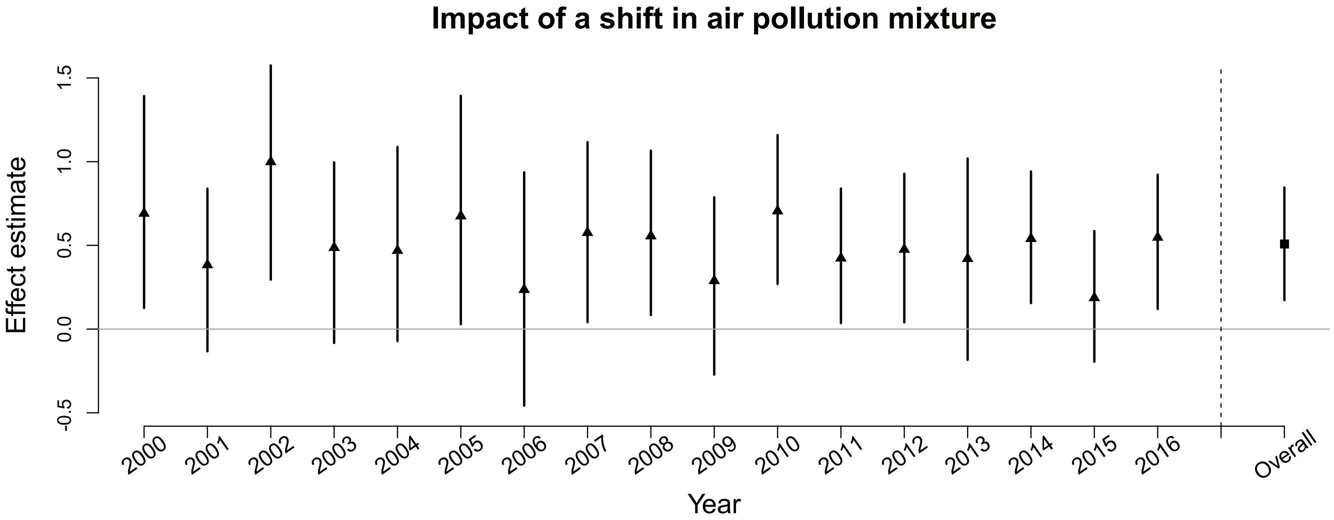


Figure 9. Yearly estimates and 95% credible intervals of the effect of increasing all components of the air pollution mixture from their first to third quartiles.

Table 6. MTE-VIM for Each Covariate Averaged Across All Years of the Study

Variable	Average MTE-VIM
Percentage female	0.065
Dual eligibility to Medicaid	0.222
Average age	0.089
Percentage White	0.193
Average BMI	0.004
Percentage smoking	0.004
Average income	0.021
Median house value	0.018
Percentage with high education	0.018
Population density	0.061
Percentage owner-occupied housing	0.004
Average summer temperature	0.033
Average summer precipitation	0.038

In **Figure 10,** we see the posterior means and corresponding 95% credible intervals for the MTE-VIM value of these three covariates across all years. Note that credible intervals were calculated as the 0.025 and 0.975 quantiles of the posterior distribution of the MTE-VIM, and they need not be symmetric around the posterior mean, particularly for such a parameter, which is required to be between 0 and 1. One key takeaway from this plot is that there is a lot more variability in these estimates compared with the marginal effect estimates seen in Figure 9. Effect modification is a more complex quantity to estimate and tends to have more uncertainty around it than marginal effect estimates. Nonetheless, these results clearly show that dual eligibility to Medicaid and race are very important for understanding effect modification for air pollution mixtures on mortality. In 2014, it is estimated that dual eligibility to Medicaid accounts for over 50% of the variability in exposure effects, while race explains nearly 30% of this variability in seven of the study years assessed. There is more uncertainty in the estimates for average age, as its MTE-VIM is effectively zero for many years but is substantially higher in five of the years studied. Part of the reason for this uncertainty could be explained by the fact that the range of average ages within zip codes in the Medicare population is very small. The 10th and 90th quantiles of average age in our study were approximately 71 and 75, respectively, which is not a large age range to consider and highlights one limitation of our study design that assessed zip code–level data rather than individual-level data. While this analysis provides insight into which variables are effect modifiers for the effect of the air pollution mixture that was studied, the extent of this effect modification may differ in individual-level studies with greater variability in covariate values. The example discussed above, showing the small amount of variability in average age within our study, highlights this challenge. There may not be large differences in the exposure effects for individuals of ages 71 and 75, but there may be very large differences in exposure effects for individuals of ages 65 and 85, which we cannot detect in our setting with zip code–level data.

While the MTE-VIM values provide information about which covariates drive the effect modification of the exposure effect, we can further examine the direction of the relationship between these covariates and susceptibility to air pollution exposure. To do this, we can visualize $E\{Y(\boldsymbol{w}_1) - Y(\boldsymbol{w}_0) \mid \boldsymbol{X} = \boldsymbol{x}\}$

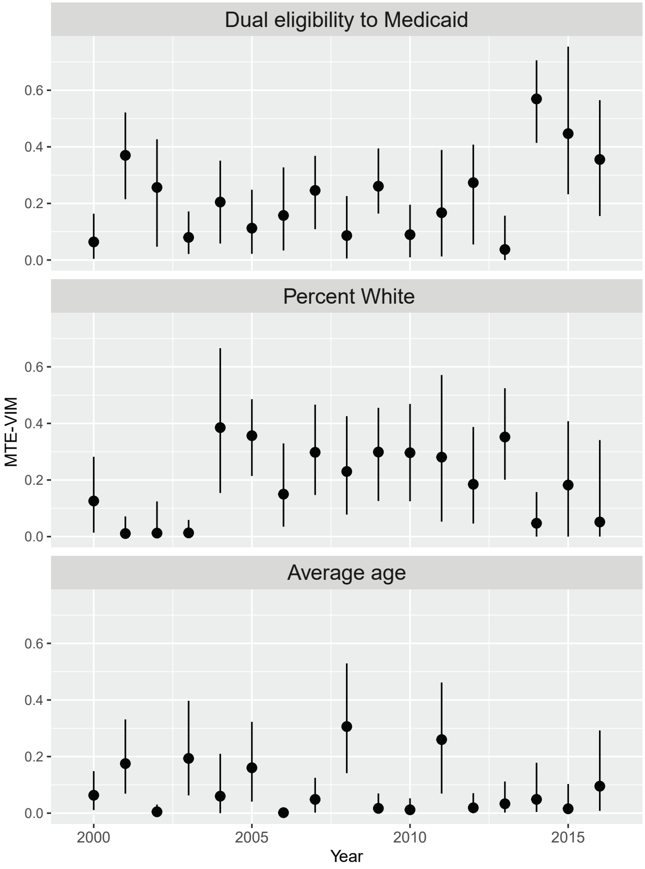


Figure 10. Posterior means and 95% credible intervals of the MTE-VIM across all study years for dual eligibility to Medicaid, race, and age.

for a particular set of $\boldsymbol{x}$ values. We can fix all but one of the covariate values at a chosen value; in this case, we choose their mean value, and then increase the value of the covariate of interest across the range of its plausible values. An increasing trend in this function would show that increasing values of this covariate lead to larger effects of the exposure on mortality. **Figure 11** shows these functions for the three covariates of interest, across four evenly spaced years in our study, along with an overall estimate that averages across all years. The pooling across years was done in the same manner as for the marginal effects, which was described above. We see a consistent set of findings across all years for dual eligibility to Medicaid. As the percentage of Medicare enrollees who have dual eligibility goes up, the exposure effects become larger and more harmful. Dual eligibility to Medicaid is an indicator of low socioeconomic status, and therefore, this result shows that the negative health effects of air pollution are more pronounced among individuals with low socioeconomic status. The results for age are also fairly intuitive, as the general trend across years is that increasing age leads to increased effects of exposure to air pollution, although these effects are much smaller than those from dual eligibility to Medicaid. The results are less clear for race, as some years indicate an increasing trend in the percentage of White individuals, while others show a decreasing trend. Such mixed findings have been found before, as previous studies of the health effects of $PM_{2.5}$ have shown smaller effects in areas with larger proportions of Black residents.[69] One potential explanation for this finding is survival bias.[89,69] The existing literature shows that minorities are disproportionately exposed to air pollution,[90,91] which can result in premature deaths of racial minorities susceptible to air pollution even before enrolling in Medicare.[92–94] Overall, our mixed findings across years, along with this prior research, show that race likely plays a modifying role in the health effects of air pollution, but further research is needed to uncover the exact nature of this relationship.

4.5 DISCUSSION AND CONCLUSIONS

In this chapter, we discussed a new methodology we have developed that tackles issues prevalent in the causal analysis of environmental mixtures. From a statistical perspective, we developed Bayesian nonparametric models that allow for nonlinear effects with interactions between exposures, while also allowing the effect of the mixture to depend on other covariates. This approach should provide more accurate estimates of the health effects of environmental mixtures, but also provides information about which parts of the population are most affected by air pollution, which is important from a public health perspective. Beyond statistical modeling, we also provided users with new estimands that help simplify the complex, heterogeneous effects of mixtures, which allows users to identify covariates that modify the mixture effect the most, leading to the identification of subgroups most susceptible to the adverse health effects of air pollution.

These new estimands and statistical approaches proved useful in the analysis of air pollution mixtures in the US Medicare population, where our models estimated that the effect of air pollution was indeed modified by certain characteristics. Consistent with prior literature, we found that increasing air pollution levels leads to increases in mortality. Additionally, we estimated that the effects of air pollution on mortality are increased for older individuals and for those of a lower socioeconomic status, as indicated by their dual eligibility for Medicaid. We also found that race modifies this effect, but further research is required to understand the nature of this effect modification. This analysis relied on multiple assumptions, which we address in subsequent chapters of this report. For one, exposure was assigned to individuals based on exposure in their home zip code, although individuals travel to different zip codes during daily activities. We evaluate the effect of this assumption and propose potential remedies to address this issue in Chapter 5. Additionally, these results relied on the assumption that there are no unmeasured confounders of the exposure–outcome relationship. Given the strength and importance of this assumption to the estimation of exposure effects such as those explored in this chapter, we assess the robustness of our findings to unmeasured confounding bias in Chapter 6.

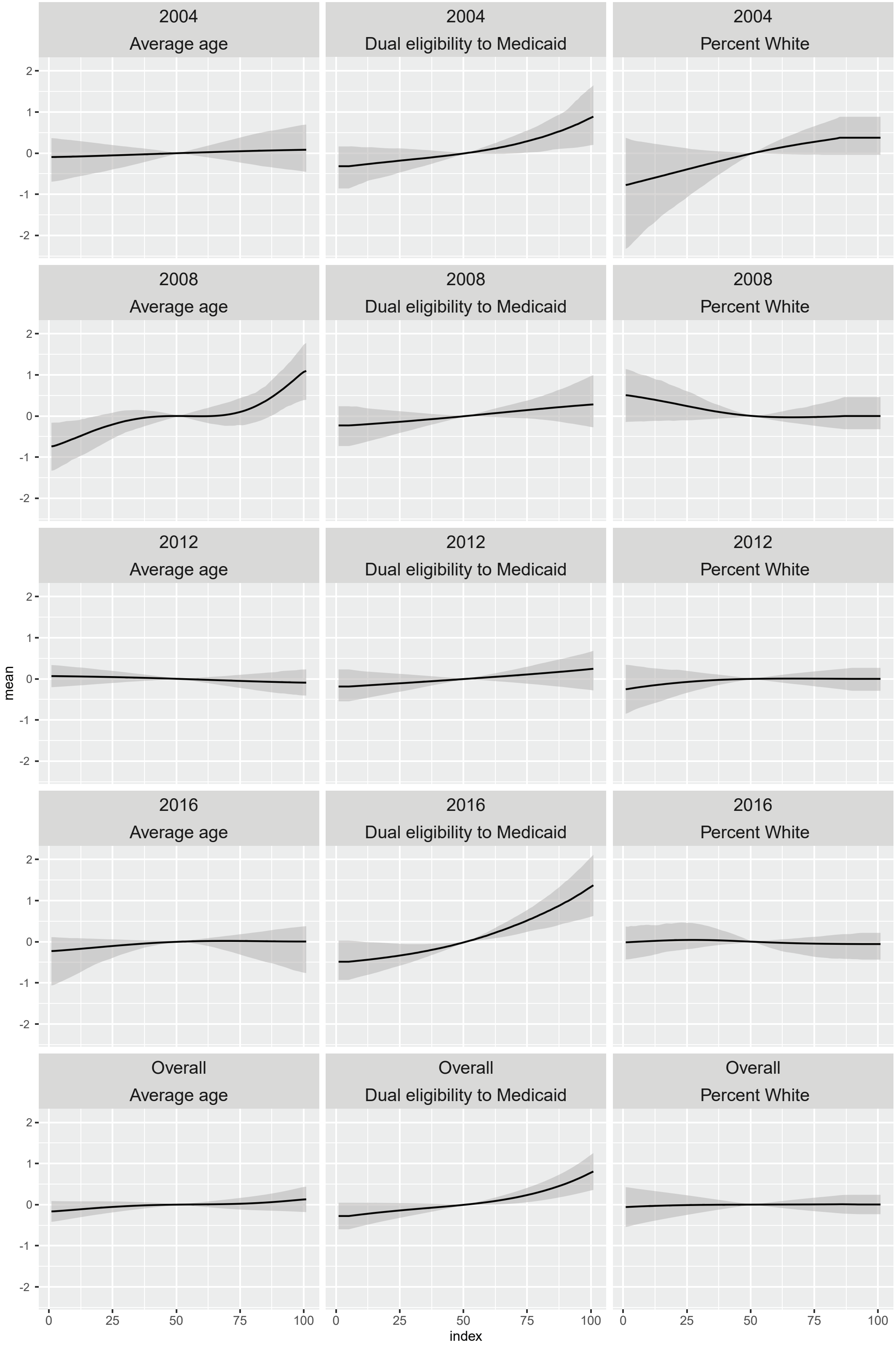


Figure 11. Posterior means and corresponding 95% credible intervals of conditional exposure effects when the covariate of interest is varied, but all other covariates remain fixed at their average value.

CHAPTER 5: ACCOUNTING FOR MOBILITY WHEN ESTIMATING THE HEALTH EFFECTS OF AIR POLLUTION MIXTURES

5.1 INTRODUCTION

Exposure assessment is a critical component of research that examines the relationship between air pollution exposure and public health. It involves assigning each study subject a value for air pollution exposure, typically informed by their home address or geographic area, such as zip code, even though people often travel for work or other activities. Several studies have examined this issue and shown that ignoring daily mobility can introduce exposure measurement error, as exposure in home regions may differ from that in areas where individuals travel.[95–97] Not surprisingly, the degree of this measurement error is strongly associated with the nature of individuals' mobility patterns and the spatial variability in the pollutant being studied.[98] Cell phone mobility data has been used to approximate mobility in New York City, and it was found that estimates of population exposure to air pollution change substantially once mobility is taken into account.[99] Recent studies have also shown that exposure to pollution through daily mobility may affect low-income populations more, which exacerbates existing disparities in air pollution exposure.[100] Another key challenge, which we focus on in this chapter, is the downstream consequences for subsequent health effect estimation. Measurement error in exposures is well known to bias estimates of exposure effects, which suggests that ignoring daily mobility may bias the causal effects of air pollution mixtures. Setton and colleagues (2011)[101] studied the effect of this measurement error and argued that it followed an extension of the classical measurement error framework that allows for correlation between the measurement error and the true, unknown level of exposure, suggesting that ignoring mobility can bias health effect estimates toward the null of no effect of the exposure on health outcomes. In this chapter, we formalize this issue within a causal inference framework and provide insights into the bias that can occur when mobility is ignored in more general settings or when a multivariate mixture is studied instead of a single exposure. We also summarize how to correct for this bias by incorporating estimates of population-level mobility obtained from cell phone mobility data.

While the focus of this chapter is on the implications that mobility has on health effect estimation, mobility has broader implications in the causal inference literature on interference. Interference occurs when the outcomes of one unit in a study are affected by the exposures of other units.[102–105] These ideas have been explored specifically in the context of air pollution to better understand the effects of interventions on power plants[106,107] and those of the United States Environmental Protection Agency's nonattainment designations.[108] In the context of power plants, interference occurs because an intervention to reduce exposure from one power plant can have impacts on pollution levels in many different areas, depending on wind direction and other atmospheric processes. We formalize the issue of mobility as an interference problem, given that exposure levels in one geographic area can affect individuals in other areas if those individuals travel to that area during daily activities. We build on the interference literature by studying what happens when interference (mobility) is ignored[109] or if the nature of the interference is misspecified and estimated with error.[110,111] Additionally, nearly all work in interference focuses on single, binary exposures, whereas we develop a methodology for interference when there are multiple, continuous exposures.

5.2 FORMALIZING THE EFFECT OF MOBILITY

As in the previous chapters, we denote the outcome for zip code i by Y_i, the exposure by $\boldsymbol{W}_i$, and the additional set of covariates by $\boldsymbol{X}_i$. A key distinction between this chapter and the earlier chapters centers on the definition of potential outcomes. Previously, we used $Y_i(\boldsymbol{w})$ to denote the outcome for zip code i that would have been observed if exposure to the air pollution mixture in zip code i were set to $\boldsymbol{w}$. This definition is not sufficient when dealing with mobility and interference because the exposure levels in other zip codes outside of zip code i might affect the outcome for zip code i, if individuals in zip code i travel to those other regions. As a simple motivating example, consider two situations in which a zip code has a particular exposure level, but in one situation, its neighboring zip codes have very high exposure levels, while in the other situation, the neighboring areas have very low exposure levels. Individuals in this zip code will likely travel to neighboring areas, and therefore we would expect outcomes to be worse in the first situation where neighboring areas have very high exposure levels, illustrating that potential outcomes in the presence of interference defined by $Y_i(\boldsymbol{w})$ are not well defined because they must depend on neighboring exposure levels as well.

In light of this issue, we begin by denoting potential outcomes by $Y_i(\boldsymbol{w}^{all})$, where $\boldsymbol{w}^{all} = (\boldsymbol{w}^1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n)'$ represents the exposure levels in *all* areas in our study. Therefore, this potential outcome represents the outcome we would observe for zip code i, had exposure across the entire study region been fixed to $\boldsymbol{w}^{all}$. While useful in the interference setting, this definition of potential outcomes allows for outcomes at zip code i to depend on all other exposures in arbitrarily complex ways, estimating any causal effect of interest impossible. To address this issue, the manner in which exposures can interfere across zip codes can be simplified reasonably. For instance, we do not expect exposure in a zip code in central Missouri to affect outcomes in northern Georgia, unless for some reason there is common travel between those two regions. To incorporate this factor, we can define a function $g(\cdot)$ along with the assumption discussed in the following section.[109,112]

5.2.1 Interference Sutva Assumption

For any $\boldsymbol{w}^{all}$ and $\boldsymbol{w}^{all'}$ such that $\boldsymbol{w}_i^{all} = \boldsymbol{w}_i^{all'}$ and $g(\boldsymbol{w}_{-i}^{all}) = g(\boldsymbol{w}_{-i}^{all'})$, we have that $Y_i(\boldsymbol{w}_i^{all}) = Y_i\left(\boldsymbol{w}_i^{all'}\right)$.

Note that $\boldsymbol{w}_{-i}^{all}$ is the exposure at all locations, except for zip code *i*. This assumption states that the outcome at zip code *i* depends on exposure in zip code *i*, and on exposures from other locations, but only through the value of $g(\boldsymbol{w}_{-i}^{all})$. This function, sometimes referred to as an exposure mapping, simplifies the interference structure significantly by assuming that a summary of the remaining $n - 1$ exposures is sufficient to explain their effect on zip code *i*. We can now define potential outcomes as $Y_i(\mathbf{w},\mathbf{g})$, which represents the outcome for zip code *i* that would have been observed if they were exposed to $\boldsymbol{w}$ in their home zip code, and their exposure from neighboring zip codes was set to $\boldsymbol{g}$.

5.2.2 Choice of g(·) and Cell Phone Mobility Data

A key component of the interference SUTVA assumption discussed above is the choice of $g(\cdot)$. In the air pollution setting where interference is caused by an individual's mobility into other zip codes, it is natural to let this function represent a weighted average of the exposures in other zip codes. Intuitively, we want to assign weight to other zip codes based on how much time is spent in those areas. Suppose that people in zip code A spend 60% of their time in zip code A, 30% in zip code B, and 10% in zip code C. We would define the neighboring exposure value for zip code A to be a weighted average of the exposures in zip codes B and C, with weights of 3/4 and 1/4, since they spend three-quarters of their time traveling in zip code B.

Now we can extend this idea to the situation in which we have *n* zip codes of interest, and we obtain mobility information from cell phone mobility data. First, let $\boldsymbol{T}$ be an $n \times n$ matrix where T_{ij} is the amount of time that cell phone users from zip code *i* spent in zip code *j*. Aggregated mobility data is provided by Cuebiq, a location intelligence platform. Data is collected from anonymized users who have opted in to provide access to their location data anonymously, through a CCPA and GDPR-compliant framework. Through its Social Impact program, Cuebiq provides mobility insights for academic research and humanitarian initiatives. The Cuebiq responsible data sharing framework enables research partners to query anonymized and privacy-enhanced data by providing access to an auditable, on-premise Data Cleanroom environment. All final outputs provided to partners are aggregated to preserve privacy.

Once we have estimates of mobility from cell phone data, we can define two key quantities that summarize both the amount and direction of mobility. We let the percentage of time individuals in zip code *i* spend in zip code *i* be given by $\tau_i = \frac{T_{ii}}{\sum_k T_{ik}}$. We also need weights that tell us how much time individuals in zip code *i* spend in every other zip code. For this, we use a_{ij} where we let $a_{ii} = 0$ and

$$\alpha_{ij} = \frac{T_{ij}}{\sum_{k \neq i} T_{ik}} \quad \text{for } i \neq j$$

We can now define our neighborhood-based exposure for zip code *i* as $\boldsymbol{G}_i = \Sigma_j a_{ij} \boldsymbol{W}_j$, where $\boldsymbol{W}_j$ is the value of the air pollution mixture for zip code *j*, which means that for every zip code in our study, we have two values of the exposure for every component of the air pollution mixture: a home zip code–based exposure and a neighborhood or mobility-based exposure. To highlight this point, **Figure 12** shows average $PM_{2.5}$ levels in Massachusetts in 2016. We can see from the home zip code–based exposures in the left panel that exposure is elevated near Boston in the east, as well as near Springfield in the southwestern part of the state. Once mobility is accounted for, we see that this elevated exposure is smoothed out across space, as people in other zip codes are exposed to those higher levels of pollution through their daily travel. We explore differences between home zip code and mobility-based exposures in more detail in Section 5.5.

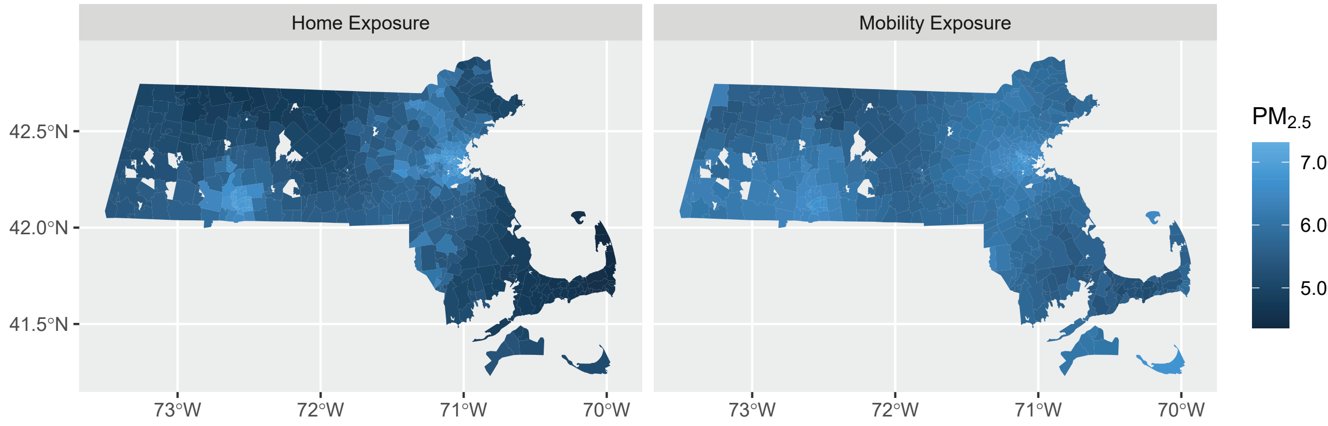


Figure 12. Home zip code and mobility-based exposure to $PM_{2.5}$ in Massachusetts in 2016.

