HEI

RESEARCH REPORT

HEALTH EFFECTS INSTITUTE

Number 195
April 2018

Impacts of Regulations on Air Quality and Emergency Department Visits in the Atlanta Metropolitan Area, 1999–2013

Armistead (Ted) G. Russell, Paige Tolbert, Lucas R.F. Henneman, Joseph Abrams, Cong Liu, Mitchel Klein, James Mulholland, Stefanie Ebelt Sarnat, Yongtao Hu, Howard H. Chang, Talat Odman, Matthew J. Strickland, Huizhong Shen, and Abiola Lawal

Includes a Commentary by the Institute's Review Committee

Impacts of Regulations on Air Quality and Emergency Department Visits in the Atlanta Metropolitan Area, 1999–2013

Armistead (Ted) G. Russell, Paige Tolbert, Lucas R.F. Henneman, Joseph Abrams,
Cong Liu, Mitchel Klein, James Mulholland, Stefanie Ebelt Sarnat, Yongtao Hu,
Howard H. Chang, Talat Odman, Matthew J. Strickland, Huizhong Shen, and Abiola Lawal

with a Commentary by the HEI Review Committee

Research Report 195
Health Effects Institute
Boston, Massachusetts

Trusted Science • Cleaner Air • Better Health

Publishing history: This document was posted at *www.healtheffects.org* in April 2018.

Citation for document:

Russell AG, Tolbert P, Henneman LRF, Abrams J, Liu C, Klein M, et al. 2018. Impacts of Regulations on Air Quality and Emergency Department Visits in the Atlanta Metropolitan Area, 1999–2013. Research Report 195. Boston, MA:Health Effects Institute.

 Cameographics, Union, Me., Compositor. Printed by Recycled Paper Printing, Boston, Mass. Library of Congress Catalog Number for the HEI Report Series: WA 754 R432.

Cover paper: made with at least 55% recycled content, of which at least 30% is post-consumer waste; free of acid and elemental chlorine. Text paper: made with 100% post-consumer waste recycled content; acid free; no chlorine used in processing. The book is of permanent archival quality.

CONTENTS

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; and
- Communicates the results of HEI's research and analyses to public and private decision makers.

HEI typically receives balanced funding from the U.S. 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 340 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 260 comprehensive reports published by HEI, as well as in more than 1,000 articles in the 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, 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 are widely disseminated through HEI's website (*www.healtheffects.org*), printed reports, newsletters and other publications, annual conferences, and presentations to legislative bodies and public agencies.

ABOUT THIS REPORT

Research Report 195, *Impacts of Regulations on Air Quality and Emergency Department Visits in the Atlanta Metropolitan Area, 1999–2013*, presents a research project funded by the Health Effects Institute and conducted by Dr. Armistead (Ted) Russell of Georgia Institute of Technology, Atlanta, Georgia, U.S.A, and his 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 Committee's comments on the study.

The Investigators' Report, prepared by Russell 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 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. This draft report is first examined by outside technical reviewers and a biostatistician. The report and the reviewers' comments are then evaluated by members of the Review Committee, an independent panel of distinguished scientists who have no involvement in selecting or overseeing HEI studies. During the review process, the investigators have an opportunity to exchange comments with the Review Committee and, as necessary, to revise their report. The Commentary reflects the information provided in the final version of the report.

PREFACE

HEI's Accountability Research Program

INTRODUCTION

The goal of most air quality regulations is to protect the public's health by implementing regulatory actions or providing economic incentives that help reduce the public's exposure to air pollutants. If this goal is met, air pollution should be reduced, and indicators of public health should improve or at least not deteriorate. Evaluating the extent to which air quality regulations succeed in protecting public health is part of a broader effort — variously termed *accountability research*, *outcomes research*, or *research on regulatory effectiveness* — designed to assess the performance of environmental regulatory policies in general. In recent decades, air quality in the United States and Western Europe has improved substantially, and this improvement is attributable to a number of factors, including increasingly stringent air quality regulations. However, the cost of the pollution-control technologies and mechanisms needed to implement and enforce these regulations is often high. It is therefore prudent to ask whether the regulations have in fact yielded demonstrable improvements in public health, which will provide useful feedback to inform future efforts.

Several U.S. government agencies have concluded that direct evidence about the extent to which air quality regulations have improved health (measured as a decrease in premature mortality and excess morbidity) is lacking. This finding is well documented by the National Research Council (NRC) in its report *Estimating the Public Health Benefits of Proposed Air Pollution Regulations* (NRC 2002) and also has been made by the California Air Resources Board, the U.S. Environmental Protection Agency (EPA), the U.S. Centers for Disease Control and Prevention (CDC), and other agencies.

In 2003, the Health Effects Institute published a monograph on accountability research, Communication 11, *Assessing Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research* (HEI Accountability Working Group 2003). This monograph was written by the members of HEI's multidisciplinary Accountability Working Group after a 2001 workshop on the topic. Communication 11 set out a conceptual framework for accountability research and identified the types of evidence required and the methods by which the evidence should be obtained. It has also guided the development of the HEI Accountability Research program, which is discussed below.

Between 2002 and 2004, HEI issued four requests for applications (RFAs), under which eight studies were funded (see Table). A ninth study was funded later, under Request for Preliminary Applications (RFPA) 05-3, "Health Effects of Air Pollution." Following this first wave of research, HEI held further workshops to discuss lessons learned, identify key remaining questions, and plan a second wave of research. These efforts led to the publication of Communication 14 (van Erp and Cohen 2009) and Communication 15 (HEI 2010b), and the issuance of RFA 11-1, "Health Outcomes Research — Assessing the Health Outcomes of Air Quality Actions." The first wave of research primarily consisted of studies evaluating relatively short-term, local-scale, and sometimes temporary interventions; RFA 11-1 solicited additional studies with a focus on longer-term, regional- and national-scale regulations, including programs targeted at improving air quality surrounding major ports, as well as further methods development.

This preface describes both the framework of accountability research as it relates to air quality regulations and HEI's Accountability Research program.

HEI's Accountability Research Program		
RFA / Investigator (Institution)	Intervention	Study or Report Title
First-Wave Studies[a]		
RFA 02-1		
Douglas Dockery (Harvard T.H. Chan School of Public Health, Boston, MA)	Coal ban in Irish cities	Effect of Air Pollution Control on Mortality and Hospital Admissions in Ireland (Research Report 176; 2013)
Annette Peters (Helmholtz Zentrum München–German Research Center for Environment and Health, Neuherberg, Germany)	Switch from brown coal to natural gas for home heating and power plants, changes in motor vehicle fleet after reunification of Germany	The Influence of Improved Air Quality on Mortality Risks in Erfurt, Germany (Research Report 137; 2009)
RFA 04-1		
Frank Kelly (King's College London, U.K.)	Measures to reduce traffic congestion in the inner city of London	The Impact of the Congestion Charging Scheme on Air Quality in London: Part 1. Emissions Modeling and Analysis of Air Pollution Measurements. Part 2. Analysis of the Oxidative Potential of Particulate Matter (Research Report 155; 2011)
RFA 04-4		
Frank Kelly (King's College London, U.K.)	Measures to exclude most polluting vehicles from entering greater London	The London Low Emission Zone Baseline Study (Research Report 163; 2011)
Richard Morgenstern (Resources for the Future, Washington, DC)	Measures to reduce sulfur emissions from power plants east of the Mississippi River	Accountability Analysis of Title IV Phase 2 of the 1990 Clean Air Act Amendments (Research Report 168; 2012)
Curtis Noonan (University of Montana, Missoula, MT)	Wood stove change-out program	Assessing the Impact of a Wood Stove Replacement Program on Air Quality and Children's Health (Research Report 162; 2011)
Jennifer Peel (Colorado State University, Fort Collins, CO)	Measures to reduce traffic congestion during the Atlanta Olympics	Impact of Improved Air Quality During the 1996 Summer Olympic Games in Atlanta on Multiple Cardiovascular and Respiratory Outcomes (Research Report 148; 2010)
Chit-Ming Wong (University of Hong Kong)	Measures to reduce sulfur content in fuel for motor vehicles and power plants	Impact of the 1990 Hong Kong Legislation for Restriction on Sulfur Content in Fuel (Research Report 170; 2012)
RFPA 05-3		
Junfeng (Jim) Zhang (University of Medicine and Dentistry of New Jersey, Piscataway, NJ)	Measures to improve air quality during the Beijing Olympics	Cardiorespiratory Biomarker Responses in Healthy Young Adults to Drastic Air Quality Changes Surrounding the 2008 Beijing Olympics (Research Report 174; 2013)

Table continues next page

[a] Abbreviations: RFA, Request for Applications; RFPA, Request for Preliminary Applications.

HEI's Accountability Research Program (*continued*)		
RFA / Investigator (Institution)	Intervention	Study or Report Title
Second-Wave Studies [a]		
RFA 11-1		
Frank Gilliland (University of Southern California)	California and federal programs to improve air quality, including control of emissions from diesel engines and other sources targeted at freight transport and ports, as well as stationary sources	The Effects of Policy-Driven Air Quality Improvements on Children's Respiratory Health (Research Report 190; 2017)
Ying-Ying Meng (University of California–Los Angeles)	2006 California Emissions Reduction Plan for Ports and Goods Movement to control emissions from road, rail, and marine transportation, focusing on the ports of Los Angeles and Long Beach	Improvements in Air Quality and Health Outcomes Among California Medicaid Enrollees Due to Goods Movement Actions (Study ongoing)
Armistead Russell (Georgia Institute of Technology)	Programs to control emissions from major stationary sources and mobile sources in the Southeast United States	Impacts of Regulations on Air Quality and Emergency Department Visits in the Atlanta Metropolitan Area, 1999–2013 (Current report)
Corwin Zigler (Harvard T.H. Chan School of Public Health)	National regulations to improve air quality focusing on State Implementation Plans for particulate matter	Causal Inference Methods for Estimating Long-Term Health Effects of Air Quality Regulations (Research Report 187; 2016)

[a] Abbreviations: RFA, Request for Applications; RFPA, Request for Preliminary Applications.

BACKGROUND

The first step in assessing the effectiveness of air quality regulations is to measure emissions of the targeted pollutants to see whether they have in fact decreased as intended. A series of intermediate assessments, described in detail below, is needed to accurately measure the adverse health effects associated with air pollution to see whether their levels also decreased in incidence or severity relative to emissions. Some accountability studies to date have used hypothetical scenarios (comparing estimated outcomes under existing and more stringent regulations) and risk estimates obtained from epidemiological studies in an attempt to quantify past effects on health and to predict future effects (U.S. EPA 1999). However, more extensive validation of these estimates with data on actual outcomes would be helpful.

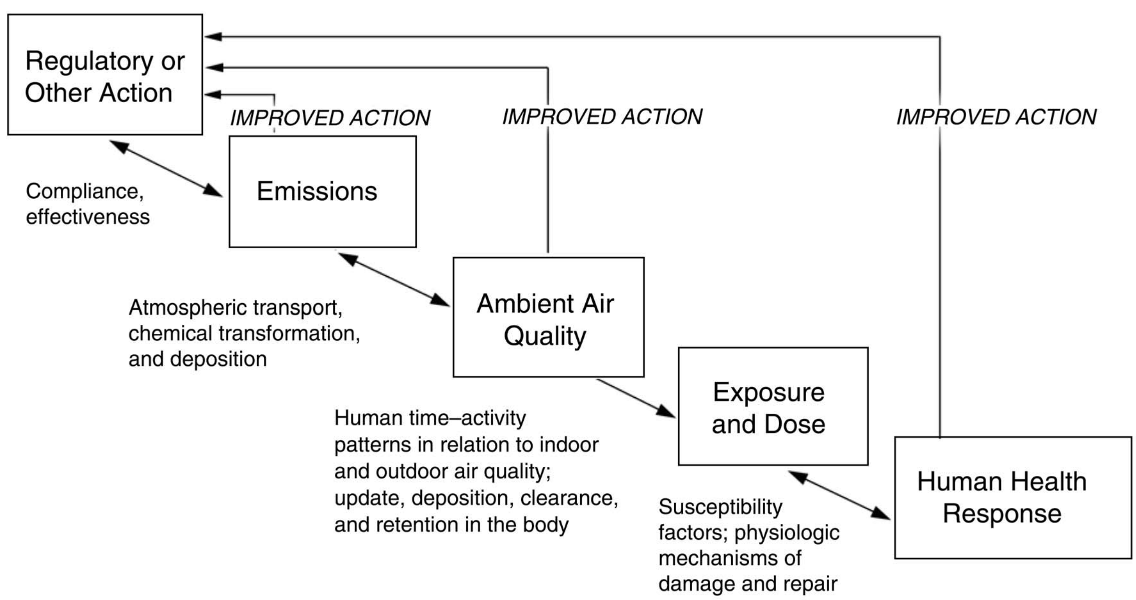
The long-term improvements in U.S. air quality have been associated with improved health in retrospective epidemiological studies (Chay and Greenstone 2003; Laden et al. 2006; Pope et al. 2009). Considerable challenges, however, are inherent in the assessment of the health effects of air quality regulations. Different regulations go into effect at different times, for example, and may be implemented at different levels of government (e.g., national, regional, or local). Their effectiveness therefore needs to be assessed in ways that take into account the varying times of implementation and levels of regulation. In addition, other changes at the same time and place might confound an apparent association between pollution reduction and improved health, such as economic trends (e.g., changes in employment), health care improvements, and behavioral changes (e.g., staying indoors when government warnings indicate pollution concentrations are high). Moreover, adverse health effects that might have been caused by exposure to air pollution can also be caused by other environmental risk factors (some of which may have changed over the same time periods as the air pollution concentrations). These challenges become more pronounced when regulations are implemented over long periods and when changes in air quality and health outcomes are not seen immediately, thus increasing the chance for confounding by other factors. For these reasons, scenarios in which regulations are expected to have resulted in rapid changes in air quality tend to be among the first, and most likely, targets for investigation, rather than evaluations of complex regulatory programs implemented over multiple years. Studies in Ireland by Clancy and colleagues (2002) and in Hong Kong by Hedley and colleagues (2002) are examples of such scenarios.

These inherent challenges are well documented in Communication 11 (HEI Accountability Working Group 2003), which was intended to advance the concept of accountability research and to foster the development of methods and studies throughout the relevant scientific and policy communities. In addition, recent advances in data collection and analytic techniques provide an unprecedented opportunity to improve assessments of the effects of air quality interventions.