5.2.3 Estimands That Incorporate Mobility

We can now define causal estimands of interest that represent meaningful and policy-relevant shifts in the air pollution mixture, while also acknowledging the presence of mobility. We focus on one estimand in this report that we find particularly relevant in this setting, but for other potential choices, see our work in Shin and colleagues (2023).[113] For this estimand, we first consider a shift in the air pollution mixture across all n zip codes given by $\Delta = [\Delta_1', \ldots, \Delta_n']$, where Δ_i is the change in air pollution levels for zip code i. This shift in exposure levels will shift both home zip code and mobility exposures for zip code i by Δ_{wi} and Δ_{gi}, respectively. Our main effect of interest is then given by

$$\omega(\Delta) = \frac{1}{n}\sum_{i=1}^{n}\left\{Y_i\left(\boldsymbol{W}_i + \Delta_{wi}, \boldsymbol{G}_i + \Delta_{gi}\right) - Y_i\left(\boldsymbol{W}_i, \boldsymbol{G}_i\right)\right\}$$

This equation represents the average effect of the shift in exposures given by Δ. In the interference literature, these types of effects are commonly decomposed into two parts: one representing the shift in **W**, and the other representing the shift in **G**. This approach is particularly useful in other scientific fields, such as vaccine studies, for which it is important to understand how much benefit an individual receives because other people take a vaccine. In the air pollution setting, however, we feel it is more relevant for policy purposes to focus on the overall change caused by the shift in air pollution levels, and it is less important whether this change is initiated from home zip code or mobility-based exposure shifts. As in previous chapters, we must make certain assumptions to estimate this quantity from observational data. We have already discussed the interference SUTVA assumption, which is the extension of SUTVA from the prior chapters. The positivity assumption now is with respect to both $\boldsymbol{W}$ and $\mathbf{G}$ instead of simply $\mathbf{W}$, meaning that there is a positive probability of observing all relevant values of $\boldsymbol{W}$ and $\mathbf{G}$, given the covariates. Lastly, we need to extend the no unmeasured confounding assumption to this setting, which states that there are no unmeasured confounders that affect both exposures ($\boldsymbol{W}$ and $\mathbf{G}$) and outcomes. Formally, this assumption is written as

$$\mathrm{Y}(\mathbf{w}, \mathbf{g}) \perp \boldsymbol{W}, \boldsymbol{G} \mid \boldsymbol{X},$$

which says that the potential outcomes are independent of the exposures once we condition on covariates $\boldsymbol{X}$.

5.3 WHAT HAPPENS WHEN MOBILITY IS IGNORED

Given that most large-scale epidemiological studies of the health effects of air pollution ignore mobility, it is important to understand its implications and the degree to which it can bias estimates of health effects. Measurement error is a well-studied issue in air pollution epidemiology,[114] as exposure measurement error inevitably occurs for a multitude of reasons. Here, we focus solely on the effect of mobility and ignore other potential sources of measurement error. Additionally, we restrict attention to a simplified setting with linear models and a single exposure, rather than a multivariate air pollution mixture. This focus simplifies calculations and provides clearer intuition about the effect of mobility on health effect estimates, although we point readers to Shin and colleagues (2023) for more general bias formulas outside of this simple scenario. Empirically, we have seen that the key points found in this scenario hold in more complex scenarios with multiple exposures and nonlinear exposure-response functions.

To simplify calculations, we assume that the potential outcomes are defined by the following linear model, which is a function of the home and neighborhood exposures,

$$Y(w, g) = \tau w \beta_w + (1-\tau) g \beta_g + \epsilon,$$

where ε represents independent random noise that is centered at zero, and τ is the percentage of time spent in the home zip code. Without loss of generality, we impose that both the home zip code and mobility-based exposures have been centered and scaled to have mean zero and variance 1, and the correlation between them is given by ρ. The parameters β_w and β_g are the effects of the home zip code and mobility-based exposures, respectively. Our target parameter will be the effect of increasing pollution everywhere by one, which leads to the true causal effect of interest being given by $\tau\beta_w + (1-\tau)\beta_g$. We examine three different approaches to estimating this quantity:

1. (Naive approach) Ignoring mobility altogether and fitting a model defined by
$$E(Y \mid W = w) = \tilde{\beta}_0 + w\tilde{\beta}_w$$
2. (Measurement error correction) Incorporating mobility by constructing a new exposure $W^* = \tau W + (1-\tau)G$ and then fitting a model defined by
$$E(Y \mid W^* = w^*) = \widetilde{\beta_0^*} + w\widetilde{\beta_w^*}$$
3. (Interference approach) Incorporating mobility by separately including W and G into a regression model

The third approach is to fit the true regression model and, therefore, should provide unbiased estimates for the causal estimand of interest. The first approach is the standard one in the literature that ignores mobility, and one can show that

$$\tilde{\beta}_w = \tau\beta_w + \rho(1-\tau)\beta_g$$

This expression looks exactly like the true value, except for the presence of ρ in the second term, which shows that ignoring mobility will lead to biased estimates of the causal effect of interest, and the degree of this bias depends on the correlation between the home zip code and mobility-based exposures. If this correlation is very high and close to 1, then there is very little bias. This result also provides insights into whether this bias is toward the null of no exposure effect. We typically expect the signs of β_w and β_g to be the same, and for ρ to be positive, which guarantees that this bias is toward the null and suggests that ignoring mobility may affect statistical power but will at least be conservative with respect to stating whether there is an effect of air pollution on a particular health outcome.

Now we can examine the measurement error correction approach, which is an alternative way one might incorporate mobility into an analysis. A natural approach is to find a “gold standard” exposure — which in this case is a weighted average of the home and neighborhood exposures — and simply use that exposure in all subsequent analyses. Doing so gives the following result:

$$\widetilde{\beta_w^*} = \frac{\tau^2\beta_w + \rho\tau(1-\tau)\beta_w + \rho\tau(1-\tau)\beta_g + (1-\tau)^2\beta_g}{\tau^2 + 2\tau(1-\tau)\rho + (1-\tau)^2}$$

This expression is much less intuitive, but there are a number of key points stemming from this result. One can show that this quantity is equal to the true causal effect if $\rho = 1$, $\tau \in \{0, 0.5, 1\}$, or $\beta_w = \beta_g$, which shows that it provides unbiased results in more scenarios than the naive approach of ignoring mobility. Moreover, we have observed that the resulting bias is generally much smaller than that of the naive approach. The first two conditions for unbiasedness are unlikely to hold in practice; however, the condition that $\beta_w = \beta_g$ merits further consideration. This condition states that the home zip code and mobility-based exposures have the same effect on the outcome, which seems plausible, if not likely, as these two quantities represent the same exposures, just obtained from different locations. There are situations, however, in which these coefficients might be expected to differ. If the exposure being examined is $PM_{2.5}$, then it is possible that the components comprising this pollution are different in home zip codes or neighboring locations in general. It is also possible that people spend more time outside while traveling than they do at their home zip code, and therefore, these coefficients would differ due to differences in indoor and outdoor pollution exposure. For this reason, we advocate for the interference approach developed here, which provides unbiased results in general. It is worth noting that the measurement error correction likely leads to similar results in most realistic situations. Either approach, however, requires mobility data and simply represents a different way of incorporating it into analysis.

5.4 ROBUSTNESS TO ERRORS IN MOBILITY ESTIMATES

Throughout this chapter, we rely on obtaining estimates of mobility between pairs of zip codes. In this section, we study the extent to which errors in estimating mobility patterns can have on resulting estimates of the causal effect of air pollution on health outcomes. There are two reasons why we might expect errors in the estimation of mobility patterns. First, because the mobility data contain a random subset of individuals in each zip code, our estimates of mobility might deviate from the truth simply due to sampling variability. A second, and potentially more problematic, reason is that the population represented by the mobility data may not reflect the population for which we are estimating health effects. Our mobility data comes from 15 million anonymous cell phone users throughout the contiguous United States in the year 2019, although specific demographic information is not collected on these users. Recent studies using the same source of mobility data have attempted to infer demographic characteristics of the population of cell phone users by linking cell phone data to publicly available census tract data on such characteristics. They have found that the users within the cell phone mobility data are largely representative of the US population as a whole, as indicated by high correlations between their data and census or county-level data.[115,116] Our health effects study, however, is of the Medicare population, which almost entirely comprises individuals over the age of 65 who likely have different mobility patterns. We focus in this section on these two realistic scenarios and how they would influence the estimation of causal effects. We center our results on errors in τ_i, the proportion of time spent in the home zip code, as it is easier from which to derive results. We then discuss potential implications of these findings for errors in estimating the weights a_{ij}, which represent the direction of mobility.

Throughout this section, we can let the true mobility values be given by τ_i^*, and the ones obtained from our cell phone mobility data be given by τ_i. We explore two scenarios for the misspecification of these values:

1. A constant shift in these values, given by

$$\tau_i = \frac{\tau_i^*}{c} \text{ for all } i \text{ and } c > 0$$

2. Random measurement error, given by

$$\tau_i = \tau_i^* + \eta_i, \text{ where } E(\eta_i) = 0 \text{ and } \operatorname{Var}(\eta_i) = \sigma_\eta^2$$

These two scenarios capture the two realistic scenarios described above. It is possible that the Medicare population systematically travels less than the general US population, which is captured by the first of these two scenarios. The second of these two scenarios depicts situations in which our mobility data provides noisy estimates of the true mobility, which can occur due to finite sample variability. Here we focus on the same linear model used in Section 5.3, and we assume the estimand of interest is $\omega(\Delta)$ with $\Delta = (1,1,\ldots,1)$. For simplicity, we denote the estimand by ω and our estimate that uses the incorrect weights τ_i as $\widehat{\omega}$. We show in Shin and colleagues (2023)[113] that the asymptotic bias under the constant shift error in scenario 1 is given by

$$\lim_{n\to\infty} E\left(\widehat{\omega} - \omega\right) = \frac{\beta_g(1-c)(1-\rho)\operatorname{Var}(\tau)\left[\rho E(\tau) - (1+\rho)E(\tau^2)\right]}{\left(E(\tau^2)E\left[(1-\tau)^2\right] - \rho^2 E^2\left[\tau(1-\tau)\right]\right)}$$

This equation shows that we get unbiased results if there is no shift (c = 1) or if there is a perfect correlation between home zip code and mobility-based exposures ($\rho = 1$). Of more interest is what happens when there is a shift, and the exposures are not perfectly correlated. We have seen empirically that this bias is very small and effectively negligible in most scenarios. This bias depends on the distribution of τ, and we have explored a wide range of distributions for this variable. We found that as long as $0.5 < c < 1.5$ and $\rho > 0.4$, both of which are reasonable to assume in practice, then this bias is at most $0.05\,\beta_g$. Another important thing to note about this bias formula is that all components of the bias, except for β_g and c, are estimable once we have mobility data and do not depend on unknown quantities, so we can empirically assess how bad the bias could be under potential shifts in τ.

Deriving results in scenario 2 under measurement error is a bit more challenging analytically. To make the calculations more tractable, we can assume that the home zip code and mobility-based exposures are independent of $\rho = 0$. While this situation is not a reasonable one in practice, we have seen that it leads to larger bias than when $\rho > 0$ in general, and so it can be thought of as a worst-case scenario that we can study to assess the effect of this form of error in τ. In this situation, we showed that

$$\lim_{n\to\infty} E\left(\hat{\omega}-\omega\right) = \beta_w\left\{\xi_w E\left(\tau^*\right) - E\left(\tau^*\right)\right\} + \beta_g\left\{\xi_g E\left(1-\tau^*\right) - E\left(1-\tau^*\right)\right\}$$

for constants ξ_w and ξ_g that are necessarily between 0 and 1. Importantly, when there is no measurement error and $\sigma_\eta^2 = 0$, both of these constants are 1, and we obtain unbiased results. Because these constants are between 0 and 1, we are guaranteed that measurement error of this type will lead to conservative estimates of the health effects of air pollution as long as β_w and β_g have the same sign, which is expected in practice. We have also explored different magnitudes of measurement error and found that even under rather large amounts of measurement error, these constants are greater than 0.9, which leads to small amounts of bias.

Overall, these findings suggest that the results will be relatively robust to incorrect estimation of the mobility weights, given by τ_i, which justifies our use of the cell phone mobility data as a proxy for mobility in the Medicare cohort. However, whenever possible, these two populations should be chosen to be as close to each other as possible. While we focused on τ_i, another place errors can occur is in the weights a_{ij} that tells us the proportion of time that people in zip code i spend in zip code j. It is more difficult to derive results with these weights, as there are n of them per zip code, and they are constrained to sum to 1. If there was a random error in a_{ij}, such as in scenario 2 as described above, we expect similar results where bias would be relatively small and toward the null. A more problematic situation would be if there were systematic errors in the direction of mobility, and these errors were correlated with pollution level, as bias could go in either direction in this scenario. It is important to remember, however, that ignoring mobility altogether is a special case of errors in estimating τ_i and a_{ij}, and it would take exceedingly bad estimates of these quantities from mobility data to do worse than ignoring mobility altogether.

5.5 DATA ANALYSIS

In this section, we apply the proposed methodology that incorporates mobility to further study the effects of air pollution mixtures on mortality in the Medicare cohort. We run several analyses solely examining $PM_{2.5}$ as the exposure of interest, and additionally run an analysis that examines the health effects of an air pollution mixture containing BC, OM, NH_4, SO_4, and NIT. We again separate analyses by year in this study, which avoids issues caused by correlated data points across time within a zip code and allows the effects to change over time. We discuss specifics of the statistical models used in what follows, but the ones presented are much faster computationally than those expounded on in Chapter 4 and could be applied to very large datasets. We also adjust for the same set of confounders $\boldsymbol{X}$ that we included in the analysis of effect modification examined in Section 4.4. As discussed in Section 5.2, cell phone mobility data are used to estimate $\boldsymbol{T}$ and are drawn from more than 15 million cell phone users in the contiguous United States over all days in 2019. Note that we do not have available mobility data for the years of our study (2000 to 2016) and, therefore, use mobility data from 2019 as an estimate of the mobility patterns across zip codes in earlier years.

5.5.1 STATISTICAL MODELING

Up to this point, we have only discussed conceptual issues with including mobility into analyses of the health effects of air pollution, but now we briefly discuss the statistical models used to estimate these quantities. All quantities of interest can be estimated once we obtain estimates of $E(Y \mid \boldsymbol{W},\boldsymbol{G},\boldsymbol{X})$, which we specify as

$$E(Y \mid \boldsymbol{W}=\boldsymbol{w}, \boldsymbol{G}=\boldsymbol{g}, \boldsymbol{X}=\boldsymbol{x}) = \tau f\left(\boldsymbol{w}\right) + \left(1-\tau\right) h\left(\boldsymbol{g}\right) + l\left(\boldsymbol{x}\right),$$

where $f(\cdot)$, $h(\cdot)$, and $l(\cdot)$ are arbitrary, unknown functions that need to be estimated. Note that here $f(\boldsymbol{w})$ captures the effect of the home zip code–based exposures and $h(\boldsymbol{g})$ captures the effect of mobility-based exposures, and these are weighted by the time spent in the home zip code, τ_i. We first assume that $l(\boldsymbol{x}) = \boldsymbol{x}\theta$, which means we adjust for the confounders linearly. We do not want to assume the effects of the exposures are linear, so we use a generalized additive model for the effects of home zip code and mobility-based exposures. Specifically, for the home zip code–based exposures, we specify

$$f\left(\boldsymbol{w}\right) = \sum_{j=1}^{q} f_j\left(w_j\right) = \sum_{j=1}^{q}\sum_{m=1}^{M}\beta_{jm}\phi_m\left(w_j\right),$$

where $\phi_m(\cdot)$ are orthogonal polynomial basis functions, and M is the number of basis functions used. Throughout, we set $M = 3$ so that we use orthogonal cubic polynomials. We use these same basis functions for the mobility-based exposures so that

$$h\left(\boldsymbol{g}\right) = \sum_{j=1}^{q}\sum_{m=1}^{M}\left(\beta_{jm} + \zeta_{jm}\right)\phi_m\left(g_j\right).$$

Note that the basic functions are the same, but now the coefficient for each function is given by $\beta_{jm} + \zeta_{jm}$. We see that β_{jm} is shared by both the $f(\cdot)$ and $h(\cdot)$ functions, but the ζ_{jm} are unique to the mobility-based exposures. These coefficients capture differences between these two functions. We parameterize the problem in this way so that we can use shrinkage prior distributions to bring these two functions toward each other, as we believe a priori that they are not likely to be very different from each other. Specifically, we place horseshoe prior distributions[117] on ζ_{jm}, which shrinks them toward zero, unless the data strongly suggest that $f(\cdot)$ and $h(\cdot)$ are different. We have shown in Shin and colleagues (2023)[113] that this can lead to big improvements in the efficiency of estimates. We use standard, noninformative prior distributions for all other parameters in the model.

5.5.2 Assessing Mobility and Its Relationship to Air Pollution

Before estimating the effects of air pollution in the Medicare cohort, we can examine distributions of both home zip code and mobility-based exposures to assess how large a role mobility can play in our analysis. We can first explore the distribution of τ_i across the country in **Figure 13**, which shows the distribution of τ_i across zip codes, separated by state. This quantity captures the percentage of time spent in the home zip code, and we see that most states have fairly similar distributions of this quantity that are centered around 0.78, with most zip codes falling between 0.7 and 0.85. The area with the smallest τ_i values are in Washington, DC, while the largest τ_i values are found in Montana.

This observation shows that people do spend reasonably large amounts of time outside of their home zip code; we can therefore examine whether they are subsequently exposed to systematically different pollution levels in their travels. **Figure 14** shows the correlation between home zip code and mobility-based exposures for each pollutant across time. The correlation is fairly constant across time, but there are big differences across pollutants. Not surprisingly, sulfates have a very high correlation as they vary slowly across space, and these values are not expected to change much across nearby zip codes. Regardless of the exposure examined, correlations are fairly high between these two exposures. To examine these correlations more closely, **Figure 15** shows a scatter plot of these two exposures for the year 2016, separated by pollutant. A general trend across all pollutants is that people in zip codes with low levels of pollution tend to travel to areas of higher pollution, and those who live in higher pollution areas tend to travel to lower pollution areas. There is also less variability in the mobility-based exposures compared with the home zip code–based exposures, which was also evident in Figure 12, in which mobility-based exposures were smoothed across space. Lastly, **Table 7** shows average exposure levels for both home zip code and mobility-based exposures, and we see that exposures are slightly higher during travel than at the home zip code for all pollutants. This finding is potentially caused by daily mobility for work or other commercial activities from suburban areas with lower pollution into urban areas with higher pollution.

Overall, this scenario paints a fairly clear picture of mobility-based exposures and their potential impact on subsequent health effect estimates. Typically, people travel into areas of higher levels of pollution, but this mobility-based exposure is quite highly correlated with home zip code–based exposure in the United States at the zip code level. This observation, coupled with our earlier results, suggests that mobility will not drastically affect the resulting estimates on the causal effects of air pollution mixtures in our study.

5.5.3 Analysis of $PM_{2.5}$ Only

We now estimate the health effects of air pollution on mortality using the proposed approaches that incorporate mobility. We use the statistical model described above, and run the MCMC algorithm for 10,000 iterations, removing the first 2,000 scans, and thinning every eighth sample. Throughout, we also fit the same statistical models without accounting for mobility, which amounts to discarding $\boldsymbol{G}$ and only using $\boldsymbol{W}$ to estimate the effect of pollution. We first

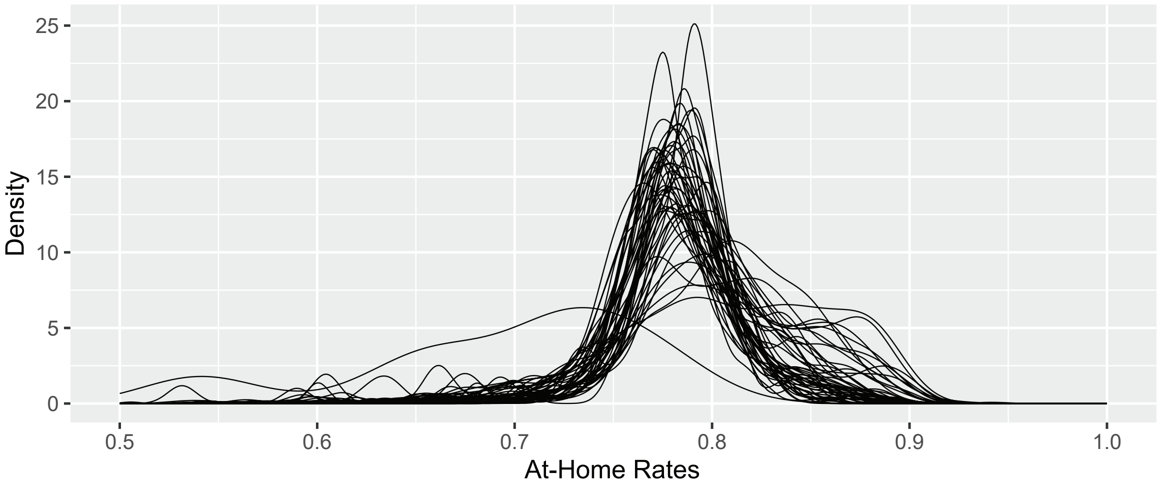


Figure 13. Density plots of the proportion of time people stay in their home zip code (τ_i) in 2019, across geographic regions. Each line corresponds to a specific state or Washington, DC.

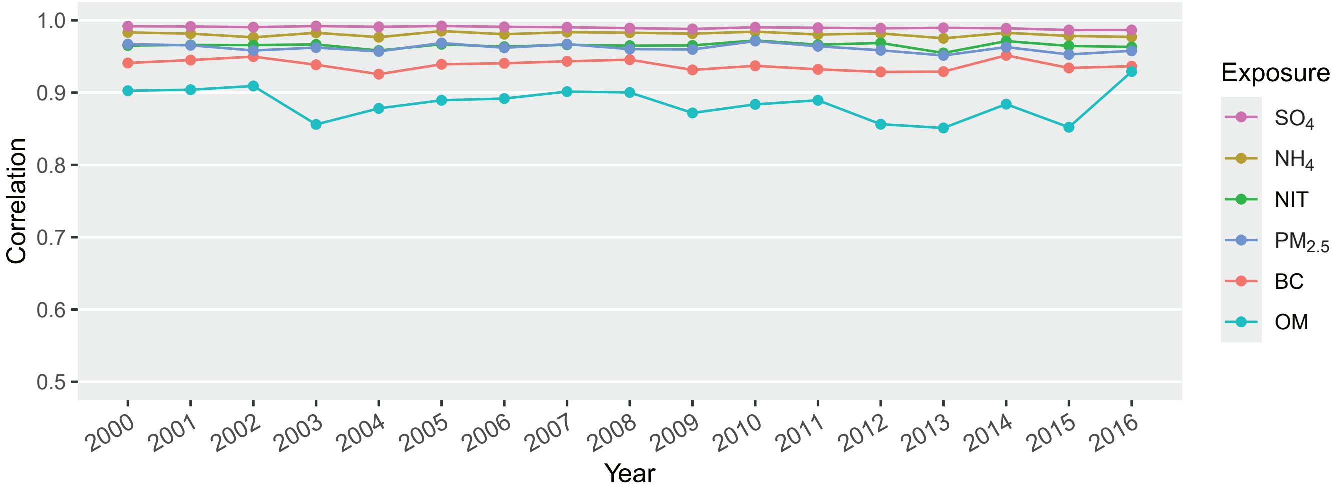


Figure 14. Correlation between home zip code and mobility-based exposures for each pollutant and year combination.

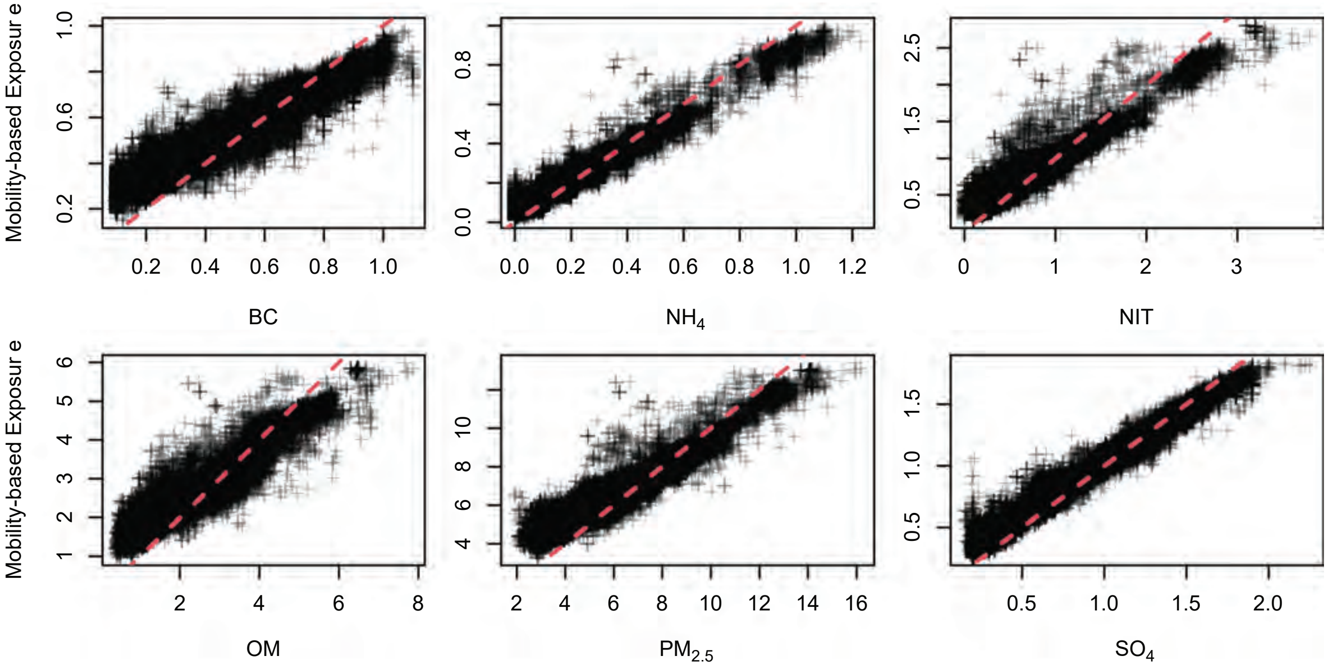


Figure 15. Scatter plot of home zip code and mobility-based exposures in the year 2016 for each pollutant.

Table 7. Exposure Levels Averaged Across All Zip Codes and Years, Separated by Pollutant and Whether Exposure is Home Zip Code or Mobility Based

	BC	NH_4	NIT	OM	$PM_{2.5}$	SO_4
Home	0.71	0.91	1.09	3.08	9.34	2.27
Mobility	0.75	0.93	1.15	3.32	9.64	2.31

explore a univariate exposure setting for which we focus on $PM_{2.5}$ and then extend results to the full air pollution mixture. We run analyses separately by year, and in each analysis the target parameter is the expected change in mortality rates under a 0.25 standard deviation increase in $PM_{2.5}$. The results can be found in the left panel of **Figure 16**, which shows a consistent, detrimental effect of air pollution on mortality across years. The results are quite similar whether mobility

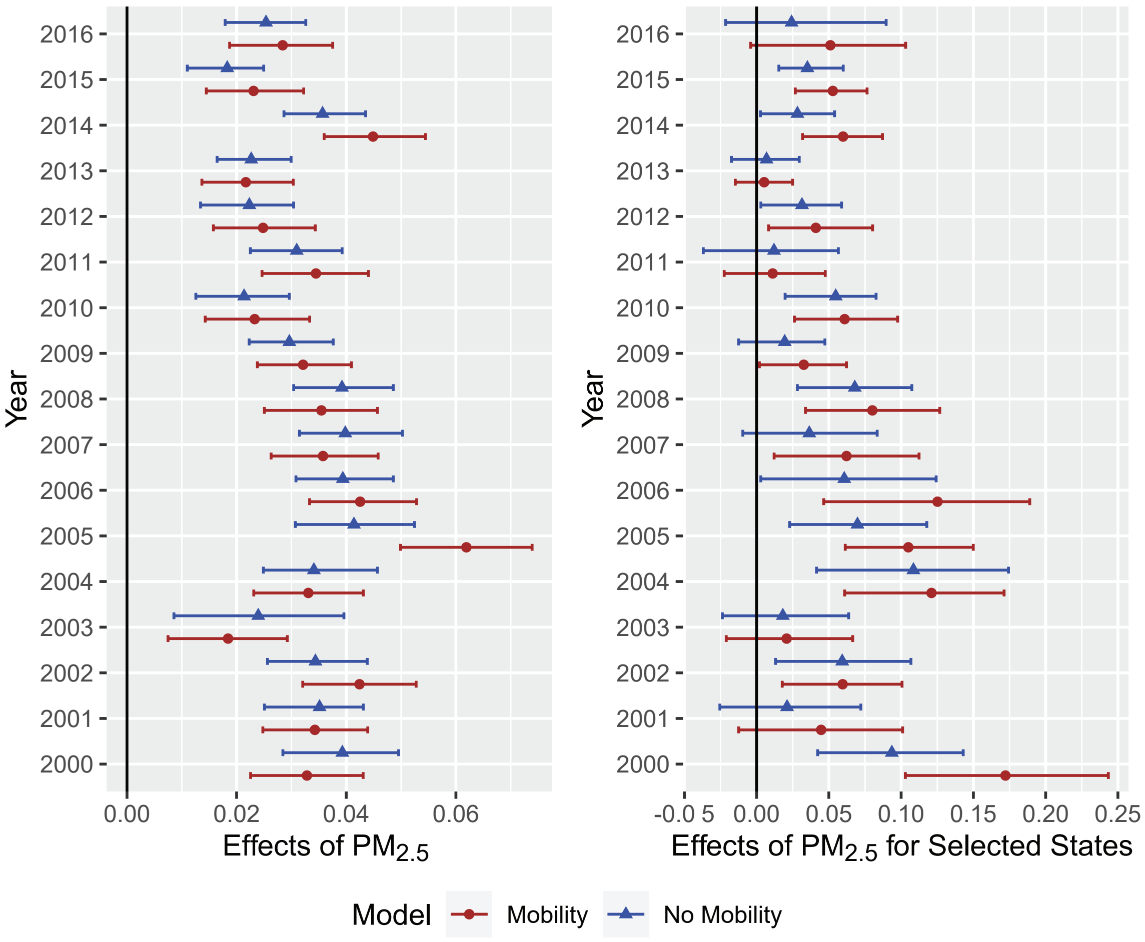


Figure 16. Point estimates and 95% credible intervals for estimating the effect of $PM_{2.5}$ on mortality rates when all states are included (left) and when only states with the lowest correlations between home and mobility-based exposures (right) are included in the analysis.

is accounted for or not, which is expected due to the high correlation between home and mobility-based $PM_{2.5}$. Given that mobility is most important to account for when the correlation between $\boldsymbol{W}$ and $\boldsymbol{G}$ is lower, we also ran the same analysis, but instead restricted our study area to all zip codes in states that had an average correlation between $\boldsymbol{W}$ and $\boldsymbol{G}$ that was less than 0.9 for $PM_{2.5}$. This removes a large portion of our study sample as it excludes the six largest states, and only 22 states are left for analysis, containing approximately 5,000 zip codes. The results can be found in the right panel of Figure 16. We again see positive estimates, although there is more uncertainty about these estimates and there are fewer that achieve statistical significance due to the reduced sample size. Interestingly, however, there are far greater differences between the analyses that adjust for mobility and those that do not. Specifically, analyses incorporating mobility have 41% larger estimates averaged over all study years. Our bias formulas suggested that ignoring mobility can bias results and that this bias would generally be toward the null, which is shown empirically here as mobility-based estimates are larger on average. This finding shows that in certain settings, accounting for mobility can provide important and non-negligible differences in the estimates of the health effects of air pollution.

5.5.4 Analysis of Air Pollution Mixtures

In this section, we extend our analysis to include the full air pollution mixture comprising six different pollutants. We use the same statistical models described previously, and our goal is to estimate the effect of increasing all six exposures simultaneously by 0.25 times their respective standard deviations on mortality rates. **Figure 17** shows the effect estimates separately by year, both with and without accounting for mobility in the analysis. We see relatively similar results to those found for the univariate $PM_{2.5}$ analysis, which is that there is a consistent positive effect estimate showing that increasing air pollution levels increases mortality. This analysis utilized zip codes from all states, where the correlation is high between home zip code and mobility-based exposures,

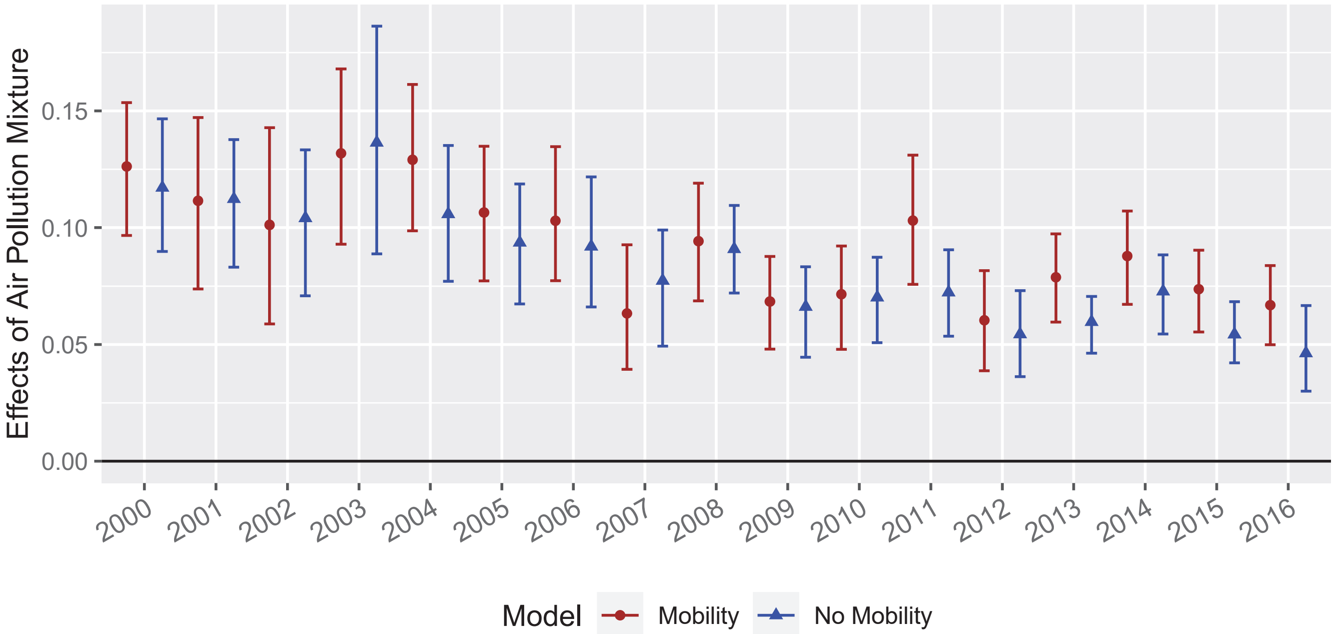


Figure 17. Point estimates and 95% credible intervals for the effect of the air pollution mixture on mortality.

which leads to similar effect estimates whether mobility is accounted for or not. While the differences are not as large as when we focused only on states with reduced correlation between $\boldsymbol{W}$ and $\boldsymbol{G}$, in 14 of the 17 analyses, the effect estimates are larger for the models that incorporate mobility. This finding is suggestive of the same issue, which is that ignoring mobility can bias estimates toward the null of no exposure effect, although this bias is small due to the high correlation between home and mobility-based exposures.

5.6 DISCUSSION AND CONCLUSIONS

In this chapter, we explored the extent to which the mobility of individuals can influence estimates of the effect of air pollution mixtures on health outcomes. By deriving bias formulas in a variety of settings, we showed that mobility can bias effect estimates, but this bias depends on a variety of factors, some of which can be estimated from observed data. The most important quantity for the impact of mobility is the correlation between home exposures and exposure during daily mobility, as bias will generally be small if this correlation is very high. Another important finding is that while bias can go in either direction when ignoring mobility, it will generally be toward the null (i.e., no exposure effect) in realistic scenarios typical of air pollution epidemiology. This outcome is beneficial because it suggests that studies ignoring mobility — which is standard practice — are somewhat protected, as their results are likely to provide conservative estimates of the effects of air pollution, albeit with reduced statistical power. This result is also beneficial as it suggests that if researchers can account for mobility, then they will likely estimate more pronounced effects of pollution and should have more statistical power when such an effect exists. Our results also suggest that if there is no effect from the exposures, then ignoring mobility does not bias results, and therefore, we would not incorrectly conclude there is an effect of an exposure when in truth there is none. Lastly, while we have formulated everything within a causal inference framework to better elucidate the role of mobility, these results have implications beyond causal analyses of the health effects of air pollution. Our findings, such as the degree and direction of bias when ignoring mobility, would also hold for conditional associations and regression parameters typically estimated in air pollution epidemiology. This detail is important for studies such as ours that have certain difficulties precluding a causal interpretation of findings, such as the ecological nature of our study and the limited set of confounders measured. This final point about unmeasured confounding will be addressed in the following chapter, in which we aim to assess how robust our findings are to this particular issue.

While our findings about the effect of mobility are generally applicable beyond the current case study, there are certain limitations with respect to the analysis performed in the Medicare cohort. As discussed previously, the population we are performing inference on is different from the one from which we obtained cell phone mobility data. While we derived promising results in Section 5.4 showing robustness of our findings to the cell phone mobility population, it is certainly preferable if these two populations were more closely aligned. For one, our results were only focused on the bias of the estimated effects when mobility is incorrectly estimated, but an equally important issue is the variability in those estimates. We have seen empirically in simulations that while bias is typically very low due to errors in mobility estimates, the uncertainty in the estimates can be higher due to incorrect assessment of mobility. Additionally, our results were only focused on errors in estimating the amount of mobility τ_i, but there could be systematic errors in estimating the direction of mobility, given by a_{ij}. If these are fundamentally different in the mobility population and the population where health effects are being estimated, it could bias results in unexpected ways. This is a more difficult problem to address mathematically and is a topic for future research. Despite these limitations, we still believe it is useful to account for mobility as it is likely to align more closely with reality than ignoring mobility altogether.

CHAPTER 6: ADDRESSING UNMEASURED CONFOUNDING WITH MULTIPLE EXPOSURES AND MULTIPLE OUTCOMES

6.1 INTRODUCTION

Confounding bias is an ever-present problem affecting the validity of observational studies of the health effects of air pollution. In certain settings, natural experiments arise that allow researchers to estimate the effects of pollution under weaker assumptions about confounding, such as those caused by reductions in pollution leading up to the Olympic Games.[118] In most cases, particularly in large-scale national studies of the health effects of pollution, such natural experiments are not available, and we must adjust for all potential confounders of the exposure–outcome association. Clearly, this assumption is strong, and one that should be assessed whenever possible. One approach to doing so is through sensitivity analysis, which aims to assess how robust a finding is to the potential presence of omitted variable bias. There are many different approaches to sensitivity analysis in the literature,[119–131] and they all have a similar underlying idea, which is to evaluate how strongly the unmeasured confounder must be associated with either the exposure or the outcome to fully explain any estimated effect. Once this approach is identified, then subject matter experts can reason about whether that amount of unmeasured confounding bias is plausible or unlikely.

While this approach is useful and certainly better than simply assuming that unmeasured confounders are not present, it is not always informative. It can be very hard for subject matter experts to reason about the potential strength of unmeasured confounding, and subsequently, whether the estimated effects are robust to unmeasured confounding bias or not. Recent work, however, has shown that certain benefits can be obtained if one has multiple exposures or multiple outcomes in their analysis.[132–138] We focus here on the benefits that multiple exposures or outcomes provide in terms of addressing unmeasured confounding bias, although there are potentially many other benefits to multiple-outcome analyses.[139] In these settings, certain parameters that govern the strength of unmeasured confounding are actually estimable from the observed data and no longer need to be reasoned about with subject matter experts. This methodology allows one to estimate bounds or partial identification regions, which are regions the causal effect must be contained within, even in the presence of unmeasured confounding. We extend our analysis to settings involving both multiple exposures and multiple outcomes, with a particular emphasis on estimating the causal effects of air pollution mixtures. Throughout this chapter, our goal is to derive partial identification regions for the causal effect of air pollution mixtures when there are unknown and unmeasured confounders present, but we are willing to make an assumption that these confounders affect multiple exposures or multiple outcomes simultaneously. This is a significantly weaker assumption than assuming that there are no unmeasured confounders, as we did in previous chapters, and can lead to useful insights about the presence, or lack thereof, of a causal effect of air pollution mixtures. We provide intuition — clear, conceptual explanations — for how the observed data can be informative in the presence of unmeasured variables and discuss additional strategies that can help reduce the widths of partial identification regions, thereby increasing the power to detect a causal effect of air pollution mixtures.

6.2 ESTIMANDS AND CORRESPONDING ASSUMPTIONS

As discussed in earlier chapters, we observe q exposures denoted by $\boldsymbol{W}_i$, and p *observed* covariates given by $\boldsymbol{X}_i$ for all $i = 1, \ldots, n$. For this chapter, we assume that we observe k distinct outcomes, which are denoted by $\boldsymbol{Y}_i$. We also assume that there exist m variables $\boldsymbol{U}_i$, which are fully unobserved, and therefore can bias the causal effect estimates we are interested in. We make the same SUTVA assumption as in Chapter 4, but the other two core assumptions we have used throughout must be modified in this setting due to the presence of the unmeasured variables $\boldsymbol{U}$. Specifically, our assumptions are given by the following:

i. *SUTVA:* Exposure for one zip code does not affect potential outcomes for other zip codes, and exposure is well defined in the sense that $\boldsymbol{Y}_i = \boldsymbol{Y}_i(\boldsymbol{W}_i)$.

ii. *Latent positivity:* $0 < f_{W \mid X,U}(\boldsymbol{w} \mid \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{U} = \boldsymbol{u})$ for all $\boldsymbol{x}$, $\boldsymbol{u}$, and $\boldsymbol{w}$ where $f_{W \mid X,U}$ is the conditional density of the exposures given the observed and unobserved covariates.

iii. *Latent ignorability:* $Y(\boldsymbol{w}) \perp \boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{U}$ for all $\boldsymbol{w}$.

Note that now the positivity and no unmeasured confounding assumption (now referred to as latent ignorability) are assumed conditional on both observed covariates ***X*** and unobserved covariates ***U***. Our focus here will be on the latent ignorability assumption as we aim to assess how robust our findings are to the presence of unmeasured variables ***U***. Importantly, we are *not* assuming that the observed variables in ***X*** are sufficient to adjust for confounding and allow for unobserved variables ***U*** to be confounders of the exposure–outcome relationship.

Our estimand of interest in this setting is a multivariate extension of the average treatment effect (ATE) seen in Chapter 4. This estimand is defined as

$$\text{ATE}_{a,w_1,w_2} = E\left[\boldsymbol{a}^T\boldsymbol{Y}\left(\boldsymbol{w}_1\right) - \boldsymbol{a}^T\boldsymbol{Y}\left(\boldsymbol{w}_2\right)\right],$$

where $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are two exposure levels of interest, and $\boldsymbol{a}$ is a vector of weights that we assign to each outcome. One example is if our interest is in estimating the effect of the exposures on only the first outcome, in which case we would set $\boldsymbol{a} = (1,0,\ldots,0)$. Under the assumptions listed above, one can identify the causal effect according to the following expression:

$$E\left[\boldsymbol{Y}\left(\boldsymbol{w}_1\right)\right] = E_{X,U}\left[E(\boldsymbol{Y} \mid \boldsymbol{W} = \boldsymbol{w}, \boldsymbol{X}, \boldsymbol{U})\right].$$

This expression is not useful to us, however, as we do not get to observe ***U*** and therefore we cannot estimate $E(\boldsymbol{Y} \mid \boldsymbol{W} = \boldsymbol{w}, \boldsymbol{X}, \boldsymbol{U})$, nor can we calculate the outer expectation as it relies on the joint distribution of ***X*** and ***U***, which is not observable. In practice, all we can hope to do is calculate the analogous quantity that uses only the observed covariates, given by

$$E_X\left[E(\boldsymbol{Y} \mid \boldsymbol{W} = \boldsymbol{w}, \boldsymbol{X})\right],$$

which is not necessarily equal to the estimand of interest, because of the presence of unmeasured confounders. In this chapter, we focus on the bias that occurs by ignoring ***U*** in this manner, and whether we can bound the causal effect using only the observed data.

6.3 PARTIAL IDENTIFICATION REGIONS WITH UNMEASURED CONFOUNDING BIAS

In this section, we show that the bias from unmeasured variables ***U*** is not identified from the observed data. However, in certain scenarios, we can identify a region of plausible values for this bias that, in turn, yields a corresponding region of values that the causal effect is contained within, which we refer to as a partial identification region. This region should not be confused with more standard intervals, such as a confidence interval. This region shows the set of values that are equally supported by the data, and without additional assumptions, we cannot hope to shrink this region further, even with an infinitely large sample size. Our results rely on what we refer to as a factor confounding assumption. One aspect of this assumption is that the observed data are generated from the following model:

$$U = \epsilon_U \tag{2}$$

$$W = h\left(\boldsymbol{X}\right) + BU + \epsilon_W \tag{3}$$

$$Y = g\left(\boldsymbol{W}, \boldsymbol{X}\right) + \Gamma \Sigma_{u|w,x}^{-1/2} U + \epsilon_Y \tag{4}$$

Here we have that $\Sigma_{u \mid w,x} = cov(\boldsymbol{U} \mid \boldsymbol{W} = \boldsymbol{w}, \boldsymbol{X} = \boldsymbol{x})$, and all of the error terms given by $(\varepsilon_U, \varepsilon_W, \varepsilon_Y)$ are assumed to follow multivariate normal distributions with diagonal covariance matrices given by $\Sigma_u, \Sigma_{w \mid x,u}, \Sigma_{y \mid w,x,u}$. If we do not observe ***U*** and we estimate the causal effect by regressing the outcomes on the exposures and observed covariates only, then we obtain a biased estimate of the $g(\cdot)$ function, which leads to a biased estimate of the causal effect of interest. The important parameters dictating the size and direction of this bias are ***B*** and **Γ**, which are coefficients dictating the strength of association between the unmeasured variables ***U*** and the exposures and outcomes, respectively. If either $\boldsymbol{B} = \boldsymbol{0}$ or $\boldsymbol{\Gamma} = \boldsymbol{0}$, then there is no confounding by unmeasured variables ***U***, and we would obtain unbiased estimates of the causal effect. If both are nonzero, however, then bias may exist, and this is the scenario that we work in throughout this chapter. Under models (2)–(4), this bias is given by

$$Bias = \boldsymbol{a}^T \Gamma \Sigma_{\mathrm{u|w,x}}^{-\frac{1}{2}} \left[E(\boldsymbol{U} \mid \boldsymbol{W} = \boldsymbol{w}_1, \boldsymbol{X} = \boldsymbol{x}) - E(\boldsymbol{U} \mid \boldsymbol{W} = \boldsymbol{w}_2, \boldsymbol{X} = \boldsymbol{x})\right].$$

One can even write down exactly what $\Sigma_{u \mid w,x}$ and $E(\boldsymbol{U} \mid \boldsymbol{W} = \boldsymbol{w}, \boldsymbol{X} = \boldsymbol{x})$ are, but they are not very simple expressions, nor do they allow for intuition, so we leave technical details to our work in Kang and colleagues (2023).[140] The important thing to note is that both are functions of ***B***, and they are such that if either $\boldsymbol{B} = \boldsymbol{0}$ or $\boldsymbol{\Gamma} = \boldsymbol{0}$, then the bias becomes zero because there is no unmeasured confounding. If we knew both ***B*** and **Γ**, we could use this expression to correct for the bias, but this approach clearly is not feasible because it would require observing ***U***, which is unavailable to us. We can, however, provide an upper bound on this bias that we can obtain from the observed data, even without getting to observe ***U***. Specifically, the squared value of the bias is bounded above by

$$\text{Bias}^2 \leq \left\|\boldsymbol{a}^T \boldsymbol{\Gamma}\right\|_2^2 \left\|\Sigma_{u|w,x}^{-\frac{1}{2}} \left[E(\boldsymbol{U} \mid \boldsymbol{W} = \boldsymbol{w}_1, \boldsymbol{X} = \boldsymbol{x}) - E(\boldsymbol{U} \mid \boldsymbol{W} = \boldsymbol{w}_2, \boldsymbol{X} = \boldsymbol{x})\right]\right\|_2^2. \tag{5}$$

Note here that the notation $\left\|\cdot\right\|_2^2$ represents the sum of the squared elements inside of quantity inside of the vertical brackets and is therefore one measure of the magnitude of its entries. In this way, we can see that the bias is bounded above by a quantity that depends on both the size of the association between the unmeasured variables and the exposures, and by the size of the association between the unmeasured variables and the outcomes. Despite this upper bound, it is not immediately obvious how it is helpful, but what we show in subsequent sections is that each of the two components of the upper bound can be estimated, even without data on ***U***.

6.3.1 Benefits of Multiple Treatments and Multiple Outcomes

The upper bound on the bias depends on the unknown coefficients ***B*** and **Γ**, which are generally not identifiable without information on ***U***. We showed in Kang and colleagues

(2023)[140] that if models (2)–(4) hold, along with additional assumptions on the true values of $\boldsymbol{B}$ and $\boldsymbol{\Gamma}$, then the upper bound in equation (5) is estimable from the observed data when we have multiple exposures and multiple outcomes. To build intuition for this result, it is easiest to focus on $\boldsymbol{\Gamma}$, which reflects the strength of dependence between the unmeasured confounders and the outcome. Using the observed data, we can estimate $\mathrm{Cov}(\boldsymbol{Y} \mid \boldsymbol{W}, \boldsymbol{X})$ as it only depends on observed variables and does not require $\boldsymbol{U}$. This quantity has the following expression:

$$\mathrm{Cov}(\boldsymbol{Y} \mid \boldsymbol{W}, \boldsymbol{X}) = \Gamma\Gamma^{\mathrm{T}} + \Sigma_{y|w,x,u}$$

Therefore, we can use standard factor analysis techniques to obtain an estimate of $\boldsymbol{\Gamma\Gamma}^{\mathrm{T}}$. Effectively, this means that the observed data cannot estimate the exact value of $\boldsymbol{\Gamma}$, but it can be used to estimate $\boldsymbol{\Gamma\Gamma}^{\mathrm{T}}$, which provides some information about the potential for unmeasured confounding bias. Specifically, the first term in the upper bound in equation (5) can be written as

$$\left\|\boldsymbol{a}^T\Gamma\right\|_2^2 = \boldsymbol{a}^T\Gamma\Gamma^{\mathbf{T}}\boldsymbol{a}.$$

Because $\boldsymbol{a}$ is known, and we can estimate $\boldsymbol{\Gamma\Gamma}^{\mathbf{T}}$ as described above, this means we can estimate this term even though we do not observe $\boldsymbol{U}$. We will not go through the details of the second term of equation (5), but the same idea applies. We are not able to estimate $\boldsymbol{B}$, but we can estimate $\boldsymbol{BB}^T$, which is sufficient for estimating the upper bound on the bias. Essentially, we are still not able to pinpoint $\boldsymbol{\Gamma}$ and $\boldsymbol{B}$ exactly, which is why we cannot pinpoint the causal effect exactly. This does, however, limit the range of possible values for $\boldsymbol{\Gamma}$ and $\boldsymbol{B}$, which in turn puts constraints on what the true causal effect can be.