THE ACCOUNTABILITY EVALUATION CYCLE

The NRC's Committee on Research Priorities for Airborne Particulate Matter set out a conceptual framework for linking air pollution sources to adverse health effects (NRC 1998). This framework can be used to identify factors along an "accountability evaluation cycle" (see Figure), each stage of which affords its own opportunities for making quantitative measurements of the intended improvements.

At the first stage (regulatory action), one can assess whether controls on source emissions have in fact been put into place. At the second stage (emissions), one can determine whether controls on sources have indeed reduced emissions, whether emitters have changed



Accountability Evaluation Cycle. Each box represents a stage in the process between regulatory action and human health responses to air pollution. Arrows connecting the stages indicate possible directions of influence. The text below the arrows identifies factors affecting the effectiveness of regulatory actions at each stage. At several of the stages, knowledge gained from studies on outcomes can provide valuable feedback for improving regulatory or other actions.

their practices, and whether there have been unintended consequences. At the third stage (ambient air quality), one can assess whether controls on sources and reductions in emissions have resulted in improved air quality. At the fourth stage (personal or population exposure), one can assess whether the improvement in air quality has reduced people's actual exposure and whether susceptible subpopulations (those most likely to experience adverse health effects) have benefited. At this stage, it is important to take into account changes in time–activity patterns that could either increase or reduce exposure. The actual dose that an individual's organs may be exposed to should also be considered (i.e., whether reductions in exposure have led to reductions in concentrations in body tissues such as the lung). Finally, at the fifth stage (human health response), one can assess whether risks to health have declined, given the evidence about changes in health outcomes such as morbidity and mortality that have resulted from changes in exposure. The challenge at this stage is to investigate the health outcomes that are most directly related to exposure to air pollution.

At each stage in the accountability evaluation cycle, the opportunity exists to collect evidence that either validates the assumptions that motivated the intervention or points to ways in which the assumptions were incorrect. The collection of such evidence can thus ensure that future interventions are maximally effective.

Ultimately, the framework for accountability research will need to encompass investigations of the broader consequences of regulations, not just the intended consequences. Unintended consequences should also be investigated, along with the possibility that risks to public health in fact increased, as discussed by Wiener (1998) and others who have advanced the concept of a portfolio of effects of a regulation.

HEI'S ACCOUNTABILITY RESEARCH PROGRAM

The first wave of HEI's Accountability Research program included nine studies (see Table). These studies involved the measurement of indicators along the entire accountability evaluation cycle, from regulatory or other interventions to human health outcomes. Many of the studies focused on interventions that were implemented over relatively short periods of time, such as a ban on the sale of coal, the replacement of old wood stoves with more efficient, cleaner ones, reductions in the sulfur content of fuels, and measures to reduce traffic. Other groups focused on longer-term, wider-ranging interventions or events; for instance, one study assessed complex changes associated with the reunification of the former East and West Germany, including a switch from brown coal to natural gas for fueling power plants and home-heating systems and an increase in the number of modern diesel-powered vehicles in eastern Germany. HEI also supported research, including the development of methods, in an especially challenging area, namely, assessment of the effects of regulations implemented incrementally over extended periods of time. In one such study, Morgenstern and colleagues (2012) examined changes that resulted from Title IV of the 1990 Clean Air Act Amendments (U.S. EPA 1990), which aimed at reducing sulfur dioxide emissions from power plants by requiring compliance with prescribed emission limitations. The first-wave studies are described in more detail in an interim evaluation of the HEI Accountability Research program (van Erp and Cohen 2009; van Erp et al. 2012).

Subsequently, HEI funded four studies as part of the second wave of HEI's Accountability program (see Table). Two studies evaluated regulatory and other actions at the national or regional level implemented over multiple years; a third study is evaluating complex sets of actions targeted at improving air quality in large urban areas and major ports with well-documented air quality problems and programs to address them; and a fourth study developed methods to support such accountability research. Gilliland and colleagues evaluated respiratory symptoms and lung function growth in children in Southern California from 1993–2012. They used data from three cohorts of the Children's Health Study, attempting to relate changes in health outcomes to major air quality regulations during that time period (Berhane et al. 2016; Gauderman et al. 2015; Gilliland et al. 2017; Lurmann et al. 2015). Russell and colleagues, as described in their Investigators' Report, assessed the effect of major stationary source and mobile source control programs on emissions and air quality in the southeastern United States, using detailed emissions and air pollution measurements and modeling combined with time-series analyses of cardiovascular and respiratory emergency department visits and hospital admissions in Atlanta. Meng and colleagues are evaluating the effects on air quality and health associated with the California Air Resources Board's Emission Reduction Plan for Ports and Goods Movement. They are examining the changes in criteria and hazardous air pollutants and characterizing health outcomes among Medicaid beneficiaries in the region surrounding the ports of Long Beach and Los Angeles. Phase 1, which focused on evaluating changes in air quality, has been completed (Su et al. 2016); Phase 2, to evaluate effects on health outcomes, is currently ongoing. Zigler and colleagues developed and applied statistical methods to evaluate long-term regulatory actions, focusing on the Clean Air Act and the role of attainment status of counties for PM_{10}, O_3, CO, and SO_2 concentrations. In particular, they focused on methods targeted on the question of whether air quality and health outcomes are causally related (Zigler and Dominici 2014; Zigler et al. 2016).

For an updated interim discussion of HEI's recent experiences in accountability research see Boogaard and colleagues (2017).

FUTURE DIRECTIONS

The second stage of accountability research was largely conceived during HEI's Strategic Plan for 2010 through 2015 (HEI 2010a). During the current Strategic Plan for 2015 through 2020 (HEI 2015), HEI continues to look closely at opportunities for unique new contributions to accountability research. We envision that future studies will again focus on large-scale, complex regulations to improve air quality and will continue to develop and implement statistical methods to tackle these complicated questions. In the interim, investigators who have identified a distinctive opportunity to evaluate the effects of environmental regulations on air pollution and human health are encouraged to contact HEI.

In addition, HEI continues to provide other researchers with access to extensive data and software from HEI-funded studies (see HEI's website, *www.healtheffects.org/research/databases*). In the same spirit, the recent State of Global Air website (HEI 2017) makes available data on air quality and health outcomes for countries around the world. The interactive site allows exploration of the data and comparisons among countries. The data currently cover 1990–2016 and are updated annually. State of Global Air 2018 is expected to be released in April 2018.

REFERENCES

Atmospheric and Environmental Research, Inc. (San Ramon, CA). HEI Air Quality Database. Available: *https://hei.aer. Com* [accessed 11 May 2016].

Berhane K, Chang CC, McConnell R, Gauderman WJ, Avol E, Rappaport E, et al. 2016. Association of changes in air quality with bronchitic symptoms in children in California, 1993–2012. JAMA 315(14):1491–1501.

Boogaard H, van Erp AM, Walker KD, Shaikh R. 2017. Accountability studies on air pollution and health: The HEI experience. Curr Environ Health Rep; 10.1007/s40572-017-0161-0.

Chay KY, Greenstone M. 2003. The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. Q J Econ 118:1121–1167.

Clancy L, Goodman P, Sinclair H, Dockery DW. 2002. Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. Lancet 360:1210–1214.

Gauderman WJ, Urman R, Avol E, Berhane K, McConnell R, Rappaport E, et al. 2015. Association of improved air quality with lung development in children. N Engl J Med 372(10):905–913.

Gilliland F, Avol E, McConnell R, Berhane K, Gauderman WJ, Lurmann FW, et al. 2017. The Effects of Policy-Driven Air Quality Improvements on Children's Respiratory Health. Research Report 190. Boston, MA:Health Effects Institute.

HEI Accountability Working Group. 2003. Assessing Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research. Communication 11. Boston, MA:Health Effects Institute.

Health Effects Institute. 2005. HEI Strategic Plan for Understanding Health Effects of Air Pollution 2005–2010. Boston, MA:Health Effects Institute.

Health Effects Institute. 2010a. HEI Strategic Plan for Understanding the Health Effects of Air Pollution 2010–2015. Boston, MA:Health Effects Institute.

Health Effects Institute. 2015. HEI Strategic Plan for Understanding the Health Effects of Air Pollution 2015–2020. Boston, MA:Health Effects Institute.

Health Effects Institute. 2010b. Proceedings of an HEI Workshop on Further Research to Assess the Health Impacts of Actions Taken to Improve Air Quality. Communication 15. Boston, MA:Health Effects Institute.

Health Effects Institute. 2017. State of Global Air 2017. Special Report. Boston, MA:Health Effects Institute.

Hedley AJ, Wong CM, Thach TQ, Ma S, Lam TH, Anderson HR. 2002. Cardiorespiratory and all-cause mortality after restrictions on sulphur content of fuel in Hong Kong: an intervention study. Lancet 360:1646–1652.

Johns Hopkins Bloomberg School of Public Health (Baltimore, MD). Internet-Based Health and Air Pollution Surveillance System (last updated 3/19/05). Available: *www.ihapss.jhsph.edu* [accessed 11 May 2016].

Kelly F, Anderson HR, Armstrong B, Atkinson R, Barratt B, Beevers S, et al. 2011a. Part 1. Emissions modeling and analysis of air pollution measurements. In: The Impact of the Congestion Charging Scheme on Air Quality in London. Research Report 155. Boston, MA:Health Effects Institute.

Kelly F, Anderson HR, Armstrong B, Atkinson R, Barratt B, Beevers S, et al. 2011b. Part 2. Analysis of the oxidative potential of particulate matter. In: The Impact of the Congestion Charging Scheme on Air Quality in London. Research Report 155. Boston, MA:Health Effects Institute.

Kelly FJ, Armstrong B, Atkinson R, Anderson HR, Barratt B, Beevers S, et al. 2011c. The London Low Emission Zone Baseline Study. Research Report 163. Boston, MA:Health Effects Institute.

Laden F, Schwartz J, Speizer FE, Dockery DW. 2006. Reduction in the particulate air pollution and mortality: extended follow-up of the Harvard Six Cities study. Am J Respir Crit Care Med 173:667–672.

Lurmann F, Avol E, Gilliland F. 2015. Emissions reduction policies and recent trends in Southern California's ambient air quality. J Air Waste Manag Assoc 65(3):324–335.

Matte TD, Cohen A, Dimmick F, Samet J, Sarnat J, Yip F, et al. 2009. Summary of the workshop on methodologies for environmental public health tracking of air pollution effects. Air Qual Atmos Health 2:177–184.

Morgenstern RD, Harrington W, Shih J-S, Bell ML. 2012. Accountability Analysis of Title IV Phase 2 of the 1990 Clean Air Act Amendments. Research Report 168. Boston, MA:Health Effects Institute.

National Research Council (U.S.). 1998. Research Priorities for Airborne Particulate Matter: I. Immediate Priorities and a Long-Range Research Portfolio. Washington, DC:National Academies Press.

National Research Council (U.S.). 2002. Estimating the Public Health Benefits of Proposed Air Pollution Regulations. Washington, DC:National Academies Press.

Noonan CW, Ward TJ, Navidi W, Sheppard L, Bergauff M, Palmer C. 2011. Assessing the Impact of a Wood Stove Replacement Program on Air Quality and Children's Health. Research Report 162. Boston, MA:Health Effects Institute.

Peel JL, Klein M, Flanders WD, Mulholland JA, Tolbert PE. 2010. Impact of Improved Air Quality During the 1996 Summer Olympic Games in Atlanta on Multiple Cardiovascular and Respiratory Outcomes. Research Report 148. Boston, MA:Health Effects Institute.

Peters A, Breitner S, Cyrys J, Stölzel M, Pitz M, Wölke G, et al. 2009. The Influence of Improved Air Quality on Mortality Risks in Erfurt, Germany. Research Report 137. Boston, MA:Health Effects Institute.

Pope CA III, Ezzati M, Dockery DW. 2009. Fine particulate air pollution and life expectancy in the United States. N Engl J Med 360:376–386.

Su JG, Meng YY, Pickett M, Seto E, Ritz B, Jerrett M. 2016. Identification of effects of regulatory actions on air quality in goods movement corridors in California. Environ Sci Technol 50(16):8687–8696.

U.S. Centers for Disease Control and Prevention. National Environmental Public Health Tracking Program. Available: *www.cdc.gov/nceh/tracking/* [accessed 11 May 2016].

U.S. Environmental Protection Agency. 1990. A Bill to Amend the Clean Air Act to Provide for Attainment and Maintenance of Health Protective National Ambient Air Quality Standards, and for Other Purposes. S 1630, 101st Cong, 2nd Sess.

U.S. Environmental Protection Agency. 1999. Benefits and Costs of the Clean Air Act 1990 to 2010: Report to Congress. EPA/410/R-99-001. Washington, DC:Office of Air and Radiation.

van Erp AM, Cohen AJ. 2009. HEI's Research Program on the Impact of Actions to Improve Air Quality: Interim Evaluation and Future Directions. Communication 14. Boston, MA:Health Effects Institute.

van Erp AM, Cohen AJ, Shaikh R, O'Keefe R. 2012. Recent progress and challenges in assessing the effectiveness of air quality interventions toward improving public health: the HEI experience. EM October: 22–28.

Wiener J. 1998. Managing the iatrogenic risks of risk management. Risk Health Safety Environ 9:39–82.

Wong C-M, Rabl A, Thach TQ, Chau YK, Chan KP, Cowling BJ, et al. 2012. Impact of the 1990 Hong Kong Legislation for Restriction on Sulfur Content in Fuel. Research Report 170. Boston, MA:Health Effects Institute.

Zeger SL, McDermott A, Dominici F, Peng R, Samet J. 2006. Internet-Based Health and Air Pollution Surveillance System. Communication 12. Boston, MA:Health Effects Institute.

Zhang J, Zhu T, Kipen H, Wang G, Huang W, Rich D, et al. 2013. Cardiorespiratory Biomarker Responses in Healthy Young Adults to Drastic Air Quality Changes Surrounding the 2008 Beijing Olympics. Research Report 174. Boston, MA:Health Effects Institute.

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(12):1133–1140.

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.

HEI STATEMENT

Synopsis of Research Report 195

Impacts of Regulations on Atlanta Air Quality and Emergency Department Visits

INTRODUCTION

Accountability research evaluates whether regulatory and other actions aimed at improving air quality have resulted in the intended decreases in air pollutant concentrations and improvements in public health. Such studies are complicated by the fact that simply comparing the changes in air pollution before and after an action may not capture what might have happened to air pollution in the absence of a regulation altogether.

A relatively recent approach to accountability research is to compare changes in air quality and health after the regulation went into effect with projected scenarios that estimate what the air quality and health outcomes would have been without the intervention. Dr. Ted Russell from the Georgia Institute of Technology and colleagues at the Georgia Institute of Technology and at Emory University proposed to examine whether national and state regulations targeting power plants and mobile sources were effective in reducing pollutant emissions, improving air quality, and ultimately reducing emergency department visits in the Atlanta area, using both measurements and modeling approaches.

APPROACH

Russell and colleagues identified major regulatory actions implemented between 1995 and 2010 and then assessed the effects of those regulations along the HEI chain of accountability by evaluating changes in emissions, effects of changes in emissions on air quality, and finally changes in air quality on emergency department visits for the period 1999–2013. The investigators estimated projected scenarios to compare what actually happened with what likely would have happened without the regulations. They focused on six sets of national- and state-level regulatory programs that they thought were likely to affect air pollutant emissions and air quality in Atlanta,

What This Study Adds

- This accountability study examined the extent to which national and state regulations targeting power plants and mobile sources were effective in reducing pollutant emissions, improving air quality, and ultimately reducing cardiorespiratory emergency department visits in the Atlanta area.
- Actual conditions in the period 1999–2013 were compared with estimated quantitative projections of emissions, air quality, and emergency department visits that likely would have occurred in the absence of regulations (called a “counterfactual scenario”).
- The study demonstrated that both the emissions and levels of all evaluated pollutants decreased by 14% to 91% over the study period. There were fewer emergency department visits for asthma and other cardiorespiratory outcomes than would have been expected without the regulations.
- Regulations targeting power plants appeared more effective in improving air quality than those targeting mobile sources. The HEI Review Committee had more confidence, however, in the results that were attributed to all regulations combined than to individual regulatory programs.
- This is one of few accountability studies to follow changes of individual regulations on emissions all the way through health outcomes, using scenarios with and without regulation. The approach is valuable and worth considering for future accountability studies.

This Statement, prepared by the Health Effects Institute, summarizes a research project funded by HEI and conducted by Dr. Armistead (Ted) G. Russell at the Georgia Institute of Technology, Atlanta, Georgia, and colleagues. Research Report 195 contains both the detailed Investigators’ Report and a Commentary on the study prepared by the Institute’s Review Committee.

Georgia: three national program sets to reduce emissions from power plants (electricity-generating units [EGUs]), and three program sets targeting motor vehicle fuel and emissions standards (mobile sources) adopted in response to national requirements.

To evaluate the effect of the regulations on emissions, the investigators used two approaches. First, they compared emissions before the regulations (from 1995 for power plants and 1993 for mobile sources) to emissions at the end of the study period (2013) for the southeastern United States. Second, because emissions could have changed for reasons unrelated to the regulations, they used daily records of how much electricity was generated by power plants and how far cars were driven in order to estimate how much higher the emissions at the end of the study period would have been if the regulations had not been implemented (called a "counterfactual scenario").

Similarly, to evaluate the effects of regulations on air quality, they compared measured levels of a large number of pollutants at a monitoring site near downtown Atlanta at the beginning of the study period (1999) with their levels at the end of the study period (2013). Because meteorology could affect the results, they adjusted the air quality measurements for the potential influence of daily meteorology. They again used a counterfactual scenario approach to project what the air pollutant levels would have been without the regulations and compared those projected levels with measured levels in order to estimate the effects of the emissions changes on air quality.

Finally, Russell and colleagues used time-series models to relate the daily numbers of Atlanta area emergency department visits to daily air pollutant levels for outcomes related to diseases of the heart (all cardiovascular disease and the subset from congestive heart failure) and lung (all respiratory disease and the subset from asthma). Following the counterfactual scenario approach, they compared actual numbers of emergency department visits with the numbers that likely would have occurred without the regulations. They presented results for the impact of each set of regulations, the three sets of regulations affecting power plants, the three sets of regulations affecting motor vehicles, and all six sets combined.

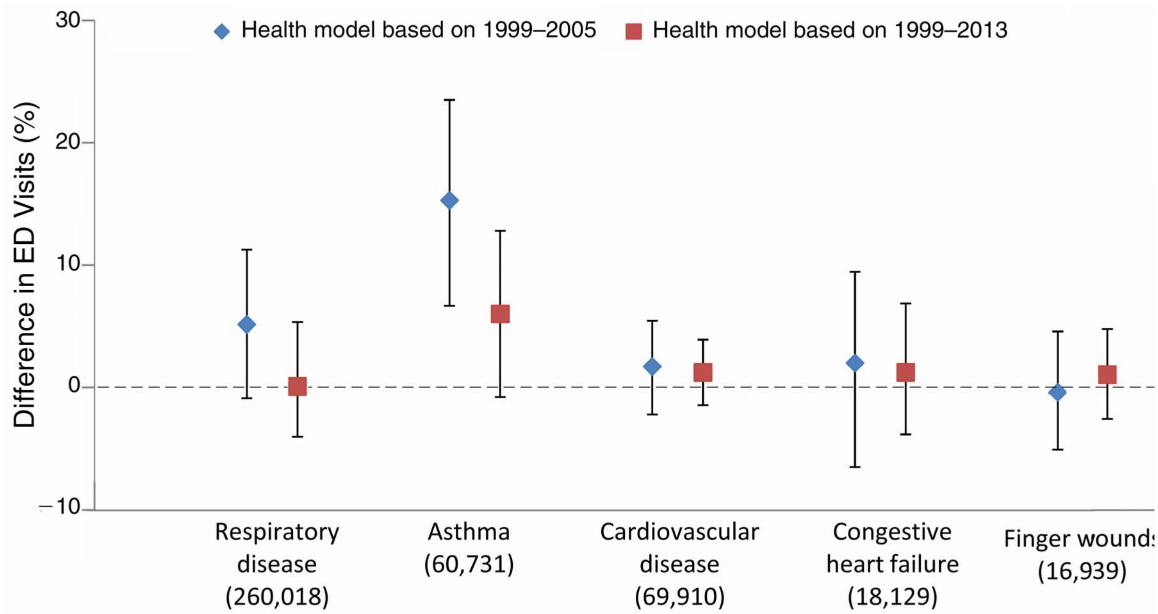
Unlike many other such studies, the uncertainty reported for the numbers of emergency department visits also included uncertainty carried forward from the emissions and air quality models, respectively. They also tested the effects of a number of assumptions on the results, including the number of pollutants (1, 5, 7, or 9) included in the health models; which years were considered when constructing the health models (1999–2005 vs. 1999–2013); and the size of the study area (5 or 20 counties in the Atlanta area).

MAIN RESULTS AND INTERPRETATION

The investigators reported that air pollutant emissions and ambient concentrations decreased over the study period 1999–2013 for most pollutants evaluated, and estimated that the pollutant levels were lower than what would have been expected without regulatory actions. Their modeling suggested that the observed improvements in air quality were associated with fewer emergency department visits for asthma and other lung outcomes compared with what would have been expected without the regulations (see Statement Figure). The health results were robust to the geographical scale of assessment (5 or 20 counties) and number of pollutants (1, 5, 7, or 9) in the health models. These results were less robust to the period evaluated. Estimates of effects of air pollutant changes on emergency department visits were larger for results with models of relationships between emergency department visits and air quality based on data from 1999 through 2005 than for models of relationships between emergency department visits and air quality based on data from 1999 through 2013. Although both analyses reported improvements in health with declining air pollution, the HEI Review Committee thought the differences in estimates for ED visits using data from two different periods suggested there was uncertainty that was not fully accounted for, and that perhaps the results for the two periods should be weighted more equally since it is not clear which health model was more appropriate.

The investigators also reported that regulations targeting power plants had a greater impact than those targeting mobile sources in improving air quality and health.

In its independent review of the report, the HEI Review Committee noted that the study was an ambitious application of HEI's accountability framework as it encompassed a broad suite of regulatory



Statement Figure. Estimated emergency department visits in Atlanta in 2012–2013 for all regulations combined compared with a scenario without the regulations for two different models. Positive numbers indicate there were fewer emergency department visits with regulatory programs in place. Whiskers represent the 95% confidence intervals. Actual numbers of emergency department visits are listed in parentheses.

programs designed to reduce multipollutant emissions from power plants and mobile sources in Georgia and nearby states over the period from 1999 to 2013. The Committee thought that the investigators had tackled an important public health question, examining whether the regulations had individually or collectively reduced emissions, improved air quality, and ultimately reduced ED visits for respiratory and cardiovascular outcomes in the Atlanta area.

The Review Committee concluded that the investigators had thoughtfully applied a counterfactual scenario approach to compare actual observations after the regulations were implemented to without-regulation scenarios. The study built on large and well-characterized data sets of air pollutant concentrations and emergency department visits from the Atlanta area. It addressed some concerns of earlier studies, such as the influence of meteorology on air quality. One of the difficulties encountered was that regulations were implemented in different years; the investigators handled this by comparing actual conditions to counterfactual conditions for each day of the study period. Together with the extensive sensitivity and uncertainty analyses in the development and application of the health models, these were all clear strengths of the study.

The Review Committee had the most confidence in the results for the link between changes in emissions and air quality because the investigators were able to rule out meteorology as an alternative explanation for the changes in air quality. The Committee thought that the link between regulations and emissions also appeared strong, although exploration of the potential effect on emissions of factors other than regulations, such as market-induced efficiency improvements, would have enhanced the analysis.

The Review Committee noted some limitations in the linkages between air quality and health effects (and therefore also in the estimates of changes in the numbers of emergency department visits). One of the strengths of the study is that it was conducted over a long period of time (i.e., 15 years); however, this leads to the possibility that potentially important factors that also changed over time were not fully captured, such as changes in healthcare access and practice. This could explain why the health models based on data from different periods yielded different results.

Overall, the Committee had more confidence in the results attributed to all regulations combined than to individual regulatory programs. The investigators' finding that regulations targeting power plants had more impact on improving air quality and reducing emergency department visits than regulations targeting mobile sources needs further study. Direct comparisons may not be appropriate because a single monitor would be more suitable to capture the regional impact of power plant regulations than the more spatially heterogeneous impact of mobile source regulations. In addition, measurements of mobile emissions were not available and some of the mobile source regulations did not go into effect until the later part of the study period (e.g., heavy-duty diesel rules targeting particulate matter and oxides of nitrogen emissions from new vehicles beginning in 2007 and 2010). The separation of attribution to different programs is inherently more difficult than linking overall emissions reductions to health outcomes, because it requires the separation of changes that overlapped in time. It is also possible that slow turnover of the vehicle fleet and lack of compliance may have hampered reaching full implementation of the fuel and technology changes by the end of the study period (2013), and further improvements may have continued since then. Thus, the ultimate effectiveness of mobile source regulations may actually be better — even if more gradual — than what the investigators were able to estimate in their study.

The Committee thought that this report by Russell and colleagues was a valuable addition to the accountability literature. It is one of few studies to follow changes of regulations on emissions all the way through to health outcomes, using scenarios based on actual observed data. This is a valuable approach worth considering for future accountability studies, though this sort of work is labor and computationally intensive. In addition, this work provides a detailed protocol for how to conduct similar investigations in other areas of the world.

In the future, other researchers could apply similar approaches to the long-term impacts of regulations on health outcomes in other locations, although it would be recommended to more thoroughly account for changes in medical practice and healthcare access, where possible. In particular, although efforts to disentangle the effects of specific regulations among a suite of regulations remain challenging, such efforts are important and should continue. This study is a strong and important contribution to HEI's accountability research portfolio because it sequentially and carefully addressed multiple links in the accountability chain. The results suggesting that reductions in emissions and improved air quality were linked to health benefits are important in terms of continued evaluation of the public health benefits of air pollution regulation in the context of implementation and compliance issues that may hamper achievement of the intended benefits.

INVESTIGATORS' REPORT

Impacts of Regulations on Air Quality and Emergency Department Visits in the Atlanta Metropolitan Area, 1999–2013

Armistead (Ted) G. Russell[1], Paige Tolbert[2], Lucas R.F. Henneman[1], Joseph Abrams[2], Cong Liu[1], Mitchel Klein[2], James Mulholland[1], Stefanie Ebelt Sarnat[2], Yongtao Hu[1], Howard H. Chang[2], Talat Odman[1], Matthew J. Strickland[2], Huizhong Shen[1], and Abiola Lawal[1]

[1]Georgia Institute of Technology; Atlanta, Georgia; [2]Emory University, Atlanta, Georgia

ABSTRACT

INTRODUCTION

The United States and Western Europe have seen great improvements in air quality, presumably in response to various regulations curtailing emissions from the broad range of sources that have contributed to local, regional, and global pollution. Such regulations, and the ensuing controls, however, have not come without costs, which are estimated at tens of billions of dollars per year. These costs motivate accountability-related questions such as, to what extent do regulations lead to emissions changes? More important, to what degree have the regulations provided the expected human health benefits?

Here, the impacts of specific regulations on both electricity generating unit (EGU*) and on-road mobile sources are examined through the classical accountability process laid out in the 2003 Health Effects Institute report linking regulations to emissions to air quality to health effects, with a focus on the 1999–2013 period. This analysis centers on regulatory actions in the southeastern United States and their effects on health outcomes in the 5-county Atlanta metropolitan area. The regulations examined are largely driven by the 1990 Clean Air Act Amendments (C). This work investigates regulatory actions and controls promulgated on EGUs: the Acid Rain Program (ARP), the NO_x Budget Trading Program (NBP), and the Clean Air Interstate Rule (CAIR) — and mobile sources: Tier 2 Gasoline Vehicle Standards and the 2007 Heavy Duty Diesel Rule.