As mentioned above, these results rely on certain additional assumptions holding. To estimate $\boldsymbol{\Gamma\Gamma}^{\mathrm{T}}$, we need for each unmeasured confounder to affect at least three outcomes and to have enough outcomes measured so that $(\mathrm{q} - \mathrm{m})^2 - q - m \geq 0$, where q is the number of outcomes and m is the number of unmeasured confounders. An analogous assumption needs to be made to estimate $\boldsymbol{BB}^T$. Similar assumptions are sometimes referred to as a shared confounding structure or a no-single-cause confounding assumption[141]. While these are assumptions in their own right, they are significantly weaker than assuming no unmeasured confounding, which would require assuming that either $\boldsymbol{B} = \mathbf{0}$ or $\boldsymbol{\Gamma} = \mathbf{0}$.

Overall, these results show a tremendous benefit of having multiple exposures and multiple outcomes. Under weaker assumptions on unmeasured confounding variables, we can estimate an upper bound on the bias caused by unmeasured confounders. Having bounds on the bias due to unmeasured confounding implies that we can also bound the causal effect of interest. We can simply estimate the causal effect, adjusting for only measured confounders $\boldsymbol{X}$, and then add (subtract) the worst-case amount of bias to obtain an upper (lower) bound on the causal effect. If our assumptions hold, then this bound should contain the true causal effect of interest, although we did not measure $\boldsymbol{U}$. If this interval does not contain zero, then that would represent strong evidence of a causal effect of the exposures on the outcome of interest. In practice, however, this bound captures the worst-case scenarios for the confounding bias and may be very wide and uninformative, containing all plausible values of the effect of interest. For this reason, we now focus on two ways to shrink the width of this bound using other information about the problem at hand.

6.3.2 Incorporating Negative Controls to Sharpen Partial Identification Regions

Negative controls represent one approach to reducing the size of the partial identification regions described in the previous section. There are many different types of negative controls, such as negative control outcomes, which are outcomes known not to be causally affected by the exposure of interest. An example negative control for air pollution could be hospitalization rates due to falls or other accidents, which are not plausibly caused by air pollution levels. Alternatively, negative control exposures are exposures known not to causally affect the outcome of interest. We can have more general negative control pairs, which are any other exposure and outcome pairs for which a causal effect is known not to exist. These pairs have been used in air pollution studies previously by using future air pollution or air pollution in a distant city as a negative control exposure.[142] These ideas have close ties to work in causal inference on proxy variables for which confounders are not measured, but a noisy proxy of the unmeasured confounder is observed.[143,144] These studies, and other works studying negative control variables, have shown that they can even be used to remove unmeasured confounding bias under certain assumptions on the negative control variables.[134,144–146] These assumptions can be weakened, however, to obtain partial identification using proxy variables instead of point identification of causal effects.[147] Similarly, negative controls have been used to provide more informative sensitivity analyses that can provide more robust evidence on the presence of a causal effect.[148] We take a similar perspective as these final two ideas in that we use negative controls to refine our partial identification region, without necessarily trying to identify the causal effect exactly. In our setting, negative controls can be useful because they can provide additional information on the $\boldsymbol{\Gamma}$ and $\boldsymbol{B}$ matrices, which can reduce the widths of our partial identification regions and provide more informative bounds on what the causal effect of the air pollution mixture is.

Full mathematical details of how to incorporate negative controls and what the resulting partial identification regions are can be found in Kang and colleagues 2023).[140] The core idea underlying negative controls is that these are variables for which it is *known* a priori that no causal effect exists. For instance, a negative control outcome is a variable that we know must not be causally affected by the exposure of interest. We can, however, still estimate the causal effect of the exposure on this negative control outcome. For this particular effect, we have the following informal equation:

$$\text{Estimated effect} = \underbrace{\text{Confounding bias}}_{\text{Function of }(\Gamma, B)}.$$

Normally, the estimated effect is equal to the true effect plus confounding bias, but in this case, we know that the true effect is zero, which leads to the equation above. The confounding bias is a function of both $\mathbf{\Gamma}$ and $\boldsymbol{B}$, and therefore this equation imposes additional constraints on the set of possible values for these coefficients. Given that we have more constraints on $\mathbf{\Gamma}$ and $\boldsymbol{B}$, the set of possible causal effects is further restricted, showing that negative controls can reduce the width of the partial identification region by ruling out certain values of the causal effect. We already had constraints on $\mathbf{\Gamma}$ and $\boldsymbol{B}$ without negative controls because we knew the value of $\mathbf{\Gamma}\mathbf{\Gamma}^T$ and $\boldsymbol{B}\boldsymbol{B}^T$, but now we have additional constraints, which provide more information about the causal effect of interest.

Figure 18 highlights the utility of negative controls in a range of simulated examples and shows that they vary greatly in terms of how much benefit they provide. In all scenarios, the true causal effect is contained inside the partial identification region, whether negative controls are used or not, but the widths of the partial identification regions can change substantially with negative controls. Scenario 1 shows one situation in which negative controls provided very little benefit, whereas scenario 3 is a situation in which negative controls are almost able to pinpoint the causal effect exactly. The regions can also be disjoint, as seen in scenario 4, in which the negative control variable provided information on the magnitude of confounding bias but not on the direction of bias. Negative controls are most useful when they have confounding mechanisms similar to the exposures and outcomes of interest. For instance, if our outcome of interest is given by $\boldsymbol{a}^T\mathbf{Y}$ and the negative control outcome is given by $\boldsymbol{b}^T\mathbf{Y}$, then similar confounding mechanisms would mean that $\boldsymbol{a}^T\mathbf{\Gamma}$ and $\boldsymbol{b}^T\mathbf{\Gamma}$ are similar. This means that the effects of the unmeasured variables on both the outcome of interest and the negative control outcome are similar. Similarly, we want the unmeasured variables to affect the exposure of interest and the negative control exposure in similar ways.

6.3.3 Joint Partial Identification Regions

A related idea that can provide more information on causal effects of interest is to examine at multiple estimands simultaneously, as in our application in the Medicare cohort, for which we are interested in assessing the effects of air pollution on multiple different outcomes. For now, consider the scenario in which we are interested in estimating two distinct causal estimands. One approach would be to apply the aforementioned process for constructing partial identification regions, but do so separately for each of the two outcomes. Alternatively, we can consider both estimands simultaneously and aim to find a bivariate partial identification region that captures the set of possible values for each of the two estimands. These two approaches do not lead to the same set of partial identification regions for the two estimands, as the second approach, which considers the estimands jointly, should lead to smaller partial identification regions. The reason is that a particular causal effect may be plausible for one estimand but only for certain values of the second estimand, and the joint approach can

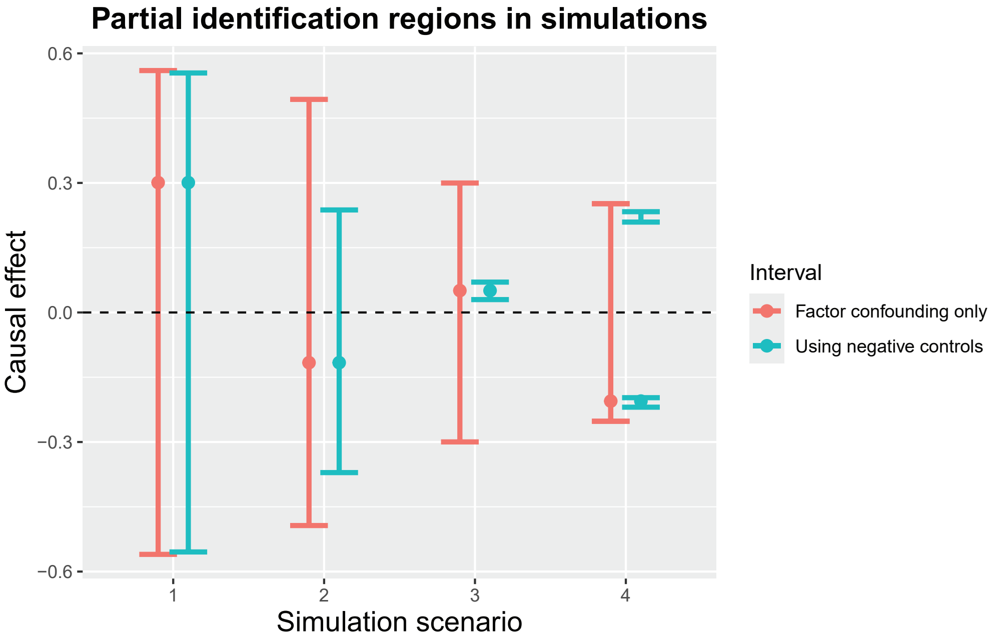


Figure 18. Partial identification regions both with and without negative control information. The points for each scenario represent the true value of the causal effect.

incorporate this restriction. **Figure 19** shows the bivariate partial identification region for two estimands in a simulated example, wherein the partial identification regions are found separately for each estimand or jointly. The joint region is a small subset of the one found separately for each estimand. It is also far more informative, as it shows that while one of the estimands can be zero, both cannot be zero simultaneously, which shows there is a nonzero causal effect for at least one of them. Therefore, if multiple estimands are of interest, which is the case in our Medicare study discussed in Section 6.4, they should be considered jointly when constructing partial identification regions.

6.4 ROBUSTNESS OF THE CAUSAL EFFECTS OF $PM_{2.5}$

We now return to the Medicare analysis described in Sections 4.4 and 5.5 to evaluate whether the effects of air pollution are robust to the presence of unmeasured confounding using the methodology described above. The analysis is run in the same population as in the previous analyses, although a few differences are worth mentioning. For one, we have combined data from all years into one analysis instead of separating analyses by year, and therefore, we will estimate a single effect for each estimand considered. We also focus on multiple outcomes in this analysis as we continue to look at mortality, but now also examine hospitalization rates for anemia, chronic obstructive pulmonary disease (COPD), lung cancer, stroke, hypertension, and asthma. The set of exposures considered in this analysis is $PM_{2.5}$, ozone, BC, EC, NH_4, SO_4, and NIT. We standardize all exposures and outcomes to have a mean of zero and a standard deviation of 1 before running the analysis so that the effect estimates are all on the same scale across outcomes and exposure combinations. For example, an estimate of 0.1 for a particular shift in exposures would be interpreted as being expected to change the outcome by 0.1 outcome standard deviations. For simplicity, we estimate the effects of the exposures (and observed confounders) on the outcome using a linear regression model. We additionally fit a nonlinear model with interactions using multivariate adaptive regression splines (MARS)[149] and found similar estimates for the effects of interest.

6.4.1 Partial Identification Regions Incorporating Negative Controls

We now use the proposed methodology to evaluate whether the health effects of $PM_{2.5}$ are robust to unmeasured confounding bias or whether they are plausibly driven entirely by confounding bias. We first must define our estimands of interest, which depend on two levels of the exposure vector given by $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$. For all exposures other than $PM_{2.5}$, we set both of these values to their median value. For $PM_{2.5}$, we set the $\boldsymbol{w}_1$ value to be the third quartile of $PM_{2.5}$ and set the $\boldsymbol{w}_2$ value to be the first quartile, which means that our estimand corresponds to a shift in $PM_{2.5}$ from the first to third quartile of its observed values, while fixing all other exposures. This

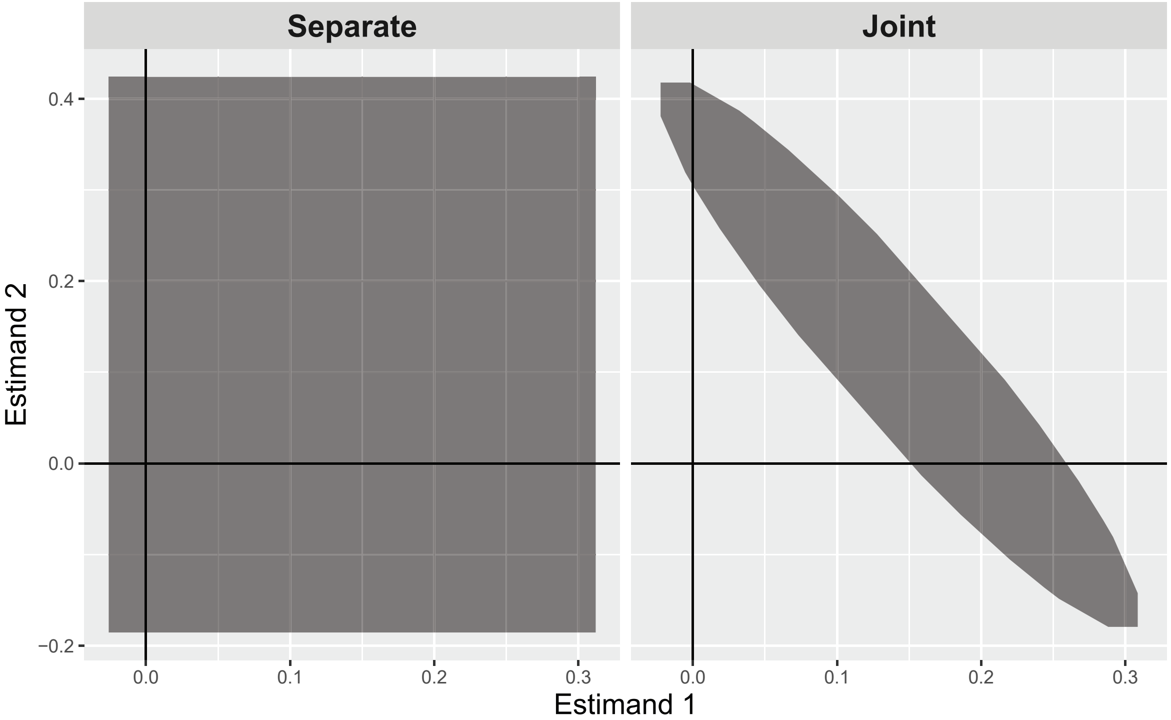


Figure 19. Comparison of bivariate partial identification regions when the regions are found separately for each estimand or when they are considered jointly.

estimand may seem contradictory given that the other exposures we are fixing are themselves components of $PM_{2.5}$, but this well-defined estimand corresponds to increasing $PM_{2.5}$ exposure from other components of particluate matter not included in the other exposures we have measured. We estimate the effect of this exposure shift on all seven outcomes considered in our analysis, and these results can be found in **Figure 20**. The blue lines correspond to standard 95% confidence intervals that assume no unmeasured confounding and provide mixed results about the effects of $PM_{2.5}$, as there are harmful effects on some outcomes and protective effects on other outcomes. These disparate results could be caused by unmeasured confounding bias; however, we now estimate partial identification regions that account for the presence of this bias. The red lines show the partial identification regions that are obtained from our proposed procedure without incorporating any information from negative controls. It is important to note how these differ from standard confidence intervals in terms of their interpretation. The upper and lower limits represent estimates we would obtain under the most extreme bias scenarios consistent with the plausible values of the $\boldsymbol{B}$ and $\boldsymbol{\Gamma}$ matrices. All points inside this partial identification region are equally plausible given the observed data. We also account for sampling uncertainty in the estimates of the upper and lower bounds, as we do for a standard confidence interval, but they are much wider due to the inclusion of a range of values of plausible causal effects instead of the single point estimate obtained when we assume no unmeasured confounding.

We can see that the partial identification regions are exceedingly wide, containing all reasonable values of the effects of $PM_{2.5}$, including some implausible values that are certainly not the true values for these effects. For one, we do not expect strong protective effects of air pollution, yet these are included in the intervals. Additionally, the partial identification regions include overly large harmful effects of pollution, making these regions relatively uninformative about the true value of the causal effect. With this in mind, we incorporate prior COPD hospitalizations as a negative control outcome to provide more informative partial identification regions. Prior COPD hospitalizations are defined as the number of COPD hospitalizations in the year before exposure was assigned, which therefore cannot be caused by the exposure. The green lines show how the partial identification region shrinks when we additionally incorporate prior COPD as a negative control outcome. The regions are generally smaller, even substantially so for certain outcomes, but they are still relatively large and cover all of the reasonable values of the causal effect of $PM_{2.5}$ for all outcomes.

6.4.2 Incorporating Additional Prior Information

The partial identification regions found in Figure 20 were effectively uninformative about the causal effects of $PM_{2.5}$. They included all reasonable positive values, which signify a harmful effect of pollution, as well as no causal effect, and even largely protective effects of pollution. This does not mean that any of these values are more favored by the data

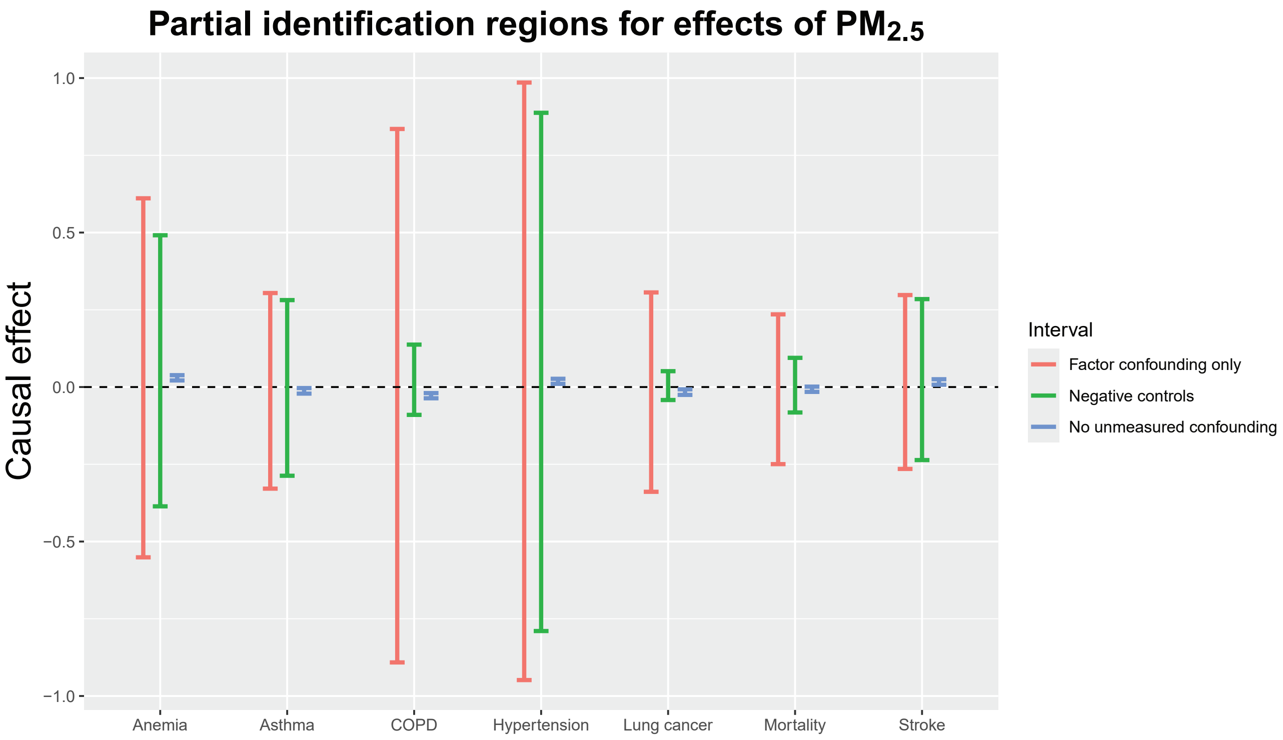


Figure 20. Estimates of the causal effect of $PM_{2.5}$ under no unmeasured confounding, as well as partial identification regions that account for the potential presence of unmeasured confounding bias.

than the others; it only indicates that none can be ruled out due to the potential presence of unmeasured confounding. The widths of these regions could be further reduced in several ways to make them more informative, such as by including more negative control outcomes or negative control exposures or by looking at joint partial identification regions as described in Section 6.3. Given the results above, we take a slightly different approach to find more informative partial identification regions.

By definition, partial identification regions represent the worst-case bounds for the causal effect across all possible values of the $\boldsymbol{B}$ and $\boldsymbol{\Gamma}$ matrices. Not all of these possible coefficient matrices are scientifically plausible, however, as they lead to overly large, unrealistic effect estimates for one or more of the estimands considered. Here, we find the partial identification regions with the worst-case bias for the causal effects, but we restrict only to $\boldsymbol{B}$ and $\boldsymbol{\Gamma}$ matrices that lead to effect estimates within a prespecified range for all estimands. In our case, we consider effect estimates to be plausible only if they lie between −0.05 and 0.1, allowing for moderately protective effects, no effects, and harmful effects of pollution, all of which are within a reasonable range that could plausibly result from such a shift in air pollution values. These results can be found in **Figure 21**, which shows far more informative bounds for the effects of interest. We cannot rule out any effects in our prespecified window for hypertension, but all other estimands have much smaller partial identification regions. We see that the intervals for both COPD and anemia are entirely contained above zero, indicating that the only plausible values are those that represent harmful effects of pollution on these outcomes. The estimates for lung cancer, mortality, and stroke are all fairly positive, although they still contain zero, indicating that zero is still plausible and cannot be entirely ruled out because of unmeasured confounding in this setting.

As described in Section 6.3, we can also investigate the bivariate partial identification regions for two estimands at a time, and this can be more informative than looking at one outcome alone. **Figure 22** shows the bivariate partial identification region for anemia and COPD, as well as for mortality and asthma. The bivariate region for COPD and anemia contains only positive values for both estimands, showing that these effects are robust to unmeasured confounding bias. We see that the region follows an upward trajectory from left to right, which shows that both estimands have either small or large causal effects. The interval for mortality and asthma can be used to help think about the effect of $PM_{2.5}$ on mortality. We see that $PM_{2.5}$ can only have no effect (or a protective effect) on mortality if $PM_{2.5}$ has a protective effect on asthma. Given prior scientific literature on the relationship between $PM_{2.5}$ and asthma, we think it is unlikely that ambient air pollution has a protective effect with respect to asthma, which implies that there is likely some harmful effect of $PM_{2.5}$ on mortality. This contrast highlights the importance of examining bivar-

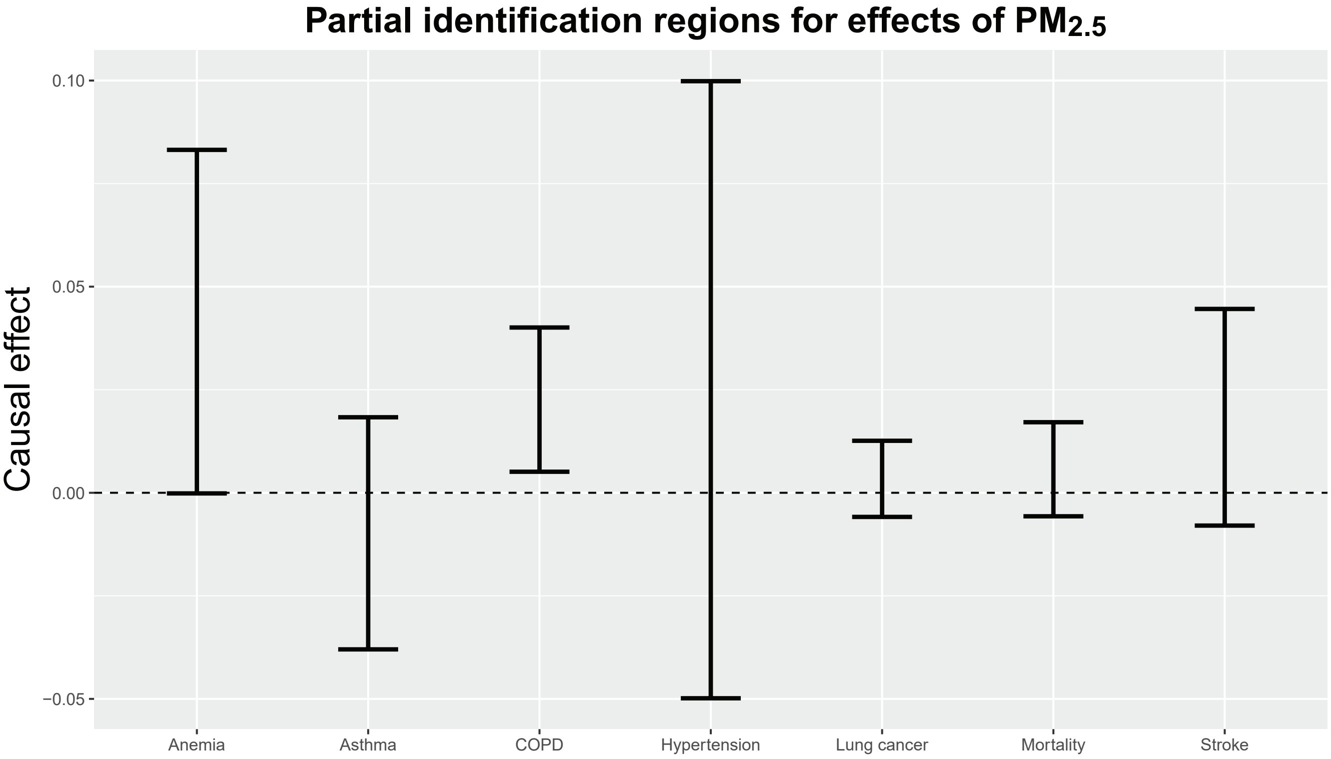


Figure 21. Partial identification regions after prior information on the set of plausible effect sizes is incorporated.

iate partial identification regions as well, as the univariate region for mortality shown in Figure 21 suggested that no effect on mortality was plausible, whereas the bivariate region shows an effect on mortality, assuming no protective effect on asthma.