METHODS

Each step in the classic accountability process (see Figure 1) was addressed using one or more methods. Linking regulations to emissions was accomplished by identifying major federal regulations and the associated state regulations, along with analysis of individual facility emissions and control technologies and emissions modeling (e.g., using the U.S. Environmental Protection Agency's [U.S. EPA's] MOtor Vehicle Emissions Simulator [MOVES] mobile-source model). Regulators, including those from state environmental and transportation agencies, along with the public service commissions, play an important role in implementing federal rules and were involved along with other regional stakeholders in the study. We used trend analysis, air quality modeling, satellite data, and a ratio-of-ratios technique to investigate a critical current issue, a potential large bias in mobile-source oxides of nitrogen (NO_x) emissions estimates.

The second link, emissions–air quality relationships, was addressed using both empirical analyses as well as chemical transport modeling employing the Community Multiscale Air Quality (CMAQ) model. Kolmogorov-Zurbenko

This Investigators' Report is one part of Health Effects Institute Research Report 195, 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. Armistead (Ted) Russell, Department of Civil and Environmental Engineering, Georgia Institute of Technology, 311 Ferst Drive, Atlanta, GA 30322; e-mail: *trussell@ce.gatech.edu*.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Assistance Award CR–83467701 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the views of the Agency, and no official endorsement by it should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

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

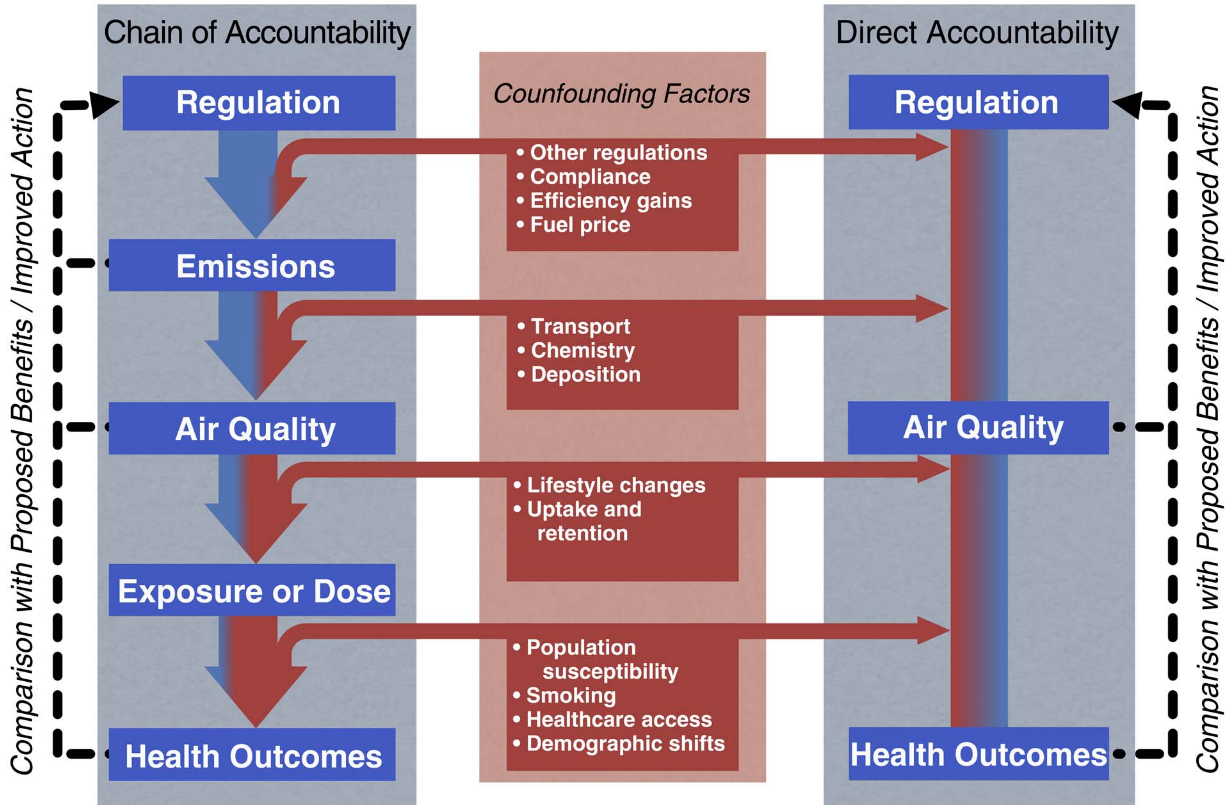


Figure 1. Two accountability frameworks: the chain of accountability and direct accountability. This study uses the chain of accountability approach, accounting for increasing confounding along the length of the chain using multiple models. (Reprinted from Henneman et al. 2016 with permission from Taylor & Francis.)

filtering accounting for day of the year was used to separate the air quality signal into long-term, seasonal, weekday–holiday, and short-term meteorological signals. Regression modeling was then used to link emissions and meteorology to ambient concentrations for each of the species examined (ozone [O_3], particulate matter ≤2.5 µm in aerodynamic diameter [$PM_{2.5}$], nitrogen dioxide [NO_2], sulfur dioxide [SO_2], carbon monoxide [CO], sulfate [SO_4^{2-}], nitrate [NO_3^-], ammonium [NH_4^+], organic carbon [OC], and elemental carbon [EC]). CMAQ modeling was likewise used to link emissions changes to air quality changes, as well as to further establish the relative roles of meteorology versus emissions change impacts on air quality trends. CMAQ and empirical modeling were used to investigate aerosol acidity trends, employing the ISORROPIA II thermodynamic equilibrium model (Ansari and Pandis 1999; Nenes 1998) to calculate pH based on aerosol composition. The relationships between emissions and meteorology were then used to construct estimated counterfactual air quality time series of daily pollutant concentrations that would have occurred in the absence of the regulations. Uncertainties in counterfactual air quality were captured by the construction of 5,000 pollutant time series using a Monte Carlo sampling technique, accounting for uncertainties in emissions and model parameters.

Health impacts of the regulatory actions were assessed using data on cardiorespiratory emergency department (ED) visits, using patient-level data in the Atlanta area for the 1999–2013 period. Four outcome groups were chosen based on previous studies identifying associations with ambient air pollution: a combined respiratory disease (RD) category; the subgroup of RD presenting with asthma; a combined cardiovascular disease (CVD) category; and the subgroup of CVD presenting with congestive heart failure (CHF).

Models were fit to estimate the joint effects of multiple pollutants on ED visits in a time-series framework, using Poisson generalized linear models accounting for overdispersion, with a priori model formulations for temporal and meteorological covariates and lag structures. Several parameterizations were considered for the joint-effects models, including different sets of pollutants and models with nonlinear pollutant terms and first-order interactions among pollutants. Use of different periods for parameter estimates was assessed, as associations between pollutant levels and ED visits varied over the study period. A 7-pollutant, nonlinear model with pollutant interaction terms was chosen as the baseline model and fitted using pollutant and outcome data from 1999–2005 before regulations might have substantially changed the toxicity of pollutant mixtures. In separate analyses, these models were fitted using pollutant and outcome data from the entire 1999–2013 study period. Daily counterfactual time series of pollutant concentrations were then input into the health models, and the differences between the observed and counterfactual concentrations were used to estimate the impacts of the regulations on daily counts of ED visits. To account for the uncertainty in both the estimation of the counterfactual time series of ambient pollutant levels and the estimation of the health model parameters, we simulated 5,000 sets of parameter estimates using a multivariate normal distribution based on the observed variance–covariance matrix, allowing for uncertainty at each step of the chain of accountability. Sensitivity tests were conducted to assess the robustness of the results.

RESULTS

EGU NO_x and SO_2 emissions in the Southeast decreased by 82% and 83%, respectively, between 1999 and 2013, while mobile-source emissions controls led to estimated decreases in Atlanta-area pollutant emissions of between 61% and 93%, depending on pollutant. While EGU emissions were measured, mobile-source emissions were modeled. Our results are supportive of a potential high bias in mobile-source NO_x and CO emissions estimates. Air quality benefits from regulatory actions have increased as programs have been fully implemented and have had varying impacts over different seasons. In a scenario that accounted for all emissions reductions across the period, observed Atlanta central monitoring site maximum daily 8-hour (MDA8h) O_3 was estimated to have been reduced by controls in the summertime and increased in the wintertime, with a change in mean annual MDA8h O_3 from 39.7 ppb (counterfactual) to 38.4 ppb (observed). $PM_{2.5}$ reductions were observed year-round, with average 2013 values at 8.9 µg/m^3 (observed) versus 19.1 µg/m^3 (counterfactual). Empirical and CMAQ analyses found that long-term meteorological trends across the Southeast over the period examined played little role in the distribution of species concentrations, while emissions changes explained the decreases observed. Aerosol pH, which plays a key role in aerosol formation and dynamics and may have health implications, was typically very low (on the order of 1–2, but sometimes much lower), with little trend over time despite the stringent SO_2 controls and SO_4^{2-} reductions.

Using health models fit from 1999–2005, emissions reductions from all selected pollution-control policies led to an estimated 55,794 cardiorespiratory disease ED visits prevented (i.e., fewer observed ED visits than would have been expected under counterfactual scenarios) — 52,717 RD visits, of which 38,038 were for asthma, and 3,057 CVD visits, of which 2,104 were for CHF — among the residents of the 5-county area over the 1999–2013 period, an area with approximately 3.5 million people in 2013. During the final two years of the study (2012–2013), when pollution-control policies were most fully implemented and the associated benefits realized, these policies were estimated to prevent 5.9% of the RD ED visits that would have occurred in the absence of the policies (95% interval estimate: −0.4% to 12.3%); 16.5% of the asthma ED visits (95% interval estimate: 7.5% to 25.1%); 2.3% of the CVD ED visits (95% interval estimate: −1.8% to 6.2%); and 2.6% of the CHF ED visits (95% interval estimate: −6.3% to 10.4%). Estimates of ED visits prevented were generally lower when using health models fit for the entire 1999–2013 study period.

Sensitivity analyses were conducted to show the impact of the choice of parameterization of the health models and to assess alternative definitions of the study area. When impacts were assessed for separate policy interventions, policies affecting emissions from EGUs, especially the ARP and the NBP, appeared to have had the greatest effect on prevention of RD and asthma ED visits.

CONCLUSIONS

This study demonstrates the effectiveness of regulations on improving air quality and health in the southeastern United States. It also demonstrates the complexities of accountability assessments as uncertainties are introduced in each step of the classic accountability process. While accounting for uncertainties in emissions, air quality–emissions relationships, and health models does lead to relatively large uncertainties in the estimated outcomes due to specific regulations, overall the benefits of regulations have been substantial.

INTRODUCTION

Tens of billions of dollars are spent annually in the United States, directly or indirectly, on controls to improve air quality. Globally, the costs are greater still. These controls are often justified by assessments quantifying the expected improvements in air quality, public health, and welfare that would accompany successful implementation (e.g., U.S. EPA 1997, 1998, 2005a, 2015). Stakeholders (e.g., those tasked with implementing the controls, government entities that promulgated the regulations, and others) ask if those controls are achieving their intended benefits. This simple question, however, is not trivial to answer given the many complexities inherent in linking specific actions (or interventions, often driven by one or more regulations) and outcomes. Accountability research in the air quality and health arena takes on this task.

The Health Effects Institute's Communication 11 (HEI Accountability Working Group 2003) — *Assessing Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research* — laid out a framework for accountability research as a chain of steps that connect regulations, emissions, air quality, exposure or dose, and health effects (Figure 1). Recent research has explored a second, complementary framework, termed *direct*, in which researchers link the outcome of interest (usually either air quality changes or public health outcomes) directly to the intervention itself by framing it as a hypothetical experiment in a potential-outcomes paradigm (Zigler and Dominici 2014). Henneman and colleagues (2016) provided a detailed review of accountability studies, including frameworks used in classic- and direct-accountability research studies.

The 2003 HEI report notes that a successful accountability study should address regulatory impacts on outcomes at each link in the chain of accountability, up to and including the outcome of interest (e.g., if the outcome of interest is air quality, one should look at the intervention's impact on emissions and air quality; if the outcome is health, the exposure or dose and health endpoints are added). Typically, researchers then apply these relationships, explicitly or implicitly, to a hypothetical scenario — often termed a *baseline* or *counterfactual* — that assumes the intervention did not occur, and attribute differences between the actual and counterfactual to the control measure. Previously, these counterfactuals have utilized information (e.g., ambient pollutant concentrations) from periods prior to or locations assumed to be unaffected by the promulgation of the intervention.

Within the direct accountability framework (Zigler and Dominici 2014; Zigler et al. 2016), researchers determine an *intervention condition* and a *control condition* such that, assuming portions of a populations were assigned randomly to these conditions, differences in the outcome of interest (typically air quality or health outcomes) between the two conditions are interpreted as the causal effects of the intervention.

The key to accountability research in both frameworks is to address factors that cloud the signal of interest (i.e., impacts of regulations on links in the chain of accountability; see Figure 1) (HEI Accountability Working Group 2003). Extraneous factors may inject uncertainty in relationships between all links in the chain. Between regulations and emissions, for example, the control implementation may be less effective than that planned by the regulators, because of the specific technology or because of enforcement, or because another regulation was implemented concurrently. Uncertainties remain in relating emissions to air quality due to incomplete knowledge of transport, chemistry, and emissions. Lifestyle changes, variability in access to care, and population changes impact the relationship between air quality, exposure, and health effects.

Early accountability studies typically focused on assessing the impact of short-term regulations and related events (such as plant closings) that caused relatively immediate changes in air pollution concentrations and allowed for the linkages along the chain of accountability to be established using before-and-after comparisons (HEI 2010; van Erp and Cohen 2009). Most early studies were limited to city scales. The field evolved by expanding to interventions spanning wider spatial and longer temporal scales, which required the development and application of newer, more powerful methods. Spatially broader interventions have the benefit of having larger potential impacts, but accountability studies assessing their impact may be subject to greater sources of error at each step in the chain of accountability. Both popular approaches, classic and direct, have been used in recent years to investigate such varied policies as coal bans in Irish cities (Clancy et al. 2002; Goodman et al. 2009), fuel sulfur bans in Hong Kong (Hedley et al. 2002; Peters et al. 1996), short-term policies enacted for Olympic Games and other major events (Friedman et al. 2001; Lin et al. 2013b; Peel et al. 2010; Rich et al. 2012), and long-term emissions reduction policies (including combinations of multiple policies) (Gauderman et al. 2015; Peters et al. 2009; U.S. EPA 2011; Zigler et al. 2012).

At the core of assessing the effectiveness of an air quality intervention in the accountability framework is a comparison of what the air quality was and what it would have been in the absence of the intervention (i.e., the *actual* versus the *counterfactual*). In a prospective study, both the actual and counterfactual must be projected; in a retrospective study, only the counterfactual needs to be estimated, since the actual case was observed (Cropper et al. 2017). There are a variety of ways to develop counterfactuals (HEI Accountability Working Group 2003; Henneman et al. 2016), but, because they are not observed, it is impossible to evaluate counterfactuals directly. Instead, judgment of their validity is based on multiple factors, such as the magnitude of counterfactual sensitivities to model inputs and assumptions (e.g., emissions changes, meteorology for air pollution, and air pollution at varying levels for health effects) and characteristics of the counterfactuals themselves in comparison to values before the intervention or in similar locations that did not have the intervention.

REGULATING AIR QUALITY IN THE SOUTHEASTERN UNITED STATES, 1990s TO THE PRESENT

The 1990 C marked a turning point in regulating air quality across the United States. The 1970 C (which amended the original 1963 Act) gave the U.S. EPA the authority to regulate air pollutants using two specific tools: air quality standards and emissions limits (National Research Council 2004). The U.S. EPA sets National Ambient Air Quality Standards (NAAQS) for six *criteria* pollutants — O_3, particulate matter (PM), CO, SO_2, NO_2, and lead — at levels designed to protect public human health and public welfare. Each pollutant has both primary (health) and secondary (welfare) standards (these are the same for many species), and PM is regulated both as $PM_{2.5}$ and as PM_{10} (particulate matter ≤10 μm in aerodynamic diameter). NAAQS are written as concentrations averaged over a specific period of time and follow specific statistical forms unique to each pollutant. U.S. EPA designates areas in exceedance of the NAAQS as nonattainment areas (NAAs), and requires the encompassing state to develop a State Implementation Plan (SIP) for reducing ambient air quality concentrations below the standards. The 1990 C clarified and expanded the U.S. EPA's previous authority related to NAAQS-setting and enforcement, mobile- and stationary-source emissions standards, emissions cap-and-trade programs, and permit requirements. The Cs expanded and modified the U.S. EPA's jurisdiction to regulate air toxics (hazardous air pollutants) and chemicals related to the stratospheric O_3 depletion.

Emissions standards aim to reduce the release of air pollutants from specific industries and source types, and are written either as emissions rates (emissions per activity, e.g., grams NO_x/mile, where NO_x is the sum of nitric oxide [NO] and NO_2) or as total allowable emissions over a specified amount of time. Some standards are applied to specific plants, while others are applied to a fleet of sources; some regulatory programs (e.g., the ARP defined in the 1990 C) set up trading markets that permit plant owners to buy and sell emission allowances (National Research Council 2004). For mobile sources, recent regulatory programs (e.g., the Tier 2 Gasoline Vehicle Standards and the 2007 Heavy Duty Diesel Rule) set standards for both engines and fuel composition. Mobile-source emissions limits are set at a federal level; however, the U.S. EPA allows one state, California, to set mobile emissions standards independent of the national levels (though they must be at least as stringent), and other states can adopt either the federal or California standards. States use other tools to reduce emissions in NAAs, such as setting limits on Reid vapor pressure in gasoline that are below the federal limits and overseeing inspection and maintenance (IM) programs.

In response to regulations contained in the 1990 C, the U.S. EPA and the Georgia Department of Natural Resources' Environmental Protection Division (EPD) have applied various regulatory tools to improve air quality, with a focus on Atlanta, which frequently exceeds the NAAQS for O_3 and $PM_{2.5}$. Assessments of the effectiveness of specific regulations, however, are made difficult by the complex interplay between national regulations and their implementation at the state and local levels. For example, the EPD has implemented multiple emissions standards on stationary sources separate from national programs. Often, the state programs, codified in SIPs, are similar in approach and timing to national programs and may be developed in negotiations between regulators, utilities, and public service commissions that govern utilities. A state may promulgate a rule to achieve multiple objectives or meet multiple national standards (e.g., O_3 and $PM_{2.5}$ share precursors). Further, a utility, whose actions are subject to public service commission rulings, may seek to identify the most cost-effective measures to address multiple regulations. Such interconnected emissions policies affect air quality in varying ways depending on multiple factors such as source industry, location and stack height, economic activity, and climate. Any assessment of the effectiveness of specific regulations implemented under the C, therefore, must begin by acknowledging the intermeshed nature of air pollution regulations.

ACCOUNTABILITY ASSESSMENT OF REGULATIONS IN THE SOUTHEASTERN UNITED STATES

The southeastern United States, with its large, discrete population centers, humid subtropical climate, and high levels of biogenic emissions, presents a unique environment for investigating the effectiveness of regulations promulgated under the 1990 C. The region sees high photochemical activity, and the effects of this activity are accentuated by both natural and anthropogenic emissions, particularly of volatile organic compounds (VOCs), PM, NO_x, and SO_2. Biogenic emission sources contribute large amounts of VOCs, which serve as precursors to O_3 and $PM_{2.5}$. Southeastern cities experience elevated pollutant levels that often exceed the NAAQS, but these cities have been the beneficiary of a number of recent control programs and have seen dramatic improvements in air quality over recent decades (Blanchard et al. 2013).

This study uses multiple models to investigate impacts of controls implemented under the C on multiple links along the chain of accountability. The primary analysis of the study focuses on the development of daily counterfactual time series of ambient concentrations that assume the absence of specific control programs over the period of interest. Analyses of these counterfactuals allow for the quantification of air quality and health benefits related to specific air pollution regulations under the 1990 C.

The methodical approach assesses nine specific regulatory scenarios focused on EGU and mobile emissions. Meetings with regulatory agency staff and regional stakeholders that helped develop and implement these regulations proved helpful in assessing impacts of regulations on emissions. It is of interest to assess these programs, considering that multiple changes occurred during the period of interest that likely had large impacts on emissions, such as the recessions in the early and late 2000s, meteorological variability (such as in 2009 and 2013, two years with relatively cool, wet summers), population shifts, and the retrofit of major coal plants near the city from coal to natural gas.

Two empirical approaches, meteorological detrending and a statistical model of emissions, and a deterministic model (CMAQ) were employed to assess emissions–meteorology–air-quality relationships. Results of this analysis are combined with assessments of emissions reductions to create daily time series of ten air pollution species for each of nine counterfactual scenarios assuming the absence of one or more regulations and to estimate uncertainty bounds. Intermediate results from all three models provide insights into both the confidence in modeling results and relevant atmospheric relationships (such as aerosol pH and NO_x emissions levels).

Air quality results of the present study are specific to the southeastern United States, but the majority of the methods are more broadly applicable to areas with long records of ambient air quality measurements.

ASSESSMENT OF THE HEALTH IMPACTS OF POLLUTION-CONTROL POLICIES

There are numerous challenges to estimating health impacts for this study as well as for air pollution accountability studies in general. Policies may not be implemented over a precise period, there may be several policies implemented concurrently, and the effects are not limited to a specific geographical boundary nor are they uniformly distributed within any area. Importantly, while policies may be aimed at lowering ambient pollutant levels below specific standards with the ultimate goal of reducing air-pollution–linked health outcomes, pollution controls affect pollutants at the emissions level, so that there are multiple levels of separation between the measured effects of policies on emissions and their eventual health impacts. Finally, studies measuring health impacts through this chain of accountability often rely on health associations for individual pollutants, which can fail to capture the overall impact of policies that can impact emissions of multiple toxic pollutants.

This study utilizes two critical analytic tools that are intended to address these limitations of accountability studies: counterfactual analysis and multipollutant modeling.

COUNTERFACTUAL STUDY DESIGN

Many observational studies assessing the effects of an intervention use a pre–post approach, which contrasts outcomes in the period after an intervention to outcomes in the period before the intervention (Henschel et al. 2012; Li et al. 2010; Lin et al. 2013a, 2014; van Erp et al. 2011). Differences between outcomes before and after an intervention could possibly be due to many factors other than the intervention itself, so that study results may be biased by temporal confounding, which can be a particularly major concern if periods of interest are relatively long.

Other studies may utilize the approach of selecting geographical locations that did not implement the interventions of interest as controls for comparison (Atkinson et al. 2009; Boogaard et al. 2012). The assumptions being made in these types of analyses are that trends in pollution-related health outcomes in different locations would have been the same in the absence of the intervention. Similar to pre–post analyses, these types of studies become increasingly susceptible to confounding over longer study periods (for these studies, through spatiotemporal effect modification), as different geographical areas may undergo important changes that lead to diverging secular trends. Since both study types utilize explicit comparisons between intervention and reference populations (defined by either time or space), these study types are well suited to interventions that take place fully during a specific, well-defined period.

This current study assesses impacts of numerous overlapping pollution-control policies, which were implemented to full effect over long periods, such that there are no clear-cut reference or intervention periods. Rather, this study can be conceptualized as measuring the effects of a progressively increasing gradient of the implementation of these policies. These effects could not be adequately captured using a classical pre–post approach, which typically measures the effect for a binary indicator of the presence or absence of an intervention; however, a counterfactual analysis is ideally suited for this task. For counterfactual accountability studies, outcomes after an intervention are compared with outcomes during the same period and in the same geographical location but in the absence of the intervention. In other words, in these studies the outcome is estimated for the hypothetical situation in which the intervention(s) of interest had not been implemented, but all external factors are held constant; then these counterfactual outcomes are compared with observed outcomes to determine the impact of the interventions.

The counterfactual design used for this study modeled continuous changes at all stages on the chain of accountability and was therefore able to more fully capture the impacts of these policies. The primary limitations of counterfactual studies involve potential estimation errors through the modeling of counterfactual exposures and outcomes. To address this potential issue, we used Monte Carlo simulations to account for model uncertainty at multiple study stages.

MULTIPOLLUTANT HEALTH IMPACT MODELING

There is an extensive body of literature linking air pollution and cardiorespiratory health outcomes, which includes population-level studies and panel studies, cohort studies, and time-series studies, studies on indoor and ambient air pollution, experimental studies, and observational studies (Brunekreef and Holgate 2002; Cohen et al. 2005; Curtis et al. 2006; Seaton et al. 1995). Since all people are exposed to some level of air pollution and there are many dimensions at which exposure can be measured, there are myriad ways to conduct studies assessing the relationship between air pollution and health outcomes. Quantifying the effects of ambient air pollution — a mixture of numerous gaseous and particulate pollutants — on human health can be difficult, and every study design has a unique set of strengths and weaknesses. Traditional assessments of single-pollutant health associations may be misleading: measured associations may actually be capturing the effects of highly correlated pollutants originating from common sources (Woodruff et al. 2009). Conversely, health associations for multipollutant models are difficult to interpret, especially as differing amounts of measurement error can lead to incorrect assumptions about the effects of correlated pollutants (Brunekreef and Holgate 2002; Tolbert et al. 2007).

However, assessing whether pollutant–health associations exist is not a goal of this study. The vast amount of air pollution research conducted over many decades has led to overwhelming evidence that ambient air pollution, in various forms, is linked to adverse cardiorespiratory health outcomes. These associations are not being tested in this study, which leads to considerable freedom in modeling health impacts; the joint impact of complex pollutant mixtures can be properly assessed without attempting to disentangle individual pollutant effects. Single-pollutant or bipollutant models may underestimate the true impact of air pollution mixtures (Mauderly et al. 2010), strengthening the argument for utilizing larger multipollutant models.

More intricate investigations of health impacts are possible if assessments of individual associations are not required. There is no reason to believe that effects of individual pollutants are independent of levels of other pollutants; rather, multipollutant mixtures may exhibit important atmospheric chemical reactions or synergistic health effects (Mauderly and Samet 2009). In addition, the dose–response curves between pollutants and health effects may not be linear (Schwartz and Zanobetti 2000). The use of pollutant interaction terms and nonlinear terms in a multipollutant health impact model may make individual pollutant contributions difficult to decipher. However, if assumptions about pollutants being independently and linearly associated with health outcomes are false, these more detailed models could result in more accurate and complete assessments of the impact of pollution-control policies.

LONG-TERM POLLUTANT AND HOSPITAL ED DATA SETS

One other key strength of this study is the availability of long-term data sets on ambient pollutant levels and hospital ED visits in the Atlanta metropolitan area. Numerous studies have used these data sets in time-series analyses to measure health associations of daily ambient pollutant levels. Early studies established an association between several pollutants (such as NO_2, CO, O_3, $PM_{2.5}$, and PM_{10}) and assorted cardiovascular and respiratory outcomes (Metzger et al. 2004; Peel et al. 2005; Tolbert et al. 2000). Subsequent analyses quantified health effects of ambient air pollutants on children and other susceptible groups (Darrow et al. 2011; O'Lenick et al. 2017; Peel et al. 2007; Strickland et al. 2010, 2014). Efforts were made to characterize and address potential methodological issues arising in time-series studies such as measurement error (Goldman et al. 2011; Sarnat et al. 2010), temporal time windows (Darrow et al. 2011), residual confounding (Flanders et al. 2011, 2017), power estimation (Winquist et al. 2012b), and spatial variability (Sarnat et al. 2013; Strickland et al. 2011). Some analyses utilized alternative exposures such as pollen counts (Darrow et al. 2012), VOCs (Ye et al. 2017), or oxidative potential (Bates et al. 2015; Fang et al. 2016), or used alternative outcomes such as hospital admissions (Winquist et al. 2012b). Additional studies have explored novel methods such as joint pollutant effects (Winquist et al. 2014) or Bayesian ensemble-based source apportionment (Gass et al. 2015). All these studies contributed to the epidemiological literature, not only for the quantification of the acute effects of ambient air pollution, but also on time-series analytic methodology. Crucially, these studies helped to improve confounder control in order to decrease bias stemming from effects of meteorology, seasonality, long-term trends, and other time-varying factors on both air pollution and ED visitation patterns.

SPECIFIC AIMS

This study approaches the accountability of regulations under the 1990 C through six objectives, the first three of which are focused on the first three links in the chain of accountability. The final three objectives assess the linkages between changes in air pollution and improvements in public health, and evaluate factors (modeling assumptions and estimation error) that can affect estimation of the overall impacts of pollution-control policies.