6.5 DISCUSSION AND CONCLUSIONS

In this chapter, we described the methodology we developed to assess how robust estimates of the health effects of air pollution are to the presence of unmeasured confounding bias. This work builds on that discussed in Chapters 4 and 5, in which we made the strong assumption that there were no unmeasured confounders, and allows us to provide stronger evidence for the presence of adverse health effects of pollution. We showed how to construct partial identification regions for the effects of interest when studying multiple exposures and multiple outcomes simultaneously. We further illustrated how incorporating additional information, such as negative control variables or prior knowledge about the magnitude of plausible effect sizes, can greatly improve these partial identification regions and yield more informative bounds for causal effects in the presence of unmeasured confounding bias. Applying these ideas to a national study of health effects in the Medicare cohort led to several important findings that strengthen the evidence for adverse effects of $PM_{2.5}$ on multiple health outcomes. Despite the potential presence of unmeasured confounders, our analysis ruled out null effects of $PM_{2.5}$ on both anemia and COPD. Further analysis suggests likely effects on other outcomes as well, including mortality, providing robust evidence of the presence of harmful effects of ambient air pollution on public health.

Observational studies have been widely used to estimate the health effects of air pollutants and are a key piece of evidence used to help inform environmental regulatory policy. Despite their limitations, observational studies have provided strong evidence of a causal effect on health outcomes by using a wide variety of study designs and by replicating findings in different cohorts at different exposure levels. Our work here provides an additional, distinct, and strong piece of evidence that further supports the notion that air pollution leads to increased hospitalizations and death. Arguably, the main limitation of observational studies is the potential presence of unmeasured variables biasing estimates of the health effects of air pollution, which can lead to an incorrect assessment of whether a particular exposure causes a health outcome. Our analysis shows that while unmeasured confounding may bias our results to some degree, it cannot fully explain the estimated effects of $PM_{2.5}$, and these findings are indeed robust to the presence of unmeasured confounding.

The main limitation of the proposed methodology is the assumption that models (2)–(4) hold, and that the unmeasured confounders affect multiple exposures or multiple outcomes simultaneously. While these are assumptions in their own right, we argue that they are far weaker than assuming there are no unmeasured confounders at all. Future work can build on these ideas and construct partial identification regions that account for unmeasured confounding bias under weaker modeling assumptions than those used here. Although this approach has limitations, we believe it is a meaningful step forward in addressing the issue of unmeasured confounding bias for air pollution mixtures and offers researchers additional tools to evaluate robustness to this crucial assumption.

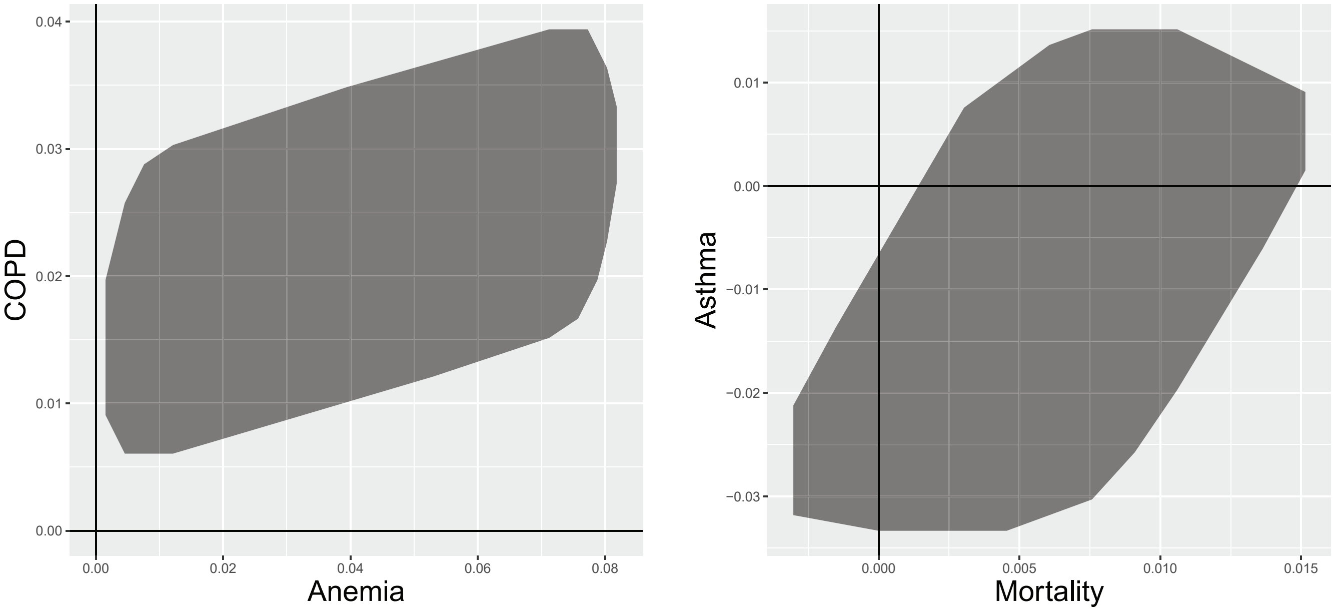


Figure 22. Bivariate partial identification regions after prior information on the set of plausible effect sizes is incorporated.

CHAPTER 7: SYNTHESIS, INTERPRETATION, AND IMPLICATIONS OF FINDINGS

Significant progress has been made toward estimating the health effects of environmental mixtures, but many challenges remain. In this report, we addressed many of these key issues and documented our development of more rigorous tools for estimating causal effects of environmental mixtures, which can provide policymakers with improved regulatory evidence necessary for regulating complex air pollution mixtures. The main hurdle to causal inference with air pollution mixtures is the set of core identifying assumptions necessary for estimating causal effects from observational data. Some of these assumptions, such as positivity, are not particularly well understood and are commonly disregarded. Others, such as the no-unmeasured-confounding assumption, are more commonly discussed but are particularly problematic as they are not testable and can lead to misleading results if not addressed. Throughout this report, we aimed to lower these barriers to causal inference in the air pollution mixture setting by providing solutions to thorny issues stemming from these assumptions that can otherwise block progress toward understanding the causal effect of air pollution on public health. In Chapter 3, we discussed in detail the assumption of mixture positivity and how it is likely to be particularly problematic in air pollution mixture settings, and we provided alternative estimands that rely on weaker assumptions yet still provide important evidence on the health effects of multivariate exposures. In Chapter 5, we tackled the no-interference assumption and its relationship to the mobility of individuals to areas with different exposures. We showed the implications of ignoring this mobility, along with statistical methodology to incorporate mobility if such information is available. In Chapter 6, we tackled the key assumption that there are no unmeasured confounders of the air pollution–outcome relationship, which is a necessary yet very strong assumption commonly made in causal analyses of the health effects of air pollution. We showed that there are significant benefits to studying multiple exposures and multiple outcomes simultaneously, as certain features of the effect of the unmeasured confounders on the exposures and outcomes can be estimated even without data on these unmeasured variables. While this approach is typically not sufficient to completely correct for biases from unmeasured variables, it can provide very informative partial identification regions that can help identify the presence of health effects of air pollution. Lastly, while not a core identifying assumption for causal inference, in Chapter 4, we alleviated a different, commonly made assumption that mixture effects are homogeneous with respect to the characteristics of individuals being exposed to pollution. It is well understood that certain subgroups of the population are more susceptible to the deleterious effects of air pollution, and our proposed methodology allows users to identify these subgroups in a flexible way that makes very limited statistical modeling assumptions.

In addition to advancing the statistical framework for air pollution mixtures, we have provided strong epidemiological evidence of the presence of adverse health effects of air pollution on a variety of public health outcomes. In each of Sections 4.4, 5.5, and 6.4, we evaluated the effects of air pollution mixtures on a large-scale, national study of Medicare enrollees in the United States and consistently found evidence that air pollution negatively affects public health. These analyses contribute to the vast literature highlighting adverse health effects of ambient air pollution and extend them in several ways. Our heterogeneity analysis showed that the effects of the air pollution mixture are more pronounced among low socioeconomic status individuals, which extends prior research on the effect modification of $PM_{2.5}$ effects to the multivariate air pollution mixture setting. Our mobility analysis has several implications for the broader field of air pollution epidemiology. For one, it showed that ignoring the mobility of individuals will likely bias health effect estimates toward the null, which leads to conservative results when such information is not available. Additionally, if this information is available, it can be incorporated into an analysis to potentially increase the magnitude of effect estimates, which can increase statistical power to detect small to moderate exposure effects. Arguably, our most important finding from our Medicare analyses was discussed in Chapter 6, in which we evaluated how robust the estimated effects were to the presence of unmeasured confounding. A common criticism of observational studies is that any significant health effects seen are solely driven by unmeasured variables and are not due to a true effect of pollution on public health outcomes. Our analysis provides strong evidence that the harmful effects of $PM_{2.5}$ cannot be driven solely by unmeasured variables and that there is indeed an adverse effect of $PM_{2.5}$ on several health outcomes.

While our results have found consistent, detrimental effects of air pollution on public health in the Medicare cohort, these results come with inherent limitations. The largest such limitation is that our analyses were performed at the zip code level, and not the individual level, limiting our ability to infer effects of air pollution on individuals in Medicare. While this is a limitation across all three analyses of the Medicare cohort performed, the degree to which it limits our findings varies across our analyses. Certainly, the degree of effect modification, which was explored in Chapter 4, would be expected to be different at the individual level. As one example, the extent to which age modifies the effects of air pollution would likely change in individual-level studies in which there is significantly higher variability in the ranges of ages observed than the small ranges observed when looking at zip code level average age. Future work should look to apply the same ideas developed in Chapter 4 to individual-level studies to better understand which variables modify the effect of air pollution the most. Aggregation to the zip code level is arguably less problematic in our study of mobility. We do not have mobility information on specific individuals in Medicare; therefore, the mobility information can only provide aggregate-level mobility information on individuals from each zip code. This approach likely provides a better assessment of the average mobility of individuals within a zip code than it does for each individual, as individual mobility patterns can vary greatly from person to person, and this is less problematic when aggregating data to the zip code level. Future work would be required to assess how to better incorporate anonymous mobility data, such as the data we used, within individual-level studies. Individual-level mobility data are available, although not for the individuals in our cohort; when available, this information could be leveraged to help address biases from ignoring mobility altogether. Lastly, the methodology developed in Chapter 6 to assess the effect of unmeasured confounding would apply immediately to settings with individual-level data, and future studies should look to incorporate these ideas within their analyses.

While our focus is on the development of tools for a rigorous framework for estimating causal effects of air pollution mixtures, the results found and the statistical methodology developed have broader implications for both statistics and environmental epidemiology. From a statistical point of view, our results have extended the literature on causal inference for multivariate, continuous exposures, which has been fairly limited up to this point. Many of the statistical ideas developed, while motivated for use in environmental epidemiology, have direct implications for use in other scientific areas. For example, unmeasured confounding is a ubiquitous problem in observational studies, and the results in Chapter 6 can be used in any study with multiple treatments and multiple outcomes. The tools developed also have multiple implications for the real-world practice of epidemiology, regardless of whether the analyses are explicitly targeting causal effects. Whenever researchers study the health effects of air pollution, they must target a particular estimand that answers their scientific question of interest. We provide simple and easy-to-use tools in Chapter 3 that can inform researchers whether these are estimands that are easy to estimate with their observed data, and provide similarly straightforward alternatives when the data does not support estimation of the original target estimand. Mixtures analysis has become very popular in recent years within environmental epidemiology and is commonly used in applied analyses. The tools discussed in Chapter 4 provide many of the same benefits of existing approaches in terms of being flexible and allowing for interactions between exposures, but can also be utilized to study questions of effect modification that are commonly of scientific interest. Our mobility findings in Chapter 5 also apply to any air pollution epidemiology study in which exposure assessment is done at one's home location or region, although they travel to areas with different exposure levels. Even if one does not have access to mobility data, our findings and derivations about biases obtained when mobility is ignored are still useful. Additionally, these findings are not only applicable to causal effects but are applicable to regression coefficients or other conditional associations typically estimated in such studies. Lastly, we believe that unmeasured confounding is one of the main hurdles to studying the health effects of air pollution on public health outcomes. The tools we developed and described in Chapter 6 can be applied in any study that has either multiple exposures or multiple outcomes, or both. Further, these ideas can be applied using standard statistical software for regression models and factor analysis, making them easy to use and fast computationally, which should make them easier to implement in practice.

Despite the gains we have made, several opportunities remain to strengthen the evidence provided by analyses of air pollution mixtures. One further advancement, which was discussed earlier in the context of our Medicare analyses, would be to apply the same ideas to individual-level studies of the health effects of air pollution. This approach would be most difficult for the nonparametric Bayesian methodology developed in Chapter 4 because of the computation time required to run analyses on such a large scale. Our analysis of heterogeneity among Medicare zip codes took over a day to run, and analyzing datasets with millions of observations is not currently feasible. Future work could potentially incorporate approximations or frequentist versions of the approach that do not rely on MCMC sampling. Another opportunity for future efforts, which could greatly broaden its applicability to the real-world practice of epidemiology, is to weaken some of the modeling assumptions necessary for the results developed around unmeasured confounding in Chapter 6. Additionally, this approach currently applies to continuous outcomes only and does not immediately extend to other commonly used outcome types such as binary or survival outcomes, and this extension would widen the scope of analyses that could benefit from the methodology. Another ubiquitous problem in air pollution epidemiology that was not discussed in detail in this report is exposure measurement error. While we believe the implications of exposure measurement error in our analyses are similar to those in other contexts that have been previously studied, future research could combine the ideas presented here with approaches to correcting for measurement error. Similarly, while we have discussed separate issues within each chapter of the report, future work could combine these and other important ideas into a single analysis to provide a more robust, comprehensive assessment of the health effects of air pollution. Combining these ideas and addressing so many issues at once is a difficult task, but one that we should strive for as we aim to provide the best science for informing regulatory policy in air pollution epidemiology.

DATA AVAILABILITY STATEMENT

We processed and unified data attributes across multiple sources and followed best practices for FAIR (findable, accessible, interoperable, and reusable) research data sharing. The main data assets can be classified into three groups: health outcomes, air pollution exposures, and socioeconomic covariates.

For the health outcomes, open access to the Medicare data is restricted under the Health Insurance Portability and Accountability Act (HIPAA). However, the pipeline with which the data was processed is available at *https://github.com/NSAPH-Data-Processing/mbsf_mortality_denom*, and the metadata with details of the derived dataset is found at *https://doi.org/10.7910/DVN/Y1WNU7*. The US air pollution estimates at the grid level[30] were transformed to obtain a dataset of derived measures at zip code tabulation area level (*https://doi.org/10.7910/DVN/2NT5CV*). Similarly, US Census Bureau variables were fetched and processed, leading to a clean dataset (*https://doi.org/10.7910/DVN/9V5WCM*).

The code and results associated with the analysis are open for use. Specifically, *https://github.com/hshin111/SepBART_Medicare_Analysis* was used in Chapter 4, *https://github.com/hshin111/Mobility_Medicare_Analysis* in Chapter 5, and *https://github.com/jantonelli111/MultiOutcomeAnalysis* in Chapter 6.

ACKNOWLEDGMENTS

We would like to sincerely thank the Health Effects Institute for its support of our work, without which this research would not have been possible. We would also like to thank several collaborators who have contributed to different components of this report. We thank Corwin Zigler for his advice and expertise when completing the work in Chapter 3, as well as his overall guidance and support throughout the entirety of this work. We thank Antonio Linero for his help and expertise with the development of Bayesian nonparametric methods in Chapter 4. We thank Kezie Irene for her programming and data management expertise for work in Chapters 4 and 5. Lastly, we want to thank Michael Daniels and Michelle Bell for their guidance throughout the completion of this work.

REFERENCES

1. Burnett R, Chen H, Szyszkowicz M, Fann N, Hubbell B, Pope CA III, et al. 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proc Natl Acad Sci 115:9592–9597, *https://doi.org/10.1073/pnas.1803222115*.
2. Dockery DW, Pope CA, Xu X, Spengler JD, Ware JH, Fay ME, et al. 1993. An association between air pollution and mortality in six US cities. N Engl J Med 329:1753–1759, *https://doi.org/10.1056/NEJM199312093292401*.
3. Peng RD, Dominici F, Pastor-Barriuso R, Zeger SL, Samet JM. 2005. Seasonal analyses of air pollution and mortality in 100 US cities. Am J Epidemiol 161:585–594, *https://doi.org/10.1093/aje/kwi075*.
4. Pope CA III, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, et al. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 287:1132–1141, *https://doi.org/10.1001/jama.287.9.1132*.
5. Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, et al. 2000. The National Morbidity, Mortality, and Air Pollution Study. Part II: Morbidity and Mortality from Air Pollution in the United States. Research Report 94. Boston, MA: Health Effects Institute.
6. Dominici F, Peng RD, Barr CD, Bell ML. 2010. Opinion: protecting human health from air pollution: shifting from a single-pollutant to a multipollutant approach. Epidemiol:187–194, *https://doi.org/10.1097/EDE.0b013e3181cc86e8*.
7. Davalos AD, Luben TJ, Herring AH, Sacks JD. 2017. Current approaches used in epidemiologic studies to examine short-term multipollutant air pollution exposures. Ann Epidemiol 27:145–153, *https://doi.org/10.1016/j.annepidem.2016.11.016*.
8. Brauer M, Brook JR, Christidis T, Chu Y, Crouse DL, et al. 2019. Mortality–Air Pollution Associations in Low-Exposure Environments (MAPLE): Phase 1. Research Report 203. Boston, MA: Health Effects Institute.
9. Dominici F, Schwartz J, Di Q, Braun D, Choirat C, Zanobetti A. 2019. Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution: Phase 1. Research Report 200. Boston, MA: Health Effects Institute.
10. Brunekreef B, Strak M, Chen J, Andersen ZJ, Atkinson R, Bauwelinck M, et al. 2021. Mortality and Morbidity Effects of Long-Term Exposure to Low-Level $PM_{2.5}$, BC, NO_2, and O_3: An Analysis of European Cohorts in the ELAPSE Project. Research Report 208. Boston, MA: Health Effects Institute.
11. Brauer M, Brook JR, Christidis T, Chu Y, Crouse DL, Erickson A, et al. 2022. Mortality–Air Pollution Associations in Low Exposure Environments (MAPLE): Phase 2. Research Report 212. Boston, MA: Health Effects Institute.
12. Dominici F, Zanobetti A, Schwartz J, Braun D, Sabath B, Wu X. 2022. Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution: Implementation of Causal Inference Methods. Research Report 211. Boston, MA: Health Effects Institute.
13. Liu JC, Peng RD. 2018. Health effects of mixtures of ozone, nitrogen dioxide, and fine particulates in 85 US counties. Air Qual Atmos Health 11:311–324, *https://doi.org/10.1007/s11869-017-0544-2*.
14. Carrico C, Gennings C, Wheeler DC, Factor-Litvak P. 2015. Characterization of weighted quantile sum regression for highly correlated data in a risk analysis setting. J Agric Biol Environ Stat 20:100–120, *https://doi.org/10.1007/s13253-014-0180-3*.
15. Keil AP, Buckley JP, O'Brien KM, Ferguson KK, Zhao S, White AJ. 2020. A quantile-based g-computation approach to addressing the effects of exposure mixtures. Environ Health Perspect 128:047004, *https://doi.org/10.1289/EHP5838*.
16. Herring AH. 2010. Nonparametric Bayes shrinkage for assessing exposures to mixtures subject to limits of detection. Epidemiology 21:S71, *https://doi.org/10.1097/EDE.0b013e3181cf0058*.
17. Bobb JF, Valeri L, Claus Henn B, Christiani DC, Wright RO, Mazumdar M, et al. 2015. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. Biostat 16:493–508, *https://doi.org/10.1093/biostatistics/kxu058*.
18. Narisetty NN, Mukherjee B, Chen Y-H, Gonzalez R, Meeker JD. 2019. Selection of nonlinear interactions by a forward stepwise algorithm: application to identifying environmental chemical mixtures affecting health outcomes. Stat Med 38:1582–1600, *https://doi.org/10.1002/sim.8059*.
19. Antonelli J, Mazumdar M, Bellinger D, Christiani D, Wright R, Coull B. 2020. Estimating the health effects of environmental mixtures using Bayesian semiparametric regression and sparsity-inducing priors. Ann Appl Stat 14:257–275, *https://doi.org/10.1214/19-AOAS1307*.
20. Ferrari F, Dunson DB. 2021. Bayesian factor analysis for inference on interactions. J Am Stat Assoc 116:1521–1532, *https://doi.org/10.1080/01621459.2020.1745813*.
21. Ferrari F, Dunson DB. 2020. Identifying main effects and interactions among exposures using Gaussian processes. Ann Appl Stat 14:1743, *https://doi.org/10.1214/20-AOAS1363*.

22. Wei R, Reich BJ, Hoppin JA, Ghosal S. 2020. Sparse Bayesian additive nonparametric regression with application to health effects of pesticide mixtures. Stat Sin 30:55–79, *https://doi.org/10.5705/ss.202017.0315*.
23. Samanta S, Antonelli J. 2021. Estimation and false discovery control for the analysis of environmental mixtures. arXiv preprint arXiv:2103.10563, *https://doi.org/10.48550/arXiv.2103.10563*.
24. Weisskopf MG, Seals RM, Webster TF. 2018. Bias amplification in epidemiologic analysis of exposure to mixtures. Environ Health Perspect 126:047003, *https://doi.org/10.1289/EHP2450*.
25. Zigler CM, Kim C, Choirat C, Hansen JB, Wang Y, Hund L, et al. 2016. Causal Inference Methods for Estimating Long-Term Health Effects of Air Quality Regulations. Research Report 187. Boston, MA: Health Effects Institute.
26. Zigler CM, Dominici F. 2014. Point: clarifying policy evidence with potential-outcomes thinking — beyond exposure-response estimation in air pollution epidemiology. Am J Epidemiol 180:1133–1140, *https://doi.org/10.1093/aje/kwu263*.
27. Dominici F, Zigler C. 2017. Best practices for gauging evidence of causality in air pollution epidemiology. Am J Epidemiol 186:1303–1309, *https://doi.org/10.1093/aje/kwx307*.
28. Carone M, Dominici F, Sheppard L. 2020. In pursuit of evidence in air pollution epidemiology: the role of causally driven data science. Epidemiology 31:1, *https://doi.org/10.1097/EDE.0000000000001090*.
29. Sommer AJ, Leray E, Lee Y, Bind M-AC. 2021. Assessing environmental epidemiology questions in practice with a causal inference pipeline: an investigation of the air pollution-multiple sclerosis relapses relationship. Stat Med 40:1321–1335, *https://doi.org/10.1002/sim.8843*.
30. Van Donkelaar A, Martin RV, Li C, Burnett RT. 2019. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. Environ Sci Technol 53:2595–2611, *https://doi.org/10.1021/acs.est.8b06392*.
31. Di Q, Amini H, Shi L, Kloog I, Silvern R, Kelly J, et al. 2019. An ensemble-based model of $PM_{2.5}$ concentrations across the contiguous United States with high spatiotemporal resolution. Environ Int 130:104909, *https://doi.org/10.1016/j.envint.2019.104909*.
32. Di Q, Wei Y, Shtein A, Hultquist C, Xing X, Amini H, et al. 2021. Daily and annual $PM_{2.5}$ Concentrations for the Contiguous United States, 1-km Grids, v1 (2000–2016). Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). Available: *https://doi.org/10.7927/0rvr-4538*.
33. Requia WJ, Wei Y, Shtein A, Hultquist C, Xing X, Di Q, et al. 2021. Daily 8-Hour Maximum and Annual O_3 Concentrations for the Contiguous United States, 1-km Grids, v1 (2000–2016). Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). Available: *https://doi.org/10.7927/10.7927/a4mb-4t86*.
34. Carlin DJ, Rider CV, Woychik R, Birnbaum LS. 2013. Unraveling the health effects of environmental mixtures: an NIEHS priority. Environ Health Perspect 121:A6–A8, *https://doi.org/10.1289/ehp.1206182*.
35. Gibson EA, Nunez Y, Abuawad A, Zota AR, Renzetti S, Devick KL, et al. 2019. An overview of methods to address distinct research questions on environmental mixtures: an application to persistent organic pollutants and leukocyte telomere length. Environ Health 18:1–16, *https://doi.org/10.1186/s12940-019-0515-1*.
36. Joubert BR, Kioumourtzoglou M-A, Chamberlain T, Chen HY, Gennings C, Turyk ME, et al. 2022. Powering research through innovative methods for mixtures in epidemiology (PRIME) program: novel and expanded statistical methods. Int J Environ Res Public Health 19:1378, *https://doi.org/10.3390/ijerph19031378*.
37. Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL. 2000. Fine particulate air pollution and mortality in 20 US cities, 1987–1994. N Engl J Med 343:1742–1749, *https://doi.org/10.1056/NEJM200012143432401*.
38. Portney PR. 1990. Policy watch: economics and the Clean Air Act. J Econ Perspect 4:173–181, *https://doi.org/10.1257/jep.4.4.173*.
39. US Environmental Protection Agency (US EPA). 2011. The Benefits and Costs of the Clean Air Act from 1990 to 2020. Washington, DC: US Environmental Protection Agency. Available: *https://www.epa.gov/sites/default/files/2015-07/documents/fullreport_rev_a.pdf*.
40. Braun JM, Gennings C, Hauser R, Webster TF. 2016. What can epidemiological studies tell us about the impact of chemical mixtures on human health? Environ Health Perspect 124:A6–A9, *https://doi.org/10.1289/ehp.1510569*.
41. Wilson A, Zigler CM, Patel CJ, Dominici F. 2018. Model-averaged confounder adjustment for estimating multivariate exposure effects with linear regression. Biometrics 74:1034–1044, *https://doi.org/10.1111/biom.12860*.
42. Traini E, Huss A, Portengen L, Rookus M, Verschuren WM, Vermeulen RCH, et al. 2022. A multipollutant approach to estimating causal effects of air pollution mixtures on overall mortality in a large, prospective cohort. Epidemiology 33:514, *https://doi.org/10.1097/EDE.0000000000001492*.
43. Hirano K, Imbens GW. 2004. Chapter 7: The propensity score with continuous treatments. In: Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives: An Essential Journey with Donald Rubin's Statistical Family (Shehwart WA, Wilks SS, eds.) New York: John Wiley & Sons, Inc. Available: *https://doi.org/10.1002/0470090456.ch7*.
44. Keil AP, Buckley JP, Kalkbrenner AE. 2021. Bayesian G-computation for estimating impacts of interventions on exposure mixtures: demonstration with metals from coal-fired power plants and birth weight. Am J Epidemiol 190:2647–2657, *https://doi.org/10.1093/aje/kwab053*.
45. Zigler CM. 2021. Invited commentary: the promise and pitfalls of causal inference with multivariate environmental exposures. Am J Epidemiol 190:2658–2661, *https://doi.org/10.1093/aje/kwab142*.
46. Little RJ, Rubin DB. 2000. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. Annu Rev Public Health 21:121–145, *https://doi.org/10.1146/annurev.publhealth.21.1.121*.
47. Smith TJS, Keil AP, Buckley JP. 2023. Estimating causal effects of interventions on early-life environmental exposures using observational data. Curr Environ Health Rep 10:12–21, *https://doi.org/10.1007/s40572-022-00388-y*.
48. Nethery RC, Mealli F, Sacks JD, Dominici F. 2021. Evaluation of the health impacts of the 1990 Clean Air Act Amendments using causal inference and machine learning. J Am Stat Assoc 116:1128–1139, *https://doi.org/10.1080/01621459.2020.1803883*.
49. Haneuse S, Rotnitzky A. 2013. Estimation of the effect of interventions that modify the received treatment. Stat Med 32:5260–5277, *https://doi.org/10.1002/sim.5907*.
50. Papadogeorgou G, Mealli F, Zigler CM. 2019. Causal inference with interfering units for cluster and population-level treatment allocation programs. Biometrics 75:778–787, *https://doi.org/10.1111/biom.13049*.
51. Petersen ML, Porter KE, Gruber S, Wang Y, Van Der Laan MJ. 2012. Diagnosing and responding to violations in the positivity assumption. Stat Methods Med Res 21:31–54, *https://doi.org/10.1177/0962280210386207*.