This report is organized as follows: the current section covers specific objectives of the project and details the study area, regulatory actions of interest, and the importance of stakeholder engagement in the conduct of this work. The next two major sections, "Study Design and Methods" and "Results," step through links in the chain of accountability to maintain consistency in the discussion and aid the reader in keeping track of the chain of events as they happened in reality. The "Discussion and Conclusions" section builds on the results, integrating across the findings. An "Implications of Findings" section identifies those issues that rise to the top in terms of potentially important conclusions that could warrant action. The main body of this report provides the key aspects of the study. Additional details can be found in the Appendices, Additional Materials (available on the HEI website), and journal articles based on this work (see "Other Publications Resulting from This Research"). The reader should note that we make use of multiple acronyms — which we define at their first use — throughout the text. We refer the reader to the "Abbreviations and Other Terms" section at the end of this report.

OBJECTIVES

The current study had six objectives, which are listed below. This report is organized according to the chain of accountability, so while all of the objectives are addressed, they are not necessarily addressed in order.

1. Quantitatively assess the impacts of controls on utility NO_x and SO_x emissions on air quality in the southeastern United States.
2. Quantitatively assess the impacts of controls on light-duty vehicle NO_x, VOC, and CO emissions and on heavy-duty vehicle NO_x and PM emissions on air quality in the Southeast.
3. Assess the impact of meteorological trends on air quality in the Southeast.
4. Assess the impacts of the regulatory programs on acute cardiorespiratory ED health outcomes.
5. Evaluate the effects of methodological choices used in estimating the impact of pollution-control policies.
6. Conduct uncertainty analyses capturing potential error in parameter estimation at multiple study stages to construct comprehensive confidence interval estimates.

STUDY AREA

Most of the measurement-based empirical analysis focuses on Atlanta, Georgia. Atlanta is both the most populous metropolitan area in Georgia and a rapidly growing urban hub; the 5-county Atlanta metropolitan area (DeKalb, Fulton, Cobb, Gwinnett, and Clayton counties) grew from 2.85 million people in 1999 to 3.54 million in 2013, a 24.2% increase (see Appendix C, available on the HEI website). The current $PM_{2.5}$ 20-county Atlanta nonattainment area (ANAA) grew from 4.01 million people in 1999 to 5.35 million in 2013 (a 33.5% increase), while the entire state of Georgia increased in population 24.2% during this time, from 8.05 to 9.99 million; for reference, the U.S. population grew 13.3% from 1999–2013.

The Jefferson Street monitoring station (JST), which is part of the SouthEastern Aerosol Research and CHaracterization (SEARCH) network (Hansen et al. 2003), is located roughly 2 miles northwest of downtown Atlanta and about 1.4 miles from a major interstate highway. This site has provided detailed air quality data since 1998. Multiple aspects of the study's air quality analyses — meteorological detrending, aerosol acidity analyses, and ambient air quality trend assessment — were extended to other SEARCH sites. These are addressed in brief in the main body of this report and in detail in publications and appendices associated with this report (Henneman et al. 2015, 2017a, 2017b; see Appendix E available on the HEI website). SEARCH site abbreviations are listed at the end of Appendix E.

The ANAA serves as the boundary for the mobile emissions analysis, while EGU emissions are split into two categories: those within the ANAA and those within six southeastern states (Alabama, Georgia, Mississippi, North Carolina, South Carolina, and Tennessee) that have been shown to be major contributors to air quality issues in Georgia, as well as in Alabama (Bergin et al. 2005).

CMAQ air quality modeling associated with this study covers the East Coast of the United States. The two approaches — observation-based studies using monitor data and regional air quality modeling encompassing a much broader area — allow for a detailed look into both spatial and temporal scales important to this accountability analysis.

SCENARIOS AND REGULATORY PROGRAMS OF INTEREST

Defining the regulatory actions of interest is an important first step in accountability analyses (HEI Accountability Working Group 2003). EGUs and on-road mobile (MOB) vehicles are major sources that impact air quality in Atlanta. Both of these source categories have been subject to regulations promulgated under the 1990 C (U.S. EPA 2011), and those regulations are suspected of achieving both emissions reductions and health benefits. This project begins by investigating the emissions reductions that have been achieved under specific programs in the Southeast.

Because of the overlapping nature of regulations at national, state, and local levels, it is difficult to separate the impacts of concurrent regulations on emissions. As discussed above, a utility may decide to install a control device for multiple reasons, and any one regulatory action may drive the implementation of various control options (e.g., changes in fuels, retrofits, or plant closings). Similarly, automobile controls include on-board (largely federally regulated) and fuel-related (both state and federally regulated) controls. The majority of the control programs of interest, therefore, are defined by both the national and local programs that drive the changes in the modeled counterfactuals.

The first regulatory scenario of interest is the EGUMOB scenario (Table 1), which considers all of the regulatory actions analyzed here on EGUs and mobile sources over the period of interest. This scenario was the most straightforward to estimate because it is based on total energy demand, vehicle activity, and emissions factors (the relationship of emissions to activity) from the beginning of the period. The majority of the regulatory scenarios listed in Table 1 are named after the relevant national programs, but

Table 1. Scenarios Used in the Analysis

Scenario Name	Regulatory Programs	Period[a]
EGUMOB	All EGU and mobile regulations	1993–2013 (Mobile) 1995–2013 (ANAA) 1997–2013 (Southeast)
EGU_{ALL}	All EGU regulations	1995–2013 (ANAA) 1997–2013 (Southeast)
EGU_{ARP}	ARP and $GRAQC_{yy}$	1995–2013 (ANAA) 1997–2013 (Southeast)
EGU_{NBP}	NBP and SIP Call and $GRAQC_{jjj}$	1999–2013
EGU_{CAIR}	CAIR and $GRAQC_{sss}$	2009–2013
MOB_{ALL}	All mobile regulations	1993–2013
MOB_{IM}	Inspection and Maintenance	1993–2013
MOB_{GSP}	Tier 2 Gasoline Program $GRAQC_{bbb}$	2000–2013
MOB_{DSP}	Heavy Duty Diesel Rule	2006–2013

[a] Years provided correspond to the dates used in the analysis, and not necessarily to the dates of the regulatory action. Differences between the dates used and the program start dates stem from data variability.

GRAQC = Georgia Rules for Air Quality Control (Georgia EPD 2013); GRAQC= Gasoline Marketing Rule; $GRAQC_{jjj}$ = NO_x Emissions from Electric Utility Steam Generating Units; $GRAQC_{sss}$ = Multipollutant Control for Electricity Utility Steam Generating Units; $GRAQC_{yy}$ = Emissions of Nitrogen Oxides from Major Sources.

refer to both the national and state programs (and continue to do so for the entirety of this report). Georgia state programs are named by their title in the Georgia Rules for Air Quality Control (GRAQC) documents published by the Georgia Environmental Protection Division (Georgia EPD 2013). The last column in Table 1 refers to the period used to estimate the effects of each regulatory program. Each counterfactual scenario describes the hypothetical conditions under which the controls of interest were not implemented and uses the first year in the period as the base year (i.e., the year before effects are estimated). Relative program implementation time lines are displayed in Figure 2.

EGU Programs

The three national EGU programs investigated in most detail in this report are the ARP, the NBP and its associated SIP Call, and the CAIR.

The ARP was finalized and promulgated in 1993 to combat increasing SO_2 and NO_x emissions throughout the United States, especially in eastern states (National Research Council 2004; U.S. EPA 2002). Under Title IV of the Clean Air Act Amendments in 1990, Congress set out to reduce annual SO_2 emissions in the United States by 50% in 2010 compared with 1980 levels. To achieve these reductions, the legislation prescribed a cap-and-trade approach for SO_2 and an emissions factor (in mass per heat input) limit for NO_x that included two phases. Phase I of the ARP, which began in 1995 for SO_2 and 1996 for NO_x, targeted the largest existing power plants. Starting in 2000, Phase 2 of the ARP required all other plants regulated under Title IV of the CAA to achieve emissions reductions. In order to ensure reductions were being made, continuous emissions monitors were required for both SO_2 and NO_x on all regulated stacks (Morgenstern et al. 2012).

To address the problem of O_3 precursors being transported across state lines in the East, the U.S. EPA issued the NO_x SIP Call in 1998. This call was meant to improve the implementation of the controls established under the ARP. The SIP Call did not place a limit on individual sources; instead, it required each state to develop a plan to reduce NO_x emissions during the O_3 season that contributed to nonattainment in downwind states, particularly in the northeastern United States (U.S. EPA 1998, 2003). The U.S. EPA began the NBP under the 1998 SIP Call to aid states in their effort to meet their emissions budgets. The NBP was a

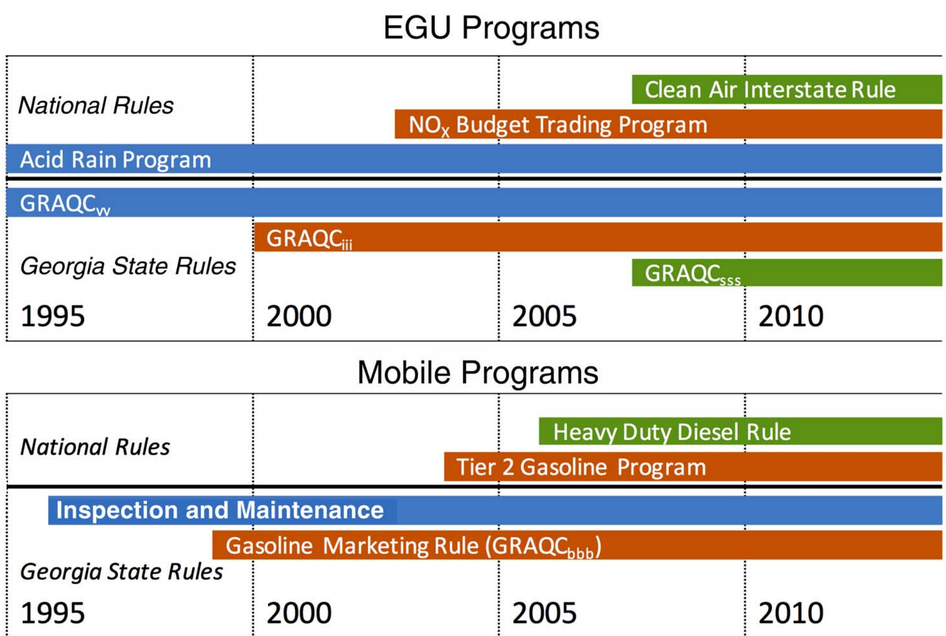


Figure 2. Implementation time lines for mobile and EGU regulatory programs at the federal and state levels. Each beginning date denotes the start of enforcement for each program and may not reflect the first time initial effects are reported in the remainder of this report. GRAQC = Georgia Rules for Air Quality Control (Georgia EPD 2013). The subscripts on GRAQC refer to sections of the Rules.

cap-and-trade strategy that was optional; however, all 20 states in the NO_x SIP Call and the District of Columbia used the program to help meet their NO_x SIP Call requirements by 2003 (U.S. EPA did not assess each state's performance until 2007). A portion of northern Georgia was included under the original draft of the NO_x SIP Call, but was later removed from the requirements of the rule because of court actions and the U.S. EPA's redesignation of Birmingham, Alabama, and Memphis, Tennessee, nonattainment areas (U.S. EPA 2008). Georgia began requiring seasonal NO_x controls on EGU sources in the ANAA beginning in 2000 under a state program ($GRAQC_{jjj}$) similar to ones adopted by other states under the SIP Call.

CAIR, promulgated in 2005, was the regulatory approach to further reducing NO_x and SO_2 emissions adopted by the U.S. EPA after the Clean Skies Act did not pass Congress. The focus of CAIR is pollutant (e.g., $PM_{2.5}$) transport across state borders. The regulation affected 28 eastern states, and set up three interstate emissions trading programs: the CAIR SO_2 annual trading program, the CAIR NO_x annual trading program, and the CAIR NO_x ozone-season trading program (U.S. EPA 2005a,c). In effect, CAIR extended the ozone-season NO_x controls under the NPB to the entire year, and required large coal plants to install SO_2 controls. The U.S. EPA reports that CAIR should have led to reductions of NO_x and SO_2 emissions of 41% and 44%, respectively, in 2015 (U.S. EPA 2005a). A 2008 court decision kept CAIR in place, but instructed the U.S. EPA to develop an alternative rule that satisfies CAA requirements related to cross-state transfer of air pollutants. The resulting rule, the Cross-State Air Pollution Rule, was promulgated in 2011 and delayed by litigation until 2015. It is not addressed in the analysis here, but it is possible that the rule, even in its proposed stage, may have impacted utilities' decisions regarding controls and fleet makeup toward the end of the study period.

Each of the major national EGU rules relates to a rule in the GRAQC. $GRAQC_{yy}$ — Emissions of Nitrogen Oxides from Major Sources — required EGUs in 13 counties originally in the Atlanta 1-hr O_3 NAA to install "reasonably available control technology" (Georgia EPD 2013) by July 1995. The rule was expanded to 32 more counties in 1999 and to smaller facilities in 2007. $GRAQC_{jjj}$ — NO_x Emissions from Electric Utility Steam Generating Units — established summertime emissions limits (1 May–30 September) on a lb/mmBtu heat input basis for plants in the 13-county NAA. Compliance was required by some units as early as summer 1999, and more plants were added each summer between 2000 and 2002. $GRAQC_{sss}$ — Multipollutant Control for Electricity Utility Steam Generating Units — was promulgated in 2007, and established dates of compliance for specific plants regarding installation of selective catalytic reduction (a NO_x emissions control device) and flue gas desulfurization (an SO_2 emissions control device) on specific units. This rule overlaps in both date and purpose with CAIR in that it requires year-round controls on NO_x and strict controls on SO_2 from coal-fired power plants.

Mobile Programs

Enhanced inspection and maintenance have been required on automobiles registered in 13 counties surrounding Atlanta since 1996. The affected counties are as follows: Cherokee, Clayton, Cobb, Coweta, DeKalb, Douglas, Fayette, Forsyth, Fulton, Gwinnett, Henry, Paulding, and Rockdale. In general, the requirement covers gasoline-powered cars and light trucks, specifically those 24 model-years old and newer.