52. Crump RK, Hotz VJ, Imbens GW, Mitnik OA. 2009. Dealing with limited overlap in estimation of average treatment effects. Biometrika 96:187–199, *https://doi.org/10.1093/biomet/asn055*.
53. Lee BK, Lessler J, Stuart EA. 2011. Weight trimming and propensity score weighting. PLoS One 6:e18174, *https://doi.org/10.1371/journal.pone.0018174*.
54. Li F, Morgan KL, Zaslavsky AM. 2018. Balancing covariates via propensity score weighting. J Am Stat Assoc 113:390–400, *https://doi.org/10.1080/01621459.2016.1260466*.
55. Antonelli J, Zigler CM. 2024. Causal analysis of air pollution mixtures: estimands, positivity, and extrapolation. Am J Epidemiol 193:1392–1398, *https://doi.org/10.1093/aje/kwae115*.
56. King G, Zeng L. 2006. The dangers of extreme counterfactuals. Polit Anal 14:131–159, *https://doi.org/10.1093/pan/mpj004*.
57. Vansteelandt S, Dukes O. 2022. Assumption-lean inference for generalised linear model parameters. J R Stat Soc Ser B Stat Method 84:657–685, *https://doi.org/10.1111/rssb.12504*.
58. Ho DE, Imai K, King G, Stuart EA. 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Polit Anal 15:199–236, *https://doi.org/10.1093/pan/mpl013*.
59. Cole SR, Hernán MA. 2008. Constructing inverse probability weights for marginal structural models. Am J Epidemiol 168:656–664, *https://doi.org/10.1093/aje/kwn164*.
60. Huber M, Lechner M, Wunsch C. 2013. The performance of estimators based on the propensity score. J Econom 175:1–21, *https://doi.org/10.1016/j.jeconom.2012.11.006*.
61. Xiao Y, Moodie EEM, Abrahamowicz M. 2013. Comparison of approaches to weight truncation for marginal structural Cox models. Epidemiol Methods 2:1–20, *https://doi.org/10.1515/em-2012-0006*.
62. Kilpatrick RD, Gilbertson D, Brookhart MA, Polley E, Rothman KJ, Bradbury BD. 2013. Exploring large weight deletion and the ability to balance confounders when using inverse probability of treatment weighting in the presence of rare treatment decisions. Pharmacoepidemiol Drug Saf 22:111–121, *https://doi.org/10.1002/pds.3297*.
63. Branson Z, Kennedy E, Balakrishnan S, Wasserman L. 2023. Causal effect estimation after propensity score trimming with continuous treatments. arXiv, *https://doi.org/10.48550/arXiv.2309.00706*.
64. Agier L, Portengen L, Chadeau-Hyam M, Basagaña X, Giorgis-Allemand L, Siroux V, et al. 2016. A systematic comparison of linear regression-based statistical methods to assess exposome-health associations. Environ Health Perspect 124:1848–1856, *https://doi.org/10.1289/EHP172*.
65. Stafoggia M, Breitner S, Hampel R, Basagaña X. 2017. Statistical approaches to address multi-pollutant mixtures and multiple exposures: the state of the science. Curr Environ Health Rep 4:481–490, *https://doi.org/10.1007/s40572-017-0162-z*.
66. Boss J, Rix A, Chen Y-H, Narisetty NN, Wu Z, Ferguson KK, et al. 2021. A hierarchical integrative group least absolute shrinkage and selection operator for analyzing environmental mixtures. Environmetrics 32:e2698, *https://doi.org/10.1002/env.2698*.
67. Scheipl F, Fahrmeir L, Kneib T. 2012. Spike-and-slab priors for function selection in structured additive regression models. J Am Stat Assoc 107:1518–1532, *https://doi.org/10.1080/01621459.2012.737742*.
68. Wang B, Eum K-D, Kazemiparkouhi F, Li C, Manjourides J, Pavlu V, et al. 2020. The impact of long-term $PM_{2.5}$ exposure on specific causes of death: exposure-response curves and effect modification among 53 million US Medicare beneficiaries. Environ Health 19:1–12, *https://doi.org/10.1186/s12940-020-00575-0*.
69. Bargagli-Stoffi FJ, Cadei R, Lee K, Dominici F. 2020. Causal rule ensemble: interpretable discovery and inference of heterogeneous treatment effects. arXiv preprint arXiv:2009.09036, *https://doi.org/10.48550/arXiv.2009.09036*.
70. Athey S, Imbens G. 2016. Recursive partitioning for heterogeneous causal effects. Proc Natl Acad Sci 113:7353–7360, *https://doi.org/10.1073/pnas.1510489113*.
71. Wager S, Athey S. 2018. Estimation and inference of heterogeneous treatment effects using random forests. J Am Stat Assoc 113:1228–1242, *https://doi.org/10.1080/01621459.2017.1319839*.
72. Hahn PR, Murray JS, Carvalho CM. 2020. Bayesian regression tree models for causal inference: regularization, confounding, and heterogeneous effects (with discussion). Bayesian Anal 15:965–1056, *https://doi.org/10.1214/19-BA1195*.
73. Shin H, Antonelli J. 2023. Improved inference for doubly robust estimators of heterogeneous treatment effects. Biometrics 00:1–13.
74. Levy J, van der Laan M, Hubbard A, Pirracchio R. 2021. A fundamental measure of treatment effect heterogeneity. J Causal Inference 9:83–108, *https://doi.org/10.1515/jci-2019-0003*.
75. Hines O, Diaz-Ordaz K, Vansteelandt S. 2022. Variable importance measures for heterogeneous causal effects. arXiv preprint arXiv:2204.06030, *https://doi.org/10.48550/arXiv.2204.06030*.
76. Li H, Hubbard A, van der Laan M. 2023. Targeted learning on variable importance measure for heterogeneous treatment effect. arXiv preprint arXiv:2309.13324, *https://doi.org/10.48550/arXiv.2309.13324*.
77. Rubin DB. 1980. Randomization analysis of experimental data: the Fisher randomization test comment. J Am Stat Assoc 75:591–593, *https://doi.org/10.1080/01621459.1980.10477517*.
78. Lei J, G'Sell M, Rinaldo A, Tibshirani R, Wasserman L. 2018. Distribution-free predictive inference for regression. J Am Stat Assoc 113:1094–1111, *https://doi.org/10.1080/01621459.2017.1307116*.
79. Zhang L, Janson L. 2020. Floodgate: inference for model-free variable importance. arXiv preprint arXiv:2007.01283, *https://doi.org/10.48550/arXiv.2007.01283*.
80. Verdinelli I, Wasserman L. 2024. Feature importance: a closer look at Shapley values and loco. Stat Sci 39:623–636, *https://doi.org/10.1214/24-STS937*.
81. Shin H, Linero A, Audirac M, Irene K, Braun D, Antonelli J. 2024. Treatment effect heterogeneity and importance measures for multivariate continuous treatments. arXiv preprint arXiv:2404.09126, *https://doi.org/10.48550/arXiv.2404.09126*.
82. Chipman HA, George EI, McCulloch RE. 2010. BART: Bayesian additive regression trees. Ann Appl Stat 4:266–298, *https://doi.org/10.1214/09-AOAS285*.
83. Dorie V, Hill J, Shalit U, Scott M, Cervone D. 2019. Automated versus do-it-yourself methods for causal inference: lessons learned from a data analysis competition. Stat Sci 34:43–68, *https://doi.org/10.1214/18-STS667*.
84. Linero AR. 2017. A review of tree-based Bayesian methods. Commun Stat Appl Methods 24:543–559, *https://doi.org/10.29220/CSAM.2017.24.6.543*.
85. Hill J, Linero A, Murray J. 2020. Bayesian additive regression trees: a review and look forward. Annu Rev Stat Appl 7, *https://doi.org/10.1146/annurev-statistics-031219-041110*.
86. Linero AR, Yang Y. 2018. Bayesian regression tree ensembles that adapt to smoothness and sparsity. J R Stat Soc Ser B Stat Methodol. 80:1087–1110, *https://doi.org/10.1111/rssb.12293*.
87. Li Y, Linero AR, Murray J. 2022. Adaptive conditional distribution estimation with Bayesian decision tree ensembles. J Am Stat Assoc 118:2129–2142, *https://doi.org/10.1080/01621459.2022.2037431*.

88. Gelman A, Rubin DB. 1992. Inference from iterative simulation using multiple sequences. Stat Sci 7:457–472, *https://doi.org/10.1214/ss/1177011136*.
89. Mayeda ER, Filshtein TJ, Tripodis Y, Glymour MM, Gross AL. 2018. Does selective survival before study enrolment attenuate estimated effects of education on rate of cognitive decline in older adults? A simulation approach for quantifying survival bias in life course epidemiology. Int J Epidemiol 47:1507–1517, *https://doi.org/10.1093/ije/dyy124*.
90. Liu J, Clark LP, Bechle MJ, Hajat A, Kim S-Y, Robinson AL, et al. 2021. Disparities in air pollution exposure in the United States by race/ethnicity and income, 1990–2010. Environ Health Perspect 129:127005, *https://doi.org/10.1289/EHP8584*.
91. Jbaily A, Zhou X, Liu J, Lee T-H, Kamareddine L, Verguet S, Dominici F. 2022. Air pollution exposure disparities across US population and income groups. Nature 601:228–233, *https://doi.org/10.1038/s41586-021-04190-y*.
92. Pope CA 3rd, Lefler JS, Ezzati M, Higbee JD, Marshall JD, Kim S-Y, et al. 2019. Mortality risk and fine particulate air pollution in a large, representative cohort of US adults. Environ Health Perspect 127:077007, *https://doi.org/10.1289/EHP4438*.
93. Dedoussi IC, Eastham SD, Monier E, Barrett SRH. 2020. Premature mortality related to United States cross-state air pollution. Nature 578:261–265, *https://doi.org/10.1038/s41586-020-1983-8*.
94. Adegunsoye A, Freiheit E, White EN, Kaul B, Newton CA, Oldham JM, et al. 2023. Evaluation of pulmonary fibrosis outcomes by race and ethnicity in US adults. JAMA Netw Open 6:e232427, *https://doi.org/10.1001/jamanetworkopen.2023.2427*.
95. Fallah-Shorshani M, Hatzopoulou M, Ross NA, Patterson Z, Weichenthal S. 2018. Evaluating the impact of neighborhood characteristics on differences between residential and mobility-based exposures to outdoor air pollution. Environ Sci Technol 52:10777–10786, *https://doi.org/10.1021/acs.est.8b02260*.
96. Nyhan MM, Kloog I, Britter R, Ratti C, Koutrakis P. 2019. Quantifying population exposure to air pollution using individual mobility patterns inferred from mobile phone data. J Expo Sci Environ Epidemiol 29:238–247, *https://doi.org/10.1038/s41370-018-0038-9*.
97. Yu X, Ivey C, Huang Z, Gurram S, Sivaraman V, Shen H, et al. 2020. Quantifying the impact of daily mobility on errors in air pollution exposure estimation using mobile phone location data. Environ Int 141:105772, *https://doi.org/10.1016/j.envint.2020.105772*.
98. Yoo, E-H, Pu Q, Eum Y, Jiang X. 2021. The impact of individual mobility on long-term exposure to ambient PM2. 5: assessing effect modification by travel patterns and spatial variability of $PM_{2.5}$. Int J Environ Res Public Health 18:2194, *https://doi.org/10.3390/ijerph18042194*.
99. Nyhan M, Grauwin S, Britter R, Misstear B, McNabola A, Laden F, et al. 2016. Exposure track: the impact of mobile-device-based mobility patterns on quantifying population exposure to air pollution. Environ Sci Technol 50:9671–9681, *https://doi.org/10.1021/acs.est.6b02385*.
100. Fan C, Chien Y-H, Mostafavi A. 2022. Human mobility disproportionately extends $PM_{2.5}$ emission exposure for low-income populations. arXiv preprint arXiv:2205.15381, *https://doi.org/10.48550/arXiv.2205.15381*.
101. Setton E, Marshall JD, Brauer M, Lundquist KR, Hystad P, Keller P, et al. 2011. The impact of daily mobility on exposure to traffic-related air pollution and health effect estimates. J Expo Sci Environ Epidemiol 21:42–48, *https://doi.org/10.1038/jes.2010.14*.
102. Hong G, Raudenbush SW. 2006. Evaluating kindergarten retention policy: a case study of causal inference for multilevel observational data. J Am Stat Assoc 101:901–910, *https://doi.org/10.1198/016214506000000447*.
103. Sobel ME. 2006. What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. J Am Stat Assoc 101:1398–1407, *https://doi.org/10.1198/016214506000000636*.
104. Hudgens MG, Halloran ME. 2008. Toward causal inference with interference. J Am Stat Assoc 103:832–842, *https://doi.org/10.1198/016214508000000292*.
105. Tchetgen EJT, VanderWeele TJ. 2012. On causal inference in the presence of interference. Stat Methods Med Res 21:55–75, *https://doi.org/10.1177/0962280210386779*.
106. Zigler C, Forastiere L, Mealli F. 2020. Bipartite interference and air pollution transport: estimating health effects of power plant interventions. arXiv preprint arXiv:2012.04831, *https://doi.org/10.48550/arXiv.2012.04831*.
107. Zigler CM, Papadogeorgou G. 2021. Bipartite causal inference with interference. Stat Sci 36:109, *https://doi.org/10.1214/19-STS749*.
108. Zirkle KW, Bind M-A, Swall JL, Wheeler DC. 2021. Addressing spatially structured interference in causal analysis using propensity scores. arXiv preprint arXiv:2101.09297, *https://doi.org/10.48550/arXiv.2101.09297*.
109. Forastiere L, Airoldi EM, Mealli F. 2021. Identification and estimation of treatment and interference effects in observational studies on networks. J Am Stat Assoc 116:901–918, *https://doi.org/10.1080/01621459.2020.1768100*.
110. Sävje F, Aronow P, Hudgens M. 2021. Average treatment effects in the presence of unknown interference. Ann Stat 49:673, *https://doi.org/10.1214/20-AOS1973*.
111. Weinstein B, Nevo D. 2023. Causal inference with misspecified interference structure. arXiv preprint arXiv:2302.11322, *https://doi.org/10.48550/arXiv.2302.11322*.
112. Aronow PM, Samii C. 2017. Estimating average causal effects under general interference, with application to a social network experiment. Ann Appl Stat 11:1912–1947, *https://doi.org/10.1214/16-AOAS1005*.
113. Shin H, Braun D, Irene K, Antonelli J. 2023. A spatial interference approach to account for mobility in air pollution studies with multivariate continuous treatments. arXiv preprint arXiv:2305.14194, *https://doi.org/10.48550/arXiv.2305.14194*.
114. Szpiro AA, Paciorek CJ. 2013. Measurement error in two-stage analyses, with application to air pollution epidemiology. Environmetrics 24:501–517, *https://doi.org/10.1002/env.2233*.
115. Deng H, Du J, Gao J, Wang Q. 2021. Network percolation reveals adaptive bridges of the mobility network response to COVID-19. PLoS One 16:e0258868, *https://doi.org/10.1371/journal.pone.0258868*.
116. Wang F, Wang J, Cao J, Chen C, Ban XJ. 2019. Extracting trips from multi-sourced data for mobility pattern analysis: an app-based data example. Transp Res C Emerg Technol 105:182–202, *https://doi.org/10.1016/j.trc.2019.05.028*.
117. Carvalho CM, Polson NG, Scott JG. 2010. The horseshoe estimator for sparse signals. Biometrika 97:465–480, *https://doi.org/10.1093/biomet/asq017*.
118. Huang C, Nichols C, Liu Y, Zhang Y, Liu X, Gao S, et al. 2015. Ambient air pollution and adverse birth outcomes: a natural experiment study. Popul Health Metr 13:1–7, *https://doi.org/10.1186/s12963-015-0050-4*.
119. Rosenbaum PR. 1987. Sensitivity analysis for certain permutation inferences in matched observational studies. Biometrika 74:13–26, *https://doi.org/10.1093/biomet/74.1.13*.
120. Rosenbaum PR. 1988. Sensitivity analysis for matching with multiple controls. Biometrika 75:577–581, *https://doi.org/10.1093/biomet/75.3.577*.

121. Manski CF. 1990. Nonparametric bounds on treatment effects. Am Econ Rev 80:319–323.

122. Rosenbaum PR. 1991. Sensitivity analysis for matched case-control studies. Biometrics 47:87–100, *https://doi.org/10.2307/2532498*.

123. Rosenbaum PR. 2002. Covariance adjustment in randomized experiments and observational studies. Stat Sci 17:286–327, *https://doi.org/10.1214/ss/1042727942*.

124. Rosenbaum PR. 2002. Overt Bias in Observational Studies. New York: Springer. Available: *https://doi.org/10.1007/978-1-4757-3692-2*.

125. Imbens GW. 2003. Sensitivity to exogeneity assumptions in program evaluation. Am Econ Rev 93:126–132, *https://doi.org/10.1257/000282803321946921*.

126. Small DS. 2007. Sensitivity analysis for instrumental variables regression with overidentifying restrictions. J Am Stat Assoc 102:1049–1058, *https://doi.org/10.1198/016214507000000608*.

127. McCandless LC, Gustafson P, Levy A. 2007. Bayesian sensitivity analysis for unmeasured confounding in observational studies. Stat Med 26:2331–2347, *https://doi.org/10.1002/sim.2711*.

128. Veitch V, Zaveri A. 2020. Sense and sensitivity analysis: simple post-hoc analysis of bias due to unobserved confounding. Adv Neural Inf Process Syst 33:10999–11009, *https://doi.org/10.48550/arXiv.2003.01747*.

129. Cinelli C, Hazlett C. 2020. Making sense of sensitivity: extending omitted variable bias. J R Stat Soc Series B Stat Methodol 82:39–67, *https://doi.org/10.1111/rssb.12348*.

130. Freidling T, Zhao Q. 2022. Sensitivity analysis with the R^2-calculus. arXiv preprint arXiv:2301.00040, *https://doi.org/10.48550/arXiv.2301.00040*.

131. Chernozhukov V, Cinelli C, Newey W, Sharma A, Syrgkanis V. 2022. Long story short: omitted variable bias in causal machine learning. arXiv:2112.13398, *https://doi.org/10.48550/arXiv.2112.13398*.

132. Zhou Y, Tang D, Kong D, Wang L. 2020. The promises of parallel outcomes. arXiv preprint arXiv:2012.05849, *https://doi.org/10.48550/arXiv.2012.05849*.

133. Zheng J, D'Amour A, Franks A. 2021. Copula-based sensitivity analysis for multi-treatment causal inference with unobserved confounding. arXiv preprint arXiv:2102.09412, *https://doi.org/10.48550/arXiv.2102.09412*.

134. Miao W, Hu W, Ogburn EL, Zhou X-H. 2022. Identifying effects of multiple treatments in the presence of unmeasured confounding. J Am Stat Assoc:1–15, *https://doi.org/10.1080/01621459.2021.2023551*.

135. Kong D, Yang S, Wang L. 2022. Identifiability of causal effects with multiple causes and a binary outcome. Biometrika 109:265–272, *https://doi.org/10.1093/biomet/asab016*.

136. Zheng J, D'Amour A, Franks A. 2021. Bayesian inference and partial identification in multi-treatment causal inference with unobserved confounding. arXiv preprint arXiv:2111.07973, *https://doi.org/10.48550/arXiv.2111.07973*.

137. Zheng J, Wu J, D'Amour A, Franks A. 2022. Sensitivity to unobserved confounding in studies with factor-structured outcomes. arXiv preprint arXiv:2208.06552, *https://doi.org/10.48550/arXiv.2208.06552*.

138. Wang J, Zhao Q, Hastie T, Owen AB. 2017. Confounder adjustment in multiple hypothesis testing. Ann Stat 45:1863, *https://doi.org/10.1214/16-AOS1511*.

139. VanderWeele TJ, Mathur MB, Chen Y. 2020. Outcome-wide longitudinal designs for causal inference. Stat Sci 35:437–466, *https://doi.org/10.1214/19-STS728*.

140. Kang S, Franks A, Antonelli J. 2023. Sensitivity analysis with multiple treatments and multiple outcomes with applications to air pollution mixtures. arXiv preprint arXiv:2311.12252, *https://doi.org/10.48550/arXiv.2311.12252*.

141. Wang Y, Blei DM. 2020. Towards clarifying the theory of the deconfounder. arXiv preprint arXiv:2003.04948, *https://doi.org/10.48550/arXiv.2003.04948*.

142. Lumley T, Sheppard L. 2000. Assessing seasonal confounding and model selection bias in air pollution epidemiology using positive and negative control analyses. Environmetrics 11:705–717, *https://doi.org/10.1002/1099-095X(200011/12)11:6<705::AID-ENV444>3.0.CO;2-H*.

143. Kuroki M, Pearl J. 2014. Measurement bias and effect restoration in causal inference. Biometrika 101:423–437, *https://doi.org/10.1093/biomet/ast066*.

144. Miao W, Geng Z, Tchetgen ET. 2018. Identifying causal effects with proxy variables of an unmeasured confounder. Biometrika 105:987–993, *https://doi.org/10.1093/biomet/asy038*.

145. Shi X, Miao W, Tchetgen ET. 2020. A selective review of negative control methods in epidemiology. Curr Epidemiol Rep 7:190–202, *https://doi.org/10.1007/s40471-020-00243-4*.

146. Hu JK, Zorzetto D, Dominici F. 2023. A Bayesian nonparametric method to adjust for unmeasured confounding with negative controls. arXiv preprint arXiv:2309.02631, *https://doi.org/10.48550/arXiv.2309.02631*.

147. Ghassami A, Shpitser I, Tchetgen ET. 2023. Partial identification of causal effects using proxy variables. arXiv preprint arXiv:2304.04374, *https://doi.org/10.48550/arXiv.2304.04374*.

148. Rosenbaum PR. 2023. Can we reliably detect biases that matter in observational studies? Stat Sci 38:440–457, *https://doi.org/10.1214/23-STS882*.

149. Friedman JH. 1991. Multivariate adaptive regression splines. Ann Stat 19:1–67, *https://doi.org/10.1214/aos/1176347963*.

HEI QUALITY ASSURANCE STATEMENT

The conduct of the study "Robust Statistical Approaches to Understanding the Causal Effect of Air Pollution Mixtures" was subjected to an independent audit by David Bush and Scott Adamson of Trinity Consultants, Inc. Mr. Bush and Mr. Adamson are experts in quality assurance for air quality monitoring studies and data management.

The audit included a review of data quality for conformance to the study protocol as detailed in the final report and the study's quality assurance plan, reviewing data quality for each of the study components. In April 2025, an off-site audit was conducted via a teleconferencing platform with primary study personnel. The audit concentrated on the study's quality assurance and data management activities and included a review of the overall process utilized to collect new data and to manage and combine the exposure, air quality, epidemiological, and modeling data. Also evaluated were the procedures and measures undertaken to ensure quality and consistency in the processed databases and modeling results. Examples of data and data processing code for the datasets and modeling files were displayed by study personnel and further reviewed in study GitHub directories for consistency, clarity, and completeness.

A written report of the audit was provided to the HEI project manager, who transmitted the findings to the Principal Investigator. The quality assurance audit demonstrated that the study was conducted by an experienced team with a high concern for data quality. Study personnel were responsive to audit questions and recommendations, providing responses that addressed all audit inquiries. The report appears to be an accurate representation of the study.

David H. Bush, Quality Assurance Officer

ABOUT THE AUTHORS

Joseph Antonelli is an Assistant Professor in the Department of Statistics at the University of Florida. His research focuses on the development of statistical methodology in the areas of causal inference and Bayesian statistics, with an emphasis on applications in environmental health and criminology. He received his PhD from the Department of Biostatistics at Harvard University in 2015.

Heejun Shin is a postdoctoral fellow at the Harvard T.H. Chan School of Public Health in the department of biostatistics. His research interests include causal inference, Bayesian statistics, and environmental statistics. He was a doctoral student in the department of statistics at the University of Florida and contributed to this work as part of his dissertation.

Suyeon Kang is an Assistant Professor in the Department of Statistics and Data Science at the University of Central Florida. Her research interests include causal inference, sensitivity analysis, and multivariate response regression. She received her PhD from the Department of Statistics at the University of California, Riverside, in 2022. She contributed to this work while she was a postdoctoral fellow in the Department of Statistics at the University of Florida.

Alexander Franks is an Associate Professor in the Department of Statistics and Applied Probability at the University of California, Santa Barbara. His research interests include covariance estimation, sensitivity analysis and causal inference, missing data and measurement error, high-throughput applications in biology, Bayesian statistics, and sports. He received his PhD from the Department of Statistics at Harvard University in 2015.

Michelle Audirac is a Senior Programmer at the Harvard T.H. Chan School of Public Health. She holds an MS in data science from the Instituto Tecnológico Autónomo de México. Her interests lie in developing data science tools and standards for open science.

Danielle Braun is a principal research scientist and director of data science for environmental and climate health in the department of biostatistics at the Harvard T.H. Chan School of Public Health. She also co-leads the BayesMendel lab. Her research focuses on risk prediction, genetic epidemiology, measurement error, survival analysis, frailty models, clinical tool development, causal inference, and comparative effectiveness research. She received her PhD from the Department of Biostatistics at Harvard University in 2013.

OTHER PUBLICATIONS RESULTING FROM THIS RESEARCH

Antonelli J, Zigler C. 2024. Causal analysis of air pollution mixtures: estimands, positivity, and extrapolation. Am J Epidemiol 193:1392–1398, *https://doi.org/10.1093/aje/kwae115*.

Jin Y, Molstad A, Wilson A, Antonelli J. 2025. Smooth and shape-constrained quantile distributed lag models. Biometrics 81:ujaf101, *https://doi.org/10.1093/biomtc/ujaf101*.

Kang S, Franks A, Audirac M, Braun D, Antonelli J. 2023. Partial identification and unmeasured confounding with multiple treatments and multiple outcomes. arXiv preprint arXiv:2311.12252, *https://doi.org/10.48550/arXiv.2311.12252*.

McGee G, Antonelli J. 2025. A general approach to modeling environmental mixtures with multivariate outcomes. arXiv preprint arXiv:2504.17195, *https://doi.org/10.48550/arXiv.2504.17195*.

Samanta S, Antonelli J. 2022. Estimation and false discovery control for the analysis of environmental mixtures. Biostatistics 23:1039–1055, *https://doi.org/10.1093/biostatistics/kxac001*.

Shin H, Braun D, Irene K, Audirac M, Antonelli J. 2023. A spatial interference approach to account for mobility in air pollution studies with multivariate continuous treatments. arXiv preprint arXiv:2305.14194, *https://doi.org/10.48550/arXiv.2305.14194*.

Shin H, Linero A, Audirac M, Irene K, Braun D, Antonelli J. 2024. Treatment effect heterogeneity and importance measures for multivariate continuous treatments. arXiv preprint arXiv:2404.09126, *https://doi.org/10.48550/arXiv.2404.09126*.

Research Report 234, *Robust Statistical Approaches to Understanding the Causal Effect of Air Pollution Mixtures*, by J. Antonelli et al.

INTRODUCTION

Ambient air pollution is a major global public health risk factor. There is now a broad expert consensus that exposure to ambient air pollution causes an array of adverse health effects, which is based on evidence from a large body of scientific literature that has grown exponentially since the mid-1990s.[1-5] Most of our knowledge about the health risks of exposure to air pollution comes from observational studies that have examined the effects of individual pollutants, such as fine particulate matter ≤2.5 µm in aerodynamic diameter ($PM_{2.5}$)* and nitrogen dioxide (NO_2). Challenges remain, therefore, with estimating and quantifying the health risks associated with exposures to the complex mixtures of pollutants to which people are exposed daily.

With the goal of developing new statistical approaches for addressing challenging questions about the link between air pollution mixtures and health, Dr. Joseph Antonelli of the University of Florida submitted an application titled "Robust Statistical Approaches to Understanding the Causal Effect of Air Pollution Mixtures" in response to HEI's Request for Applications *19–2*: Walter A. Rosenblith New Investigator Award. This award was established to provide support for outstanding new investigators at the assistant professor level to conduct research in the area of air pollution and health; it is unrestricted with respect to the specific research topic. Antonelli proposed to develop statistical methods to study air pollution mixtures in epidemiological studies that would focus on causal inference-based techniques, including methods to account for changes in individual exposures across space. These methods would be applied to a Medicare (i.e., US health program for adults 65+ years old) cohort with existing national air pollution exposure estimates to demonstrate proof-of-concept in a real-world setting.

HEI's Research Committee recommended funding Antonelli's application because it believed that the proposed methods would help address important limitations of studies assessing multipollutant mixtures and could be broadly applicable to environmental epidemiology studies beyond air pollution. The study started in 2021.

This Commentary provides the HEI Review Committee's independent evaluation of the study. It is intended to aid the sponsors of HEI and the public by highlighting both the strengths and limitations of the study and by placing the results presented in the Investigators' Report into a broader scientific and regulatory context.

SCIENTIFIC AND REGULATORY BACKGROUND

Recent advances in the development of statistical methods have enabled researchers to answer questions regarding causality — such as whether exposure to outdoor air pollution is causally linked to mortality — using epidemiological datasets.[6] HEI has long played an important role in supporting research into the development of causal inference approaches,[7–9] as well as into multipollutant statistical approaches.[10–13] Generally speaking, causal models aim to describe the distribution of health outcomes that would be observed if individuals were assigned to hypothetical alternative exposure values rather than to their actual observed exposure values.[14] For example, causal models typically address counterfactual or "what if" questions, such as, "What would happen to the mortality rate in a population if we were to decrease all exposures to a given pollutant by one unit?" Causal inference, therefore, is the process of inferring causal effects from such hypothetical experiments with data.

The primary challenge to conducting causal inference analyses, however, is satisfying specific sufficient conditions for estimating causal effects from observational data. The *positivity* assumption, for example, requires that every value of exposure potentially of interest is possible for all kinds of individuals.[15] This condition means that for every combination of covariates (e.g., male, within a given age range), all exposure concentrations of interest may actually be observed in the data. The condition would be violated if, for example, no men in the study dataset who were 40–45 years of age were exposed to a given air pollution concentration of interest, as it would then not be possible to learn about the risk of the outcome corresponding to that exposure concentration among those men.

Dr. Joseph Antonelli's 3-year study, "Robust Statistical Approaches to Understanding the Causal Effect of Air Pollution Mixtures," began in June 2021. Total expenditures were $484,105. The draft Investigators' Report from Antonelli and colleagues was received for review in June 2024. A revised report, received in December 2024, was accepted for publication in December 2024. During the review process, the HEI Review Committee and the investigators had the opportunity to exchange comments and clarify issues in the Investigators' Report and the Commentary.

This report has not been reviewed by public or private party institutions, including those that support the Health Effects Institute, and may not reflect the views of these parties; thus, no endorsements by them should be inferred.

* A list of abbreviations and other terms appears at the end of this volume.

Model misspecification is another potential problem that can occur in causal (and "traditional" epidemiological) models. Misspecification occurs when the statistical model used to estimate causal effects does not correctly represent the true relationship between variables. This problem can occur through incorrect functional forms of the statistical model (e.g., assuming linearity when the relationship is nonlinear), excluding important confounders, or using incorrect interaction terms. Model misspecification might lead to biased effect estimates (e.g., incorrect magnitude-of-effect estimates, wrong signs, or spurious significance) because the misspecified model systematically over- or under-adjusts for confounding, thereby attributing other sources of variation in the outcome to the exposure. Another challenge to conducting epidemiological studies of outdoor air pollution exposure and health — including both causal and traditional approaches — is the difficulty of accurately assigning estimates of long-term exposures to study participants and minimizing the influence of exposure *measurement error* on any estimated health risks. Measurement error in an explanatory variable (e.g., an estimate of exposure) can bias health effect estimates (most often toward a null association) in epidemiological analyses.[12–14] It is important, therefore, for epidemiological studies of outdoor air pollution to estimate exposure as accurately as possible so that they can better inform our understanding of health risks and inform the regulatory process.

The *exchangeability* (or no residual confounding) assumption is an additional challenge for causal models that require that no unmeasured variables simultaneously affect both exposure assignment and the outcome. This assumption can fail to be met if individuals with different exposure levels differ with respect to unmeasured baseline characteristics that are predictive of the risk for the outcome. A lack of information on those characteristics (i.e., confounders) could limit the researcher's ability to estimate accurately the true risk of the outcome that occurs as a result of the exposure. In other words, for the exchangeability assumption to be met, all relevant confounders must be measured accurately and controlled for in the analysis.

All the assumptions and concerns noted typically cannot be tested empirically, given that causal inference involves comparing unobserved counterfactual outcomes that would have happened under hypothetical circumstances. As such, investigators need to justify their approaches using theory and existing evidence about the real-world processes under study.[15] Antonelli and colleagues aimed to lower the barriers to conducting causal inference analyses with multipollutant mixtures by developing and justifying approaches that address these challenges.

STUDY OBJECTIVES

Antonelli and colleagues proposed to develop causal inference methods to address four specific challenges associated with conducting epidemiological analyses. Their focus was on proposing new methods to improve the analysis of environmental mixtures, specifically mixtures of multiple air pollutants. They identified the following aims:

1. Develop an approach to identify the effect of exposure to a given pollutant on health when multiple pollutants are present and when limited data are available about the distributions of these pollutants. This aim seeks to address potential violations of the positivity assumption.
2. Develop a method to estimate the effects of exposure to a mixture of pollutants on the risk of mortality. This aim seeks to identify subgroups of the population that are most susceptible to the effects of exposure to air pollution and to reduce the detrimental impacts of model misspecification.
3. Evaluate the effect of controlling for daily mobility patterns on estimates of the health effects associated with long-term exposures to air pollution. This aim seeks to minimize exposure measurement error.
4. Develop a method to account for bias caused by residual confounding in large epidemiological analyses. This aim seeks to address the exchangeability assumption.

The proposed aims advance the field by highlighting the challenges inherent in studying the health effects of multiple exposures and expanding the range of research questions that can be answered regarding air pollution mixtures. The investigators' overarching goal was to enhance the policy relevance of findings from epidemiological studies.

In some cases, the new approaches and statistical models are described for theoretical applications, and in others, Antonelli and colleagues apply their methods to real-world data using a nationwide study of air pollution and health in the US Medicare population. Those analyses are based on the health information from over 30 million Medicare beneficiaries living in about 30,000 zip codes during the years 2000–2016. For air pollution data, the investigators used annual estimates of $PM_{2.5}$, black carbon, ammonium, nitrates, organic matter, sulfate, ozone, elemental carbon, and organic carbon from existing spatiotemporal models, all at a spatial resolution of about 1 km × 1 km,[16,17] aggregated to the zip code.

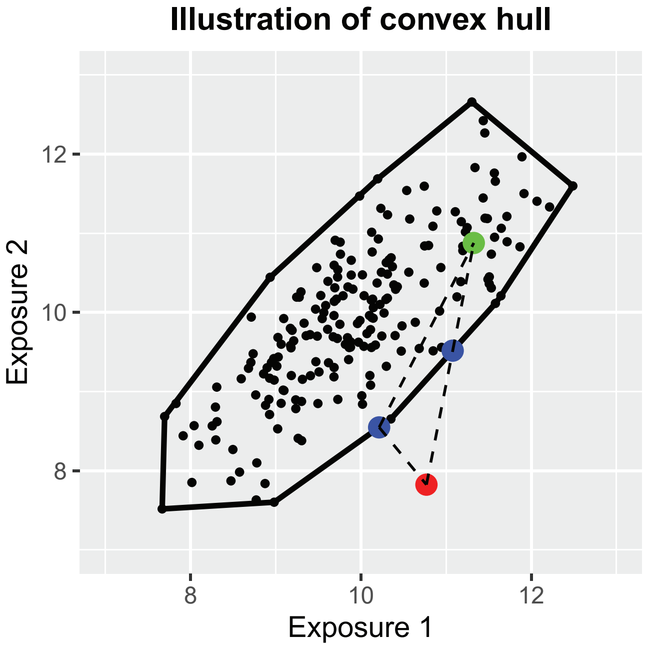
SUMMARY OF ANALYSES

AIM 1: CAUSAL ANALYSIS OF AIR POLLUTION MIXTURES: ESTIMANDS, POSITIVITY, AND MODEL EXTRAPOLATION

The first aim of the study addresses a key challenge unique to analyzing air pollution mixtures — one that cannot be resolved by simply extending univariate methods. It focuses on a) providing data-driven tools to assess whether a given dataset of study participants with multiple exposure estimates can support reliable causal effect estimation, and b) defining causal estimands that are appropriate for such datasets.

A causal estimand is the specific scientific causal quantity of interest. For example, a causal estimand might represent the expected change in risk of mortality resulting from a specified reduction in all pollutant exposure concentrations for all persons in the population. The investigators explain how the task of defining causal estimands for air pollution mixtures can lead to violating the positivity assumption. The challenge of meeting that assumption is that one might want to answer questions about circumstances for which real data are either not available or not observed (i.e., one might want to know about effects on health associated with exposures outside the range of observed concentrations). In such a situation, a researcher risks extrapolating findings to data points that do not exist. Note that whereas statistical *interpolation* involves estimating an unknown value that falls within the range of known values, *extrapolation* involves estimating an unknown value that falls outside that range. As an example, there is no empirical evidence for effects on health from annual mean exposures to concentrations of $PM_{2.5}$ of 1 µg/m^3, and it would therefore require model extrapolation to estimate the effect of an intervention that reduces annual mean exposures from, for example, 10 µg/m^3 to 1 µg/m^3. It becomes more complicated to avoid violating the positivity assumption when one is considering more than one exposure, because in such cases, there are more combinations of exposures that might not exist in a given dataset or in the real world more generally. For example, one might not be able to estimate a reduction in the concentration of one pollutant while maintaining another pollutant at the same concentration without extrapolation because both pollutants might come from a common source, like tailpipe emissions. Such violations are common in multipollutant studies and can lead to biased effect estimates (that is, estimates that are systematically too high or too low).

In this context, Antonelli and colleagues proposed a data-driven approach to detect violations of the positivity assumption in multipollutant settings. Their solution was to identify the distribution of exposure values that fall within what is known as the “convex hull.” In the case of a single variable, the convex hull would comprise the range of values bound by the minimum and maximum observed values of that variable. In the case of two variables (e.g., multipollutant models), the convex hull would be defined as the smallest polygon that contains all of the points of observed exposure among all of the exposure variables[18] (**Commentary Figure 1**).



Commentary Figure 1. Illustration of a convex hull around a set of observed values for two hypothetical exposures. The green point is an observed exposure of interest. The red point is the corresponding interventional point of interest. The blue points represent two different feasible values of the exposure that fall in the convex hull. Source: Investigators’ Report, Figure 4.

The investigators then proposed strategies to redefine causal estimands in a way that respects the limitations of the given data. Specifically, they proposed that if inspection of the convex hull indicated that estimating a value of interest would require model extrapolation (e.g., an intervention that reduces the exposure value of the green point to that of the red point in Commentary Figure 1), then alternatively defined estimands that were defined based on the convex hull could be substituted. One proposed approach would be to find a “feasible” concentration of exposure that is similar to the value of interest (i.e., the value at the red point) but that is not subject to model extrapolation (e.g., one of the two blue points in Commentary Figure 1). A second proposed approach would be to restrict the analysis to the subset of exposure values for which observed concentrations are available.

Notably, the investigators acknowledge that these approaches for modifying causal estimands impose a compromise between the feasibility of analysis and the interpretability of the results. That is, although it is feasible to estimate the modified causal estimands given that they are based on an observed sample of exposures (i.e., no extrapolation is required), the modified estimands might lose relevance and interpretability if the original estimands had been determined for scientific or policy reasons. In summary, the investigators have shown that their approaches provided feasible solutions for avoiding violation of the positivity assumption and model extrapolation in multipollutant contexts.

AIM 2: ADDRESSING THE HETEROGENEITY OF THE HEALTH EFFECTS OF AIR POLLUTION MIXTURES

Current methods for studying exposure to mixtures of pollutants assume that the effects on health are the same for everyone. But this assumption is not likely to hold, especially in air pollution research in which health effects can differ in relation to factors such as age, sex, or income.

To address this issue, Antonelli and colleagues introduced a flexible method using Bayesian statistics that reduces the detrimental effects of model misspecification and allows one to better understand how air pollution mixtures affect different groups of people. Their statistical models allow for nonlinear associations between exposures and health outcomes, for interactions between exposures, and also for the effect of the pollutant mixture to depend on other covariates in the model. Generally, their new approach allows one to identify and recognize how much the causal effect of a mixture of exposures varies with respect to measured covariates included in the model. The approach, therefore, allows one to identify, for example, what characteristic influences the magnitude of the effect of the mixture the most. By extension, one can infer which population subgroups are most affected by exposure to air pollution or who might benefit most from an intervention to reduce exposures.

AIM 3: ACCOUNTING FOR MOBILITY WHEN ESTIMATING THE HEALTH EFFECTS OF AIR POLLUTION MIXTURES

One of the major challenges in conducting epidemiological studies on outdoor air pollution exposure and health is the difficulty of accurately assigning estimates of long-term exposures to study participants. Estimating exposures is challenging because pollution concentrations can vary substantially over short distances and over time, and because people move around their neighborhoods, cities, and beyond on a daily and yearly basis. Nonetheless, most epidemiological studies that investigate long-term effects of air pollution on health assign concentrations at participants' residential addresses as their estimate of exposure. Whether that approach leads to exposure measurement error that can, in turn, affect the health effects estimation remains unclear. Some studies have shown that using exposures assigned only at residential locations and ignoring daily mobility patterns can lead to exposure measurement error,[19,20] which in turn can bias health effect estimates, most often toward a null association (i.e., showing no effect of the exposure on the health outcome).[21–23] Others, however, have found that assessing long-term air pollution exposure at only the residential address does not lead to substantial bias nor to loss of precision in health effects estimates.[24–26]

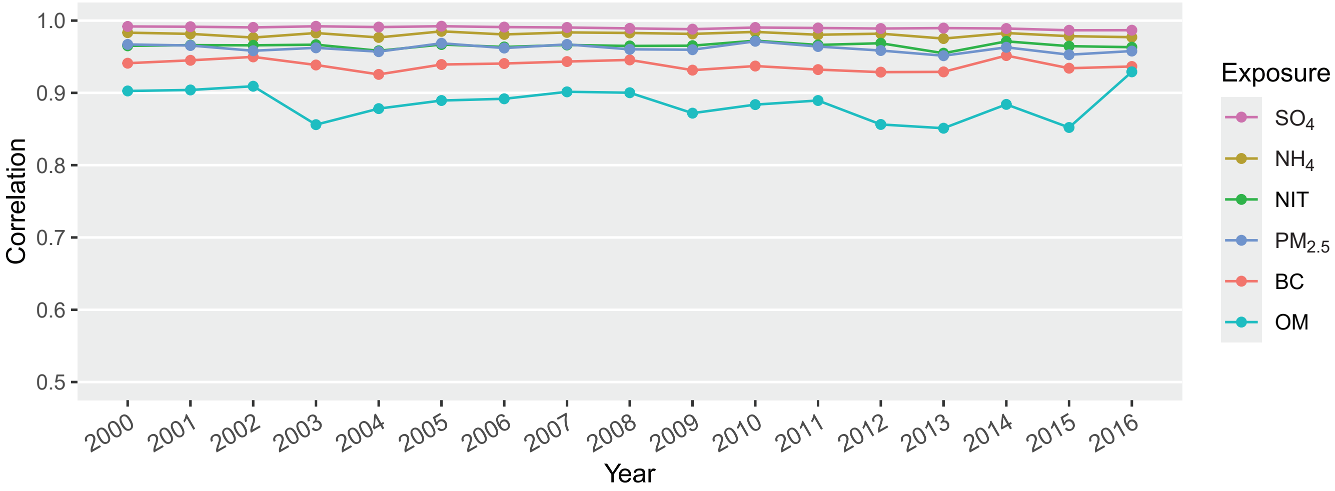
Antonelli and colleagues address the problem of potential exposure measurement error within a causal inference framework by exploring how bias might arise when mobility is ignored in estimating exposures to air pollution, especially in settings with multiple pollutants. They also present methods to correct such bias by using population-level mobility estimates derived from cell phone data. They applied their approaches in national-level analyses on the Medicare cohort mentioned earlier. For this specific analysis, the investigators acquired aggregated mobility data for the year 2019 that were collected from about 15 million anonymous cell phone users by Cuebiq, a location intelligence platform. Cuebiq compiles data from cell phone users across the contiguous United States who have opted in to provide access to their GPS location data anonymously.

The zip code of residence is the only residential information available for Medicare cohort participants. Zip codes vary in size in relation to population density and can cover a neighborhood in dense urban areas or represent an entire town, community, or larger area. For example, zip codes are on average 24 km² in Los Angeles County, California, and 268 km² in the state of Texas. Many previous epidemiological studies using the Medicare cohort have assigned concentrations of air pollution at the residential zip code only.[27–30] Here, Antonelli and colleagues derived two estimates of exposure to several pollutants for cohort participants:

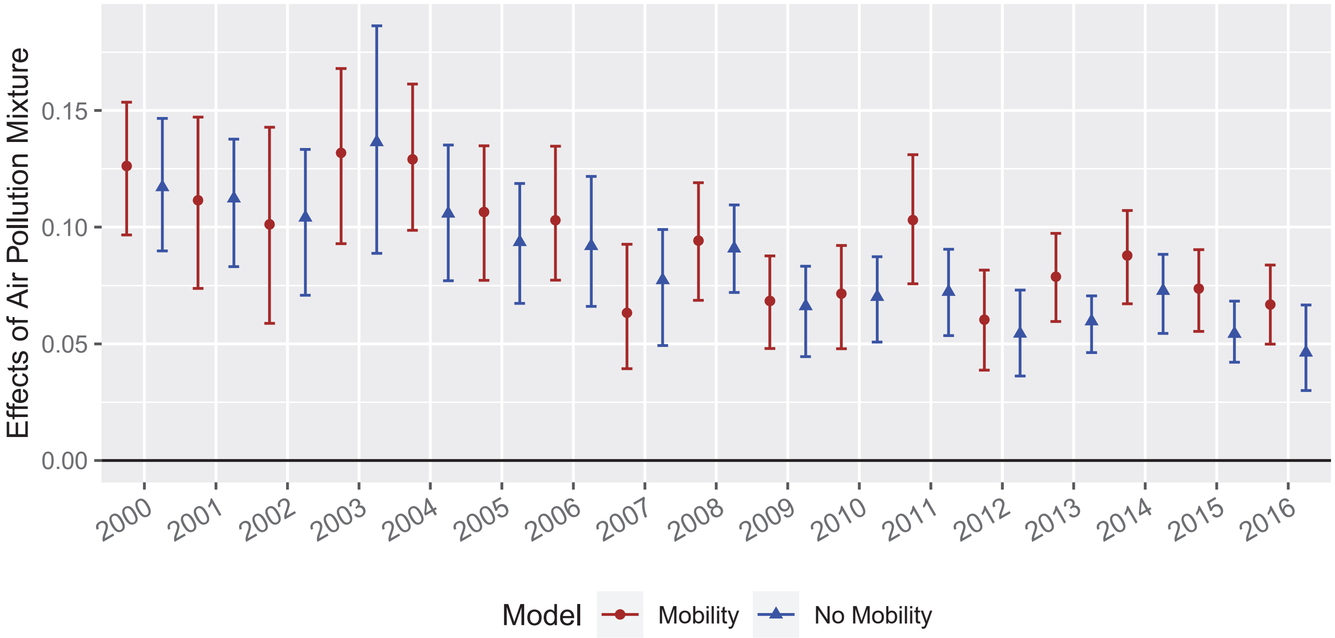
1. An annual mean concentration averaged across participants' home zip code (i.e., the conventional approach of residential-only exposure).
2. An annual mean concentration that was weighted by the amount of time residents of each zip code (as estimated via the Cuebiq data) spent in their home zip code and by the amount of time they spent in every other zip code. For example, for a given zip code, its residents might spend 70% of their time in that zip code, 20% of their time in a single adjacent zip code, and the remaining 10% in several other nearby zip codes.

The results showed that the percentage of time spent in the home zip code for residents of each state (plus Washington, DC) was generally similar, with most people spending about 78% of their time in their home zip code (Investigators' Report, Figure 13). As such, people did appear to spend a fair amount of time in other zip codes. **Commentary Figure 2** shows the correlations between the two estimates of exposures, specifically home zip code only and the mobility-based exposures, for several pollutants across time. In all cases, correlations were generally high ($r > 0.85$) between the two estimates. The annual correlations between the two estimates of sulfates were the highest, indicating very little spatial variability in that pollutant among the home and nearby zip codes.