Two separate gasoline programs are included in this analysis. Beginning in 1999, Georgia required gasoline sold in an expanded region of 25 counties to have a summertime Reid vapor pressure of 7 psi or less and a seasonal average sulfur content of less than 150 ppm (by weight). The seasonal sulfur limit was reduced in 2003 to 90 ppm and was reduced further in 2004 to a 30-ppm year-round average. In 2006, the federal Tier 2 limit required an annual average of 30 ppm sulfur fuel or less. Therefore, before 2006, benefits are attributed to the Georgia Gasoline Marketing Rule and afterward to the Tier 2 program.

Lowering Reid vapor pressure is an approach to reduce evaporative VOC emissions. Reducing sulfur both reduces SO_2 and SO_4^{2-} emissions and improves the efficiency of selective catalytic reduction controls that reduce NO_x emissions. The focus of both the Georgia Gasoline Marketing Rule and the Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Requirement is to reduce mobile contributions to ambient O_3 and PM levels by reducing NO_x and VOC emissions. The rule has the further effect of reducing SO_x emissions, which contribute to the formation of secondary PM through atmospheric conversion to SO_4^{2-} (U.S. EPA 1999).

The Tier 2 Program included updated engine emissions standards applicable to all passenger cars, light trucks, and medium-duty passenger vehicles. The standards were phased in between 2004 and 2009. Average manufacturer fleet vehicle emissions from new vehicles under this program are required to meet a standard of 0.07 g/mile NO_x, and nonmethane hydrocarbons (NMHCs, a component of VOCs) are regulated based on which of several certification bins each car model fits into. NO_x standards before this rule was in place ranged from 0.30 g/mile to 1.53 g/mile depending on the type of vehicle.

The 2007 Heavy Duty Highway Rule was promulgated in 2001 (U.S. EPA 1997, 2000). Like the Tier 2 gasoline rule, this program sets standards for both engines and fuel. The goal of this regulation was to reduce O_3 levels by reducing O_3 precursor emissions (NO_x and NMHCs). One major aspect of the rule was limiting sulfur content to 15 ppm or less by June 2006. This improves the efficiency of post-combustion selective catalytic reduction at removing NO_x and VOCs. Diesel sales in Georgia went from being comprised of 91% fuel in 2006 with a sulfur content of between 15 and 500 ppm to 35% in 2007, with the difference being declining sales of diesel with sulfur content greater than 500 ppm and increasing sales of fuel with sulfur content less than 15 ppm. By 2008, no diesel with sulfur content greater than 500 ppm was sold, and by 2012, 100% of fuel sold in Georgia had less than 15 ppm sulfur (U.S. Energy Information Administration 2014).

A second major component of the Heavy Duty Highway Rule is the reduction of NO_x and NMHC emissions standards applicable to all highway heavy-duty engines. These standards were enforced beginning on model-year 2004 vehicles. The previous standards were for total hydrocarbons at 1.3 g/bhp-hr and 4.0 g/bhp-hr for NO_x. The updated standard is a limit of 2.4 g/bhp-hr for combined NO_x and NMHCs or 2.5 g/bhp-hr with a limit of 0.5 g/bhp-hr on NMHCs (U.S. EPA 1997). In 2007, standards were introduced, to be phased in from 2007 to 2010, that limited NMHC emissions to 0.14 g/bhp-hr and NO_x emissions to 0.2 g/bhp-hr.

REGULATOR AND STAKEHOLDER ENGAGEMENT

A goal at the outset of this study was to involve the numerous regulators and stakeholders that are involved in developing, interpreting, and implementing air pollution regulations. As discussed above, many of the policies addressed in this report are complicated by their formulation, adoption, execution, and relationships with other policies. These complications have a direct bearing on interpreting regulatory impacts on emissions, air quality, exposure, and health. To gain a deeper understanding of the regulatory framework, we approached both current and former employees of many regulators and stakeholders in the region, including Georgia Power and Southern Company (Georgia Power is an operating unit of Southern Company), the U.S. EPA Region IV, and the Georgia EPD for information on developing air quality policies and regulating emissions in Georgia and the Southeast (Southern Company has another operating unit in Alabama) in general, and how the specific regulations we studied were developed in particular. These discussions pointed us to important additional sources for information — such as archived Georgia SIPs and Utility Integrated Resource Plans — and informed our methods, particularly in estimating counterfactual EGU and mobile-source emissions.

STUDY DESIGN AND METHODS

STUDY DESIGN OVERVIEW

This study approaches accountability by estimating effects of regulations on multiple outcomes, in particular, emissions, air quality, and health effects. First, a set of policies impacting two of the historically largest sources of anthropogenic emissions in the Southeast — on-road mobile sources and EGUs — was identified and operationally defined. It is important to note that over the period studied, emissions from both of those source categories have decreased appreciably (along with emissions from most other anthropogenic sources), and they continue to decrease (Table 2, Appendix Figures A.1 and A.2, available on the HEI website).

EGU and mobile emissions changes were quantified according to each of the scenarios in Table 1. For each program, estimates were made of daily counterfactual emissions that assume the regulatory program did not go into place, but that the associated activity levels (e.g., electricity demand and vehicle miles traveled [VMT]) are the same as in the base (actual) case. Empirical statistical methods and a chemical transport model (CTM) were used to estimate sensitivities of air pollution concentrations to emissions. These models were run with program-specific emissions changes to estimate counterfactual daily time series of air pollutant concentrations and the effects of each regulatory action on air quality. CTM results were used to investigate other questions related to accountability, including the model's ability to capture changes in spatial and pollutant variability in regulatory impact across the period of interest. Appendices include more detailed emissions and air quality analyses, employing empirical, CTM, and satellite methods to investigate potential bias in mobile NO_x emissions (Appendix D) and pH trends (Appendix E) across the Southeast, which are briefly summarized in this report.

Poisson generalized linear models were utilized to estimate associations between observed ambient pollutant levels and daily counts of ED visits for cardiorespiratory outcomes. Parameter estimates obtained from these models are combined with the difference between counterfactual and observed pollutant levels in order to estimate the difference between observed and counterfactual ED visits (hereafter "ED visits prevented") for each pollution-control scenario. Uncertainties in the emissions change

Table 2. Emissions Changes for Various Species and Sources Between Reference Year[a] and 2013

Source (y*) / Species	Emissions[b] in y* (tons/day)	Emissions in 2013 (tons/day)	Decrease (%)
ANAA EGU[c] (1995)			
NO_X	303	43	86
SO_2	920	142	85
Regional EGU[d] (1997)			
NO_x	2,710	475	82
SO_2	5,604	943	83
MOB (1993)			
NO_x	567	127	78
SO_2	15	1	93
$PM_{2.5}$	30	10	67
CO	4,306	1,421	67
VOC	326	123	62
EC	12	4	61
OC	18	7	67

[a] Reference years (y*) are selected as the first available year with complete data for each source.

[b] Emissions are reported as daily averages.

[c] Atlanta nonattainment area (ANAA) EGUs are within the 20-county Atlanta nonattainment area,

[d] Regional EGUs are located in seven Southeastern states (and do not include ANAA EGUs).

estimates, empirical air quality models, and health impact models are combined via Monte Carlo simulations for a total uncertainty estimate.

Each of the three major sections in this report ("Study Design and Methods," Results," and "Discussion and Conclusions") is organized in the order of the chain of accountability, focusing on each of the links in the chain as appropriate to emphasize the passing of information from one to the other. Figure 3 shows the flow of information for the central work of the study.

EMISSIONS

The first link in the chain of accountability (Figure 1) represents the relationship between regulations and emissions. This relationship can be difficult to estimate because of the number of factors that influence emissions, including energy demand (which is impacted by population, economic activity, and energy efficiency, etc.), fuel type (which is impacted by availability and price), regulations on other sources (utilities will trade or buy electricity from other utilities if it is cheaper than producing it themselves), and others.

Two sources of uncertainties were introduced by the emissions portion of this study: uncertainty in the actual emissions used to relate regulations to emissions and emissions to air quality, and uncertainty in attributing emissions changes to regulations. The first source has been addressed previously, though it is still much larger in mobile-source than in EGU emissions. The second is large and is assumed to dominate emissions uncertainties in the modeling.

This section describes the sources (EGU and mobile) for emissions data and includes a discussion of associated uncertainties. The final two subsections address the approaches for estimated changes in emissions attributable to specific regulatory programs for each of the scenarios in Table 1.

EGU Emissions Data

EGU emissions data for six southeastern states (Alabama, Georgia, Mississippi, North Carolina, South Carolina, and Tennessee) were downloaded from the U.S. EPA Air Markets Database (U.S. EPA 2016a). Continuous emissions monitoring (CEM) data, required under the ARP, are

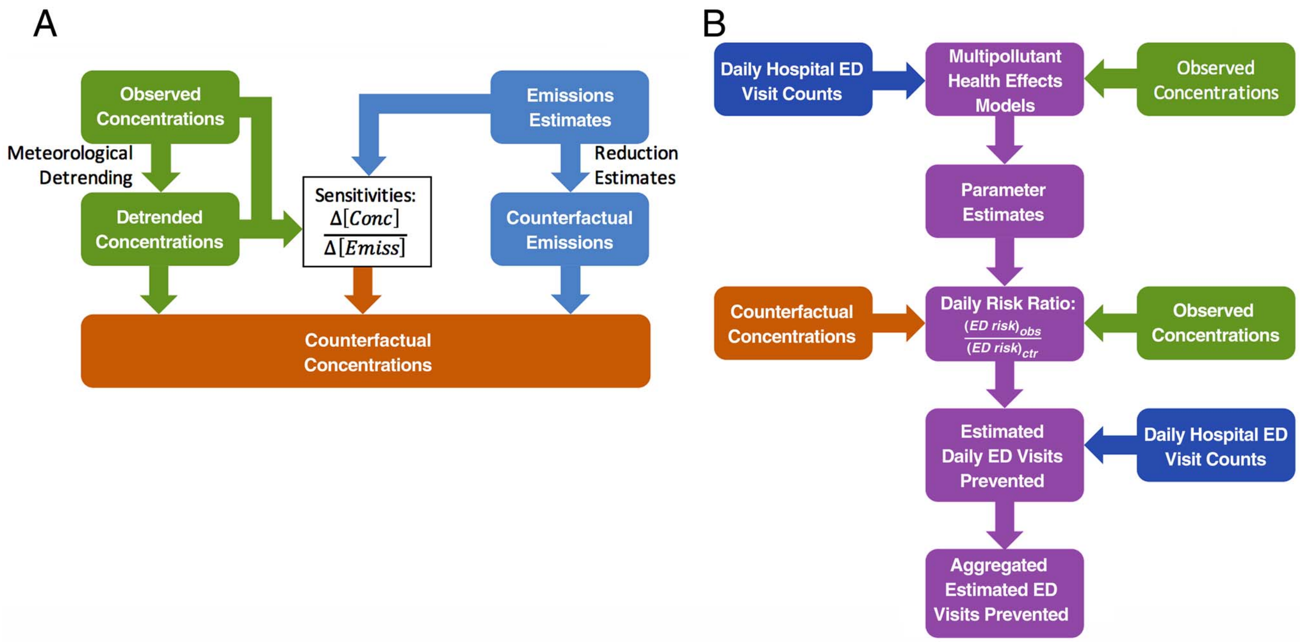


Figure 3. Outline of steps used in the air quality (A) and health impact (B) analyses.

available online for hourly emissions at the unit level for large emitters of NO_x, SO_2, and CO_2 beginning in 1995. A review of the data led to the conclusion that full compliance in the region did not occur until 1997 (Alabama and Mississippi appear to be missing data for 1995–1996). These daily emissions were split into two groups: those from plants within the ANAA and those outside (denoted REG, for "regional").

EGU Emissions Data Uncertainty

Uncertainty in EGU emissions is small relative to uncertainty in other data used in this project because CEM data are measured directly at the source. Napelenok and colleagues (2011) used an uncertainty of ±3% for major point source NO_x emissions, and Hanna and colleagues (2001) used an uncertainty of ±50% in major point sources, though the work did not assume a universal application of CEM. Given the use of CEM data in this work, uncertainty in EGU emissions was assumed to be negligible compared with other sources of uncertainty, such as that of estimating counterfactual emissions or linking emissions with changes in air quality.

Mobile Emissions Data

U.S. EPA's MOVES 2010b (U.S. EPA 2012b) software was applied to model daily on-road emissions for the ANAA from 1993–2012. When possible, the MOVES entries in the default input database were replaced with Atlanta-specific data from the EPD. The EPD supplied MOVES-ready IM inputs inferred from registration data in Atlanta, vehicle population, speed distribution, fuel formulation, road-type distribution, and vehicle-type age distribution. They provided detailed annual data for 2002, 2008, and 2010. Population change in Fulton County, which contains the majority of the area of the city of Atlanta, was used to scale vehicle populations in years for which detailed annual data were not available, and other inputs were interpolated linearly between known years. Vehicle-type age distribution was available for each year of interest from the EPD. Estimated average annual vehicle miles traveled was obtained from the Georgia Department of Transportation database (Georgia Department of Transportation 2017).

While the current study does not focus on non-road mobile emissions (from sources such as aircraft, construction equipment, and lawn mowers), they contribute significantly to total mobile-source emissions. In 2000, non-road emissions represented 47.7%, 34.2%, 40.6%, and 66.0% of total mobile-source hydrocarbon, CO, NO_x, and PM emissions, respectively (National Research Council 2004). Each of these fractions is expected to increase by 2020 as newer and cleaner on-road vehicles replace older vehicles.

Presumably, non-road emissions would respond to changes in fuel composition, but with a much less well-characterized response, and this was not studied here.

MOVES uses multiple models to account for, for example, changing transportation patterns, fuel types, and temperature effects throughout the year on various time scales; the modeling for this project included daily, monthly, annual, and decadal variability. These models produce results with unrealistic step functions between days, months, and years. To account for this, raw daily MOVES outputs used in empirical models were smoothed using linear models with terms for linear, square, sine and cosine (period of 1 year), and time interaction with cosine, plus an indicator for weekday or weekend (Appendix Figure A.5). This trends model was used to project 2013 emissions.