Ultimately, although most people spent about 22% of their time in neighboring zip codes, incorporating mobility information at that scale did not lead to appreciably different estimates of exposures (Investigators' Report, Table 7), nor appreciably different health effect estimates (**Commentary Figure 3**). Commentary Figure 3 shows the estimated number of deaths per 100,000 person-years averaged over the study period associated with increasing each pollutant (black carbon, ammonium, sulfates, nitrates, organic matter, and fine particulate matter) simultaneously by 0.25 times their respective standard deviations. The figure shows the point estimates over the period 2000–2016 for models with (red symbols) and without (blue symbols) accounting for mobility. In all cases, increases in the pollutant mixture are associated with an increase of about 0.05 to 0.14 deaths per 100,000 person-years (with confidence intervals ranging from about 0.03 to 0.18 deaths). In most years of the analysis, the point estimates were slightly larger for the models that incorporated mobility, but in no cases was this difference statistically significant. The figure also shows a decreasing trend in the point



Commentary Figure 2. Correlation between home zip code and mobility-based exposures for six pollutants over the study period. Source: Investigators' Report, Figure 14.



Commentary Figure 3. Point estimates and 95% credible intervals for the effect of increasing exposures to black carbon, ammonium, nitrates, organic matter, fine particulate matter, and sulfate simultaneously by 0.25 times their respective standard deviations on number of deaths per 100,000 person-years, with and without accounting for mobility. Source: Investigators' Report, Figure 17.

estimates, which is due in part to decreasing air pollution concentrations over time. Because the concentrations are decreasing each year, so too are their standard deviations, and thus the point estimates for each subsequent year are calculated per increasingly smaller increments of exposure.

Overall, the investigators' results suggest that ignoring mobility can bias health effect estimates toward showing no effect of exposure, although the bias in this case was small because of the high correlations between the residential-only and mobility-based exposure estimates.

AIM 4: ADDRESSING RESIDUAL CONFOUNDING WITH MULTIPLE EXPOSURES AND MULTIPLE OUTCOMES

The fourth aim of Antonelli's study addresses the issue of residual confounding, which threatens the exchangeability assumption and can thus undermine the validity of findings from observational research on the health effects of air pollution. Here, the investigators proposed a framework designed to assess the magnitude of the potential bias from residual confounding that leverages information on multiple exposures *and* multiple outcomes simultaneously. The approach is based on the idea that having information about multiple outcomes for each study participant can be more informative about the nature of confounding than it is to have information about only one outcome in isolation. This expectation is reasonable in cases where the various exposures and outcomes share a common source of confounding. The investigators developed a "factor confounding assumption," which is less stringent and more realistic than assuming no residual confounding, to derive worst-case bounds on the potential bias caused by unmeasured or imperfectly measured confounders. Worst-case bounds are limits on the range within which the true causal effect might lie under minimal or no assumptions about residual confounding or other biases. The investigators also derived theoretical and computational tools for refining the bounds on the residual confounding bias when additional information is available.

Their approach to acquiring information about the unmeasured or imperfectly measured confounders is similar to a factor analysis that generates information about a latent (i.e., unmeasured) factor or construct. That is, the approach assumes that all the exposures and all the outcomes are correlated with some unmeasured or imperfectly measured confounder. The investigators acknowledge that although their approach would not typically be sufficient to correct for biases from all unmeasured variables completely, it can provide informative partial identification that can help identify the presence of health effects associated with exposure to air pollution. It is important to note that their approach works only for confounders that are known and that are not measured well and for unknown confounders (but known to share patterns with measured features); it does not work for confounders that are completely unrelated to measured variables of interest.

Overall, the approaches presented here can enhance the potential robustness of causality in air pollution research by acknowledging uncertainty from the presence of residual confounding.

HEI REVIEW COMMITTEE'S EVALUATION

This study evaluated several important challenges and limitations of observational studies and presented creative new causal inference-based solutions to them. In its independent evaluation of the Investigators' Report, the HEI Review Committee commended the investigators for developing original study aims and for tackling important issues for environmental epidemiology with which many statisticians continue to struggle. The Committee members agreed that the investigators did an excellent job setting the context and explaining the rationale for pursuing causal inference approaches to address multipollutant exposures. Although the report is dense with statistical formulas, Antonelli and colleagues did a commendable job explaining complicated concepts.

The strengths of this work are many and are related to the wide array of methods presented that advance the causal interpretation of environmental epidemiology studies. The focus on issues of positivity, model specification, measurement error, and residual confounding is important for informing policy. There is a need to address these issues generally, and especially in the modeling of air pollution mixtures. The work presented by Antonelli and colleagues includes a rigorous assessment of the assumptions required to draw causal conclusions when using complex mixture modeling. The Committee was impressed with the originality and innovative theoretical approaches for estimating causal effects of environmental mixtures presented.

The work represents strong first steps toward solving challenging problems, although additional complexities remain that will require continued methods development. For example, an overarching limitation of this study is that, in practice, the four challenges addressed by the investigators do not exist in isolation. The methods proposed, however, were presented independently, and addressing each challenge individually does not fully resolve the broader issues. It would be beneficial for future work to integrate multiple approaches in a shared framework. Another remaining challenge is that individual pollutants are not present in complete isolation from one another. Therefore, although this work aims to address causal questions to inform policy, it is not clear how one might actually intervene to address or reduce exposures to a mixture of measured and unmeasured pollutants. Policymakers can create regulations or policies that limit emissions of specific air pollution sources, which can, in turn, affect air pollution mixtures, although it is generally impossible to regulate mixtures directly. Finally, although the theoretical work is quite general, the application of the investigators' techniques was focused on the Medicare cohort. Although the Medicare cohort is a very large dataset that comprises nearly all older adults who live in the United States, it lacks key individual-level covariates related to socioeconomic status and health behaviors. Another limitation of this dataset is that it does not include residential addresses for cohort participants, and so exposures and outcomes were aggregated to zip codes, which might understate the effect of mobility on exposure. Although those features of the dataset might be problematic for the interpretation of the applied modeling results for each aim, they should not influence the rigor of the statistical methods proposed.

Committee members appreciated the originality of the approaches to detect and address violations of the positivity assumption (Aim 1). They noted that this approach has the added benefit of showcasing that the observed sample that might be used to define the convex hull is only a sample of a larger population. As such, it emphasizes the importance of how representative that sample is, and therefore, how generalizable the findings might be to the larger population or how transferable they might be to other datasets. Evaluations of the possibility of selection bias and a lack of generalizability are always prudent with epidemiological analyses, and their approach presents a helpful graphical representation of the limits of the data. One potential limitation of the approaches described to address Aim 1, however, is that they focus on situations without measured covariates. As such, it is not clear how the approaches might work in practice. Nonetheless, the Committee felt that their methods held potential and drew attention to an issue that is often overlooked.

The approach presented in Aim 2 sought to identify which covariates are the main drivers of the association between air pollution exposure and a given health outcome. The Committee acknowledged that identifying the covariates that modify the magnitude of an association most is a well-documented problem and noted that their approach has the potential to help identify groups that are especially sensitive or resilient to exposures. Overall, the Committee concluded that although the solution developed by the investigators was not perfect, it represented an interesting, original effort.

The Committee noted that the approaches using cell phone data developed to address Aim 3 were very creative and have the potential to eventually affect practice. As presented here, however, the Committee did not agree that the investigators had engaged sufficiently with practical data and examples to conclude the importance, or lack thereof, of mobility in understanding air pollution relationships. The authors acknowledged the limitations of the mobility data derived from cell phone use, in particular that these data represent only a random subset of individuals in each zip code and that the population from which these data were derived might not be representative of the population for which the investigators were trying to estimate health effects. Mobility within zip codes can also affect exposure patterns, especially for exposures that vary on a small scale. The analyses presented here are unable to capture such within-zip-code-level mobility or related variability in exposures.

An interesting observation related to the mobility analyses was that including mobility information led to the exposure estimates regressing toward the mean. That result might have occurred because those who lived in higher-exposure zip codes tended to travel to lower-exposure zip codes (thereby producing mean exposures that were lower than the residential-only estimates), and those who lived in lower-exposure zip codes tended to travel to higher-exposure zip codes (thereby producing mean exposures that were higher than the residential-only estimates). Ultimately, the Committee members found the findings from these analyses interesting, intuitive, and useful to an extent.

CONCLUSIONS

In summary, this study is among the first to address a series of major challenges faced regularly by researchers assessing the health effects of exposures to air pollution mixtures. As discussed by the investigators, previous efforts in statistical modeling in this context were not accompanied by a rigorous assessment of the assumptions required by the complex models needed to draw causal conclusions. The work presented by Antonelli and colleagues here is original and innovative and presents several interesting theoretical approaches that advance the field in estimating causal effects of environmental mixtures, although they do not completely resolve the important challenges described herein.

An overarching limitation of the study, however, relates to the applied research used to demonstrate the effectiveness of the methods. The theoretical foundations presented here were evaluated in a very specific dataset with unique features (i.e., older adults and data available only at the zip code level) that might prevent valid inferences for reasons beyond the issues addressed by the investigators. As such, the Committee advises that caution should be used in extrapolating the results of the applied analyses presented here to more general datasets. Although there is more work needed to resolve these complex issues fully, this study provides a strong foundation for future studies to extend the concepts described here to other settings.

ACKNOWLEDGMENTS

The HEI Review Committee thanks the ad hoc reviewers for their help in evaluating the scientific merit of the Investigators' Report. The Committee is also grateful to Eva Tanner for oversight of the study, to Dan Crouse for assistance with review of the report and in preparing its commentary, to Tara Hamilton for editing this report and its commentary, and to Kristin Eckles for her role in preparing this Research Report for publication.

REFERENCES

1. HEI Panel on the Health Effects of Long-Term Exposure to Traffic-Related Air Pollution. 2022. Systematic Review and Meta-analysis of Selected Health Effects of Long-Term Exposure to Traffic-Related Air Pollution. Special Report 23. Boston, MA: Health Effects Institute.
2. International Agency for Research on Cancer (IARC). 2016. Volume 109: Outdoor Air Pollution. Available: *https://www.iarc.who.int/news-events/volume-109-outdoor-air-pollution/*.
3. United States Environmental Protection Agency (US EPA). 2016. Integrated Science Assessment (ISA) for Oxides of Nitrogen — Health Criteria (Final Report, Jan 2016). Available: *https://assessments.epa.gov/isa/document/&deid=310879*.
4. United States Environmental Protection Agency (US EPA). 2019. Integrated Science Assessment (ISA) for Particulate Matter (Final Report, Dec 2019). Available: *https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=347534*.
5. World Health Organization (WHO). 2021. WHO Global Air Quality Guidelines: Particulate Matter ($PM_{2.5}$ and PM_{10}), Ozone, Nitrogen Dioxide, Sulfur Dioxide, and Carbon Monoxide. Available: *https://www.who.int/publications/i/item/9789240034228/*.
6. National Academies of Sciences, Engineering, and Medicine; Division on Engineering and Physical Sciences; Division on

Earth and Life Studies; Board on Mathematical Sciences and Analytics; Board on Environmental Studies and Toxicology; Committee on Assessing Causality from a Multidisciplinary Evidence Base for National Ambient Air Quality Standards. 2022. Advancing the Framework for Assessing Causality of Health and Welfare Effects to Inform National Ambient Air Quality Standard Reviews. Washington, DC: The National Academies Press. Available: *https://doi.org/10.17226/26612*.

7. Dominici F, Zanobetti A, Schwartz J, Braun D, Sabath B, Wu X. 2022. Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution: Implementation of Causal Inference Methods. Research Report 211. Boston, MA: Health Effects Institute.
8. Dominici F. 2004. Time-Series Analysis of Air Pollution and Mortality: A Statistical Review. Research Report 123. Boston, MA: Health Effects Institute.
9. Zigler CM, Kim C, Choirat C, Hansen JB, Wang Y, Hund L, et al. 2016. Causal Inference Methods for Estimating Long-Term Health Effects of Air Quality Regulations. Research Report 187. Boston, MA: Health Effects Institute.
10. Coull BA, Bobb JF, Wellenius GA, Kioumourtzoglou M-A, Mittleman MA, Koutrakis P, et al. 2015. Part 1. Statistical Learning Methods for the Effects of Multiple Air Pollution Constituents. In: Development of Statistical Methods for Multipollutant Research. Research Report 183. Boston, MA: Health Effects Institute.
11. Molitor J, Coker E, Jerrett M, Ritz B, Li A. 2016. Part 3. Modeling of Multipollutant Profiles and Spatially Varying Health Effects with Applications to Indicators of Adverse Birth Outcomes. In: Development of Statistical Methods for Multipollutant Research. Research Report 183. Boston, MA: Health Effects Institute.
12. Park ES, Symanski E, Han D, Spiegelman C. 2015. Part 2. Development of Enhanced Statistical Methods for Assessing Health Effects Associated with an Unknown Number of Major Sources of Multiple Air Pollutants. In: Development of Statistical Methods for Multipollutant Research. Research Report 183. Boston, MA: Health Effects Institute.
13. Sarnat JA, Russell A, Liang D, Moutinho JL, Golan R, Weber RJ, et al. 2018. Developing Multipollutant Exposure Indicators of Traffic Pollution: The Dorm Room Inhalation to Vehicle Emissions (DRIVE) Study. Research Report 196. Boston, MA: Health Effects Institute.
14. Fay MP, Li F. 2024. Causal interpretation of the hazard ratio in randomized clinical trials. Clin Trials 21:623–635, *https://doi.org/10.1177/17407745241243308*.
15. Igelström E, Igelström E, Craig P, Lewsey J, Lynch J, Pearce A, et al. 2022. Causal inference and effect estimation using observational data. J Epidemiol Community Health 76:960–966, *https://doi.org/10.1136/jech-2022-219267*.
16. Di Q, Amini H, Shi L, Kloog I, Silvern R, Kelly J, et al. 2019. An ensemble-based model of $PM_{2.5}$ concentration across the contiguous United States with high spatiotemporal resolution. Environ Int 130:104909, *https://doi.org/10.1016/j.envint.2019.104909*.
17. van Donkelaar A, Martin RV, Li C, Burnett RT. 2019. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. Environ Sci Technol 53:2595–2611, *https://doi.org/10.1021/acs.est.8b063*.
18. King G, Zeng L. 2006. The dangers of extreme counterfactuals. Political Analysis 14:131–159, *https://doi.org/10.1093/pan/mpj004*.
19. Nyhan MM, Kloog I, Britter R, Ratti C, Koutrakis P. 2019. Quantifying population exposure to air pollution using individual mobility patterns inferred from mobile phone data. J Expo Sci Environ Epidemiol 29:238–247, *https://doi.org/10.1038/s41370-018-0038-9*.
20. Yu X, Ivey C, Huang Z, Gurram S, Sivaraman V, Shen H, et al. 2020. Quantifying the impact of daily mobility on errors in air pollution exposure estimation using mobile phone location data. Environ Int 141:105772, *https://doi.org/10.1016/j.envint.2020.105772*.
21. Richmond-Bryant J, Long TC. 2020. Influence of exposure measurement errors on results from epidemiologic studies of different designs. J Expo Sci Environ Epidemiol 30:420–442, *https://doi.org/10.1038/s41370-019-0164-z*.
22. Szpiro AA, Paciorek CJ, Sheppard L. 2011. Does more accurate exposure prediction necessarily improve health effect estimates? Epidemiology 22:680–685, *https://doi.org/10.1097/ede.0b013e3182254cc6*.
23. Wei Y, Qiu X, Yazdi MD, Shtein A, Shi L, Yang J, et al. 2022. The impact of exposure measurement error on the estimated concentration-response relationship between long-term exposure to $PM_{2.5}$ and mortality. Environ Health Perspect 130:77006, *https://doi.org/10.1289/ehp10389*.
24. Christidis T, Pinault LL, Crouse DL, Tjepkema M. 2021. The influence of outdoor $PM_{2.5}$ concentration at workplace on nonaccidental mortality estimates in a Canadian census-based cohort. Environ Epidemiol 5:e180, *https://doi.org/10.1097/ee9.0000000000000180*.
25. Hoek G, Vienneau D, de Hoogh K. 2024. Does residential address-based exposure assessment for outdoor air pollution lead to bias in epidemiological studies? Environ Health 23:75, *https://doi.org/10.1186/s12940-024-01111-0*.
26. Ndiaye A, Vienneau D, Flückiger B, Probst-Hensch N, Jeong A, Imboden M, et al. 2025. Associations between long-term air pollution exposure and mortality and cardiovascular morbidity: a comparison of mobility-integrated and residential-only exposure assessment. Environ Int 198:109387, *https://doi.org/10.1016/j.envint.2025.109387*.
27. Di Q, Wang Y, Zanobetti A, Wang Y, Koutrakis P, Choirat C, et al. 2017. Air pollution and mortality in the Medicare population. N Engl J Med 376:2513–2522, *https://doi.org/10.1056/nejmoa1702747*.
28. Dominici F, Peng RD, Bell ML, Pham L, McDermott A, Zeger SL, et al. 2006. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. JAMA 295:1127–1134, *https://doi.org/10.1001/jama.295.10.1127*.
29. Wu X, Braun D, Schwartz J, Kioumourtzoglou MA, Dominici F. 2020. Evaluating the impact of long-term exposure to fine particulate matter on mortality among the elderly. Sci Adv 6:eaba5692, *https://doi.org/10.1126/sciadv.aba5692*.
30. Yitshak-Sade M, Nethery R, Schwartz JD, Mealli F, Dominici F, Di Q, et al. 2021. $PM_{2.5}$ and hospital admissions among Medicare enrollees with chronic debilitating brain disorders. Sci Total Environ 755(Pt 2):14252, *https://doi.org/10.1016/j.scitotenv.2020.142524*.

ABBREVIATIONS AND OTHER TERMS

ATE	average treatment effect
BART	Bayesian additive regression trees
BC	black carbon
BKMR	Bayesian kernel machine regression
BMI	body mass index
CCPA	California Consumer Privacy Act
COPD	chronic obstructive pulmonary disease
EC	elemental carbon
FAIR	findable, accessible, interoperable, and reusable
GDPR	General Data Protection Regulation
HIPAA	Health Insurance Portability and Accountability Act
MARS	multivariate adaptive regression splines
MCMC	Markov chain Monte Carlo
MTE-VIM	Multivariate treatment effect variable importance metric
NAAQS	National Ambient Air Quality Standards
NH_4	ammonium
NIT	nitrates
OM	organic matter
$PM_{2.5}$	particulate matter ≤ 2.5 µm in aerodynamic diameter
SO_4	sulfate
SUTVA	stable unit treatment value assumption

PDFs are available for free downloading at www.healtheffects.org/publications.

RELATED HEI PUBLICATIONS

Number	Title	Principal Investigator	Date
Research Reports			
227	Investigating the Consequences of Measurement Error of Gradually More Sophisticated Long-Term Personal Exposure Models in Assessing Health Effects: The London Study (MELONS)	K. Katsouyanni	2025
211	Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution: Implementation of Causal Inference Methods	F. Dominici	2022
187	Causal Inference Methods for Estimating Long-Term Health Effects of Air Quality Regulations	C. Zigler	2016
183	Development of Statistical Methods for Multipollutant Research		
	Part 1. Statistical Learning Methods for the Effects of Multiple Air Pollution Constituents	B.A. Coull	2015
	Part 2. Development of Enhanced Statistical Methods for Assessing Health Effects Associated with an Unknown Number of Major Sources of Multiple Air Pollutants	E.S. Park	2015
	Part 3. Modeling of Multipollutant Profiles and Spatially Varying Health Effects with Applications to Indicators of Adverse Birth Outcomes	J. Molitor	2016
94-1	National Morbidity, Mortality, and Air Pollution Study. Part I: Methods and Methodologic Issues	J. Samet	2000
86	Statistical Methods for Epidemiologic Studies of the Health Effects of Air Pollution	W. Navidi	1999

HEI BOARD, COMMITTEES, and STAFF

BOARD OF DIRECTORS

Richard A. Meserve, Chair *Senior of Counsel, Covington & Burling LLP; President Emeritus, Carnegie Institution for Science; former Chair, US Nuclear Regulatory Commission, USA*

Stephen Corman *President, Corman Enterprises, USA*

Martha J. Crawford *Operating Partner, Macquarie Asset Management, USA*

Ana V. Diez Roux *Dana and David Dornsife Dean and Distinguished University Professor of Epidemiology, Dornsife School of Public Health, Drexel University; Director, Drexel Urban Health Collaborative, USA*

Michael J. Klag *Dean Emeritus and Second Century Distinguished Professor, Johns Hopkins Bloomberg School of Public Health, USA*

Alan I. Leshner *CEO Emeritus, American Association for the Advancement of Science, USA*

Catherine L. Ross *Regents' Professor Emerita, City and Regional Planning and Civil and Environmental Engineering, Georgia Institute of Technology; Chairman of the Board of Directors of the Auto Club Group, American Automobile Association, USA*

Martha E. Rudolph *Environmental Attorney, Former Director of Environmental Programs, Colorado Department of Public Health and Environment, USA*

Karen C. Seto *Frederick C. Hixon Professor of Geography and Urbanization Science, Yale School of the Environment, Yale University, USA*

Jared L. Cohon *President Emeritus and Professor, Civil and Environmental Engineering and Engineering and Public Policy, Carnegie Mellon University, USA, In Memoriam 1947–2024*

RESEARCH COMMITTEE

David A. Savitz, Chair *Professor of Epidemiology, School of Public Health, and Professor of Obstetrics and Gynecology and Pediatrics, Alpert Medical School, Brown University, USA*

Benjamin Barratt *Professor, Environmental Research Group, School of Public Health, Imperial College London, United Kingdom*

David C. Dorman *Professor, Department of Molecular Biomedical Sciences, College of Veterinary Medicine, North Carolina State University, USA*

Christina H. Fuller *Associate Professor, School of Environmental, Civil, Agricultural and Mechanical Engineering, University of Georgia College of Engineering, USA*

Marianne Hatzopoulou *Professor, Civil and Mineral Engineering, University of Toronto, Research Chair in Transport Decarbonization and Air Quality, Canada*

Heather A. Holmes *Associate Professor, Department of Chemical Engineering, University of Utah, USA*

Neil Pearce *Professor of Epidemiology and Biostatistics, London School of Hygiene and Tropical Medicine, United Kingdom*

Evangelia (Evi) Samoli *Professor of Epidemiology and Medical Statistics, Department of Hygiene, Epidemiology and Medical Statistics, School of Medicine, National and Kapodistrian University of Athens, Greece*

Alexandra M. Schmidt *Professor of Biostatistics, School of Population and Global Health, McGill University, Canada*

Neeta Thakur *Associate Professor of Medicine, University of California San Francisco, USA*

Gregory Wellenius *Beverly A. Brown Professor for the Improvement of Urban Health and Director, Center for Climate and Health, Boston University School of Public Health, USA*

Continues next page

HEI BOARD, COMMITTEES, and STAFF

REVIEW COMMITTEE

Sara D. Adar *Professor of Epidemiology and Global Public Health, Department of Epidemiology, University of Michigan School of Public Health, USA*

Katherine B. Ensor *Noah G. Harding Professor of Statistics, Rice University, USA*

Ulrike Gehring *Associate Professor, Institute for Risk Assessment Sciences, Utrecht University, Netherlands*

Michael Jerrett *Professor, Department of Environmental Health Sciences, Fielding School of Public Health, University of California Los Angeles, USA*

John Volckens *Professor, Department of Mechanical Engineering, Walter Scott Jr. College of Engineering, Colorado State University, USA*

Scott Weichenthal *Professor, Department of Epidemiology, Biostatistics, and Occupational Health, School of Population and Global Health, McGill University, Canada*

STAFF AND CONSULTING SCIENTISTS

Elena Craft *President and CEO*
Ellen K. Mantus *Director of Science*
Donna J. Vorhees *Director of HEI Energy*
Thomas J. Champoux *Director of Science Communications*
Jacqueline C. Rutledge *Director of Finance and Administration*
Emily Alden *Corporate Secretary*
Daniel S. Greenbaum *President Emeritus, In Memoriam 1952–2024*
Robert M. O'Keefe *Vice President Emeritus*
Annemoon M. van Erp *Deputy Director of Science Emerita*

Amy Andreini *Science Communications Specialist*
Hanna Boogaard *Consulting Principal Scientist*
Jacki Collins *Senior Staff Accountant*
Dan Crouse *Senior Scientist*
Cloelle Danforth *Senior Scientist*
Philip J. DeMarco *Compliance Manager*
Kristin C. Eckles *Senior Editorial Manager*
Hlina Kiros *Research Assistant*
Lissa McBurney *Senior Science Administrator*
Samantha Miller *Research Assistant*
Victor Nthusi *Consulting Research Fellow*
Pallavi Pant *Head of Global Initiatives*
Allison P. Patton *Senior Scientist*
Yasmin Romitti *Staff Scientist*
Anna Rosofsky *Senior Scientist and Community Health and Environmental Research Initiatives Lead*

Continues next page

(*Staff and Consulting Scientists, continued*)

Abinaya Sekar *Consulting Research Fellow*
Robert Shavers *Operations Manager*
Eva Tanner *Staff Scientist*
Alexis Vaskas *Digital Communications Manager*

RESEARCH REPORT

NUMBER 234
DECEMBER 2025

Health Effects Institute

Health Effects Institute
75 Federal Street
Suite 1400
Boston, Massachusetts 02110, USA
+1-617-488-2300

www.healtheffects.org