Mobile Emissions Data Uncertainty

MOVES inputs and internal models each impart uncertainty on the final emissions value; however, research is limited as to the biases that may be present in estimated emissions rates and activities for large fleets made up of multiple vehicle types (such as Atlanta's). Snyder and colleagues (2014) found that differences in VMT estimates from two models translated into +45% to −35% uncertainty in emissions. They noted that uncertainty in the ratios of heavy-duty to light-duty automobiles would affect certain pollutants, such as $PM_{2.5}$ and NO_x, more than others, such as CO. Adjusting the percentage of diesel vehicles from 5.2% to 9.2% on one specific type of roadway in their simulations changed total estimated $PM_{2.5}$ and NO_x by 53% and 29%, respectively. Yazdani Boroujeni and Frey (2014) investigated road grade parameterizations, and Sandhu and colleagues (2014) addressed refuse truck emissions rates. No studies to date have combined uncertainty estimates of each of the models within MOVES to estimate uncertainty across multiple inputs. Hanna and colleagues (2001) estimated uncertainty in various air quality inputs using an expert elicitation and applied ±100% for mobile sources. Using a variety of methods and comparisons — including fuel and mileage-based emissions factors, ground-based observations, satellite inverse modeling, and regional chemical transport modeling — multiple studies have found large variability between modeled and a priori estimates (Dallmann and Harley 2010; Deguillaume et al. 2007; Konovalov et al. 2006; Napelenok et al. 2008; Parrish 2006).

Although these uncertainties are large, multiple studies have shown that modeled emissions capture trends sufficiently as to allow air quality models to provide adequate estimates of air quality and changes over time (Foley et al. 2015a,b; Gégo et al. 2008; Gilliland et al. 2008; McDonald et al. 2012, 2013; Simon et al. 2012). For statistical modeling, capturing trends in emissions changes is more important than the absolute emissions amount. Previous studies have shown that modeled mobile emissions trends match expected trends based on ambient air quality observations (Blanchard et al. 2012; Pachon et al. 2012).

Given previously reported concerns in NO_x emissions estimates, particularly that the estimated emissions from mobile sources are biased high (Anderson et al. 2014; Goldberg et al. 2016; Souri et al. 2016), this study included further analysis using multiple approaches to assess potential bias (discussed in detail in Appendix D). First, trends in CEM and (MOVES-modeled) mobile NO_x emissions, ambient ground-level measurements (at JST), and satellite products were compared at monthly and annual time scales. Each satellite product recorded tropospheric NO_2 columns during its operating period, and the overall time series of the three satellites covered the period between 1997 and 2013 (GOME 06/1996–06/2003, SCIAMACHY 08/2002–03/2012, OMI 10/2004–12/2013) (Boersma et al. 2004, 2011; European Space Agency 2016). The satellite crossing times over Atlanta were approximately 3:00 PM (local time), and these were compared with daily surface concentrations monitored at the same time.

A second approach compares ratios of exhaust components in the emissions to those of ambient observations (hereafter referred to as the "ratio-of-ratios approach"). For example, the ratio $[NO_x{:}CO]_{EMISS}$ /$[NO_x{:}CO]_{OBS}$ should be close to unity if atmospheric losses are similar and remain relatively constant over time and if the relative emissions are estimated correctly (after accounting for background CO). Here, background CO is taken as the concentration at Yorkville, a SEARCH site 60 km northwest of JST. Two major assumptions are at play here: first, in the city, mobile emissions contributions dominate ambient levels of certain species; second, these species have lifetimes that do not appreciably affect the ratios between the source and receptor, or the lifetimes are accounted for in the analysis.

The third approach compares results for different species from the Research LINE-source (R-LINE) dispersion model, a steady-state dispersion model that simulates physical dispersion processes using wind speed, wind direction, Monin-Obukhov length for turbulence, surface friction velocity, and other meteorological parameters to estimate line-source impacts on ambient air quality (Snyder et al. 2013). Ten years of annual average concentration fields for $PM_{2.5}$, CO, and NO_x in the 20-county ANAA were developed using an annual average approach that accounts for the frequencies of different meteorological conditions and emission diurnal changes (Zhai et al. 2016). The emissions used are link emissions based on the

Atlanta Regional Commission's 20-county activity-based travel demand model from the year 2010 at 43,712 links and previously used by the Atlanta Regional Commission in the Atlanta Roadside Emissions Exposure Study (D'Onofrio et al. 2016). Emissions were estimated using MOVES emission factors changes relative to a 2010 base year with Atlanta traffic volume and speed and vehicle fleet composition information. For $PM_{2.5}$, the R-LINE estimates are calibrated to mobile-source impact estimates that are estimated using chemical mass balance with gas constraints based on observational data. Regression approaches were explored in linear and log-transformed forms using the jackknife resampling method (Sahinler and Topuz 2007), which estimates regression parameters with each available observation data point withheld one at a time. In total, available observations in 10 years were used at three sites for $PM_{2.5}$, five sites for CO, and seven sites for NO_x.

In a fourth approach, NO_x concentrations and O_3 sensitivities to mobile NO_x were estimated using CMAQ in July 2011 for two cases: a base case with actual emissions and an adjusted case with mobile emissions reduced by 50%. Modeled and observed concentrations were compared. In addition to assessing potential biases, this exercise tests both the impact of the magnitude of NO_x emissions on O_3 levels and the model's ability to capture nonlinearities in the relationship of concentrations to emissions.

Other Emissions Sources

While EGU and mobile sources contribute a large fraction of air pollution emissions to the southeastern airshed, other sources are important as well, and many have undergone concurrent emissions changes caused by regulatory and other changes (Appendix Figure A.1). Source apportionment studies have shown the importance of multiple sectors, including other (nonutility) industry, biomass burning, road dust, and meat cooking (Balachandran et al. 2012; Chen et al. 2012; Hu et al. 2014). An estimate of changing industrial emissions (IND) of $PM_{2.5}$ (taken from the 2002, 2005, 2008, and 2011 National Emissions Inventory, and interpolated between years [U.S. EPA 2016c]) was included in the empirical modeling for this work, but the other sources were not. Other sources are, however, captured in the CTM.

Assessment of Impacts of Regulations on EGU Emissions

Variability in EGU emissions is due to numerous factors, including fuel type, plant size and age, meteorology, economic activity (both relating to fuel costs and electricity demand), population shifts, and distribution efficiency. With these factors in mind, utilities assess many drivers that determine which controls to install, plants to build, or plants to shutter, including compliance, electricity cost to the consumer, avoidance of New Source Review, relationship with plant co-owners, and profit. In Georgia, each of these serves as a point of interest to the Public Service Commission, which approves Georgia Power's pricing structure and major facility changes. Georgia Power must periodically submit an Integrated Resource Plan detailing the company's reasoning for any changes (Georgia Power Company 2007, 2013). While the Integrated Resource Plan provides some guidance regarding the reason for closing or retrofitting plants, these actions may be taken in response to multiple drivers. The analysis here used emissions factors and dates of required compliance as reported in national and Georgia rules, supplemented by Georgia Power Integrated Resource Plans, to attribute emissions changes to specific regulations. A more detailed discussion of decision-making processes in the utility industry is provided in Additional Materials 1 (available on the HEI website).

Gégo and colleagues (2007) estimated emissions changes based on usage. First, an average emissions ratio (ER_{y^*}) was developed for base year y^*. For ANAA emissions, y^* = 1995, and for regional emissions, y^* = 1997. Annual ER was preferred over daily ER to yield a more consistent counterfactual:

$$ER_{y^*} = \left\langle \frac{E^{EGU_{ACT}}\left(d, y^*\right)}{L^{EGU_{ACT}}\left(d, y^*\right)} \right\rangle_{y^*}. \quad (1)$$

Here, $E^{EGU_{ACT}}(d, y^*)$ is the daily emissions in year y^* (in tons), and $L^{EGU_{ACT}}$ represents the gross load (MW-h). The brackets signify the average over the entire year. It is assumed that, without controls, ER would remain constant at the value of ER_{y^*}. Daily counterfactual emissions were calculated by multiplying ER_{y^*} by each day's load:

$$E_{d,y}^{EGU_{ALL}} = ER_{y^*} \times L^{EGU_{ACT}}(d, y^*). \quad (2)$$

The fact that Georgia Power (which operates a majority of the EGUs in the state) must show that a control technology is necessary in order to argue for cost recovery from increased rates on customers, allows for the assumption that all reduced emissions between the actual case and the EGU_{ALL} case correspond to a combination of control policies and improvements in efficiency. The approach for estimating emissions in each case, then, involved splitting the difference between $E^{EGU_{ALL}}$ and $E^{EGU_{ACT}}$ into different control programs.

There are two options for tracking electricity demand from the utility plants of interest: heat input and load.

Both are a measure of the demand on each plant, and emissions factors from both are functions of the type of fuel. Three major factors contribute to changing emissions: controls, increased efficiency (less fuel burned for same power output), and reduced demand. An emissions factor of NO_x per load changes over time because of both controls and increased efficiency. Changes in an emissions factor of NO_x per heat is less dependent on efficiency. For example, if 50% of the load were converted from coal to natural gas and controls were put on the coal, the NO_x per load ratio would change because of controls and increased efficiency, while the NO_x per heat ratio would change only because of the controls. Here, fleet-wide averages (such as those from the two sectors used in the current study, Atlanta-area EGUs and regional EGUs) are a better approach than individual plants because utilities will dispense load among their facilities to produce electricity as cheaply as legally possible (Additional Materials 1).

The value of $E^{EGU_{ACT}}$ as a fraction of $E^{EGU_{ALL}}$ (Appendix Figure B.1) can be interpreted as the fraction of emissions reductions achieved since y^*. Lines were fit to this fraction based on known dates of regulatory compliance and observed trends in the data. These lines were then converted back to emissions amounts, and emissions reductions were interpolated to the end of 2013 (it is assumed that older control programs would have continued in cases when newer programs replaced them). More detail is provided in Additional Materials 1 on the specific dates used for each control scenario.

Assessment of Impacts of Regulations on Mobile Emissions

Mobile sources have unique features that complicate the estimation of counterfactuals. First, the ANAA fleet is made up of vehicles of varying ages, manufacturers, types, and fuel types that are subject to different regulations. The counterfactual mobile emissions scenarios required alternative specifications within and outside of MOVES to estimate counterfactual emissions. MOB_{IM} scenario emissions were estimated by clearing the IM table from the input database — functionally, this equates to eliminating the IM program — and rerunning MOVES for the entire period of interest. IM was required beginning in 1993 in a 13-county subsection of the ANAA; the model was rerun for this subset.

The other scenarios required a different approach, and the following explanation requires a clarification of the terms. Model year is associated with when the vehicle first entered the vehicle population, and emission year (EY) is the year of interest. Emissions factors for automobiles of a certain model year generally increase with each passing EY at varying rates depending on the vehicle type, fuel type, and other factors according to models within MOVES. Because of updates to automobile engines, changes in fuel composition, and other changes, cars with later model years tend to have lower emissions factors as well.

For the MOB scenario (and related EGUMOB scenario), emissions factors for EY 1993 were applied to all future years. This assumes that no changes were made to vehicles produced after 1993 that would impact their emissions rates. The emissions factors were assigned by pollutant, fuel type, process type (e.g., running exhaust, refueling displacement vapor loss), source type (e.g., passenger car, motorcycle, transit bus), month, day (weekend or weekday), and model year. For full lists of subsectors assigned emissions factors, see the MOVES User Guide (U.S. EPA 2012a). In EY 2000, for example, a model-year 1998 gasoline-powered passenger car is assigned the complementary emissions factor for a model-year 1991 gasoline-powered passenger car in EY 1993. This approach was corroborated by running MOVES with the Rate of Progress option, which models a scenario with no C by applying 1993 emission rates to all vehicles after this year. Results from the two approaches were identical.

The approach for the MOB_{GSP} and MOB_{DSP} is similar; the above method for applying emissions factors for previous EYs was applied to only vehicles that used the fuel type of interest. In the MOB_{GSP} scenario, emissions factors for EY 1999 were applied for all future years, and in the MOB_{DSP} scenario, emissions factors for EY 2005 were applied for future years.

AIR QUALITY

This section will discuss the sources of empirical air quality data and associated uncertainty. Further, it will describe methods for meteorological detrending, empirical and deterministic air quality modeling, and uncertainty propagation from regulation to emissions to air quality.

Air Quality Data

The majority of the air quality and meteorology data used for this study are from JST (33.78°N, 84.42°W), a SEARCH site located near downtown Atlanta (Blanchard et al. 2013). For gaseous species and meteorological variables, hourly measurements were converted to daily metrics using averaging times that vary between pollutants based on air quality standards and methods used in previously published literature (Table 3; see also detailed discussion in Henneman et al. 2015).

In some cases, when data at JST had significant numbers of missing measurements in a row, meteorological data were replaced by observations at Hartsfield-Jackson International

Table 3. Daily Pollutant Species and Meteorological Metrics (1999–2013) from JST Converted from Hourly Measurements

Species	Metric	Period	Exclusion Criteria[a] (for daily aggregation) (hr)	Days[b] (n)	Days[b] Missing (%)
Gaseous (ppb)					
O_3	Maximum of 8-hr mean	12 AM–11 PM	—	5,396	1.51
NO_x	Daily mean	11 AM–7 PM	≥ 5/9	5,073	7.05
CO	Daily mean	11 AM–7 PM	≥ 5/9	5,270	3.45
SO_2	Daily maximum	12 AM–11 PM	≥ 12/24	5,329	2.37
Particulate (μg/m³)					
$PM_{2.5}$	Daily mean	12 AM–11 PM	—	4,975	8.83
$SO_4{}^{2-}$	Daily mean	12 AM–11 PM	≥ 12/24	5,151	5.62
$NH_4{}^+$	Daily mean	12 AM–11 PM	≥ 12/24	5,039	7.67
$NO_3{}^-$	Daily mean	12 AM–11 PM	≥ 12/24	5,002	8.34
EC	Daily mean	12 AM–11 PM	—	3,825	29.8
OC	Daily mean	12 AM–11 PM	—	3,821	29.9
Meteorology[c]					
SR (W/m^2)	Daily total	12 AM–11 PM	Any, 7 AM–6 PM	4,950	9.29
SR^M(W/m^2)	Daily maximum	12 AM–11 PM	—	5,398	1.11
T^m(°C)	Daily mean	11 AM–3 PM	≥ 3/5	5,368	1.66
T^M(°C)	Daily maximum	12 AM–11 PM	—	5,425	0.62
WS (m/sec)	Daily mean	11 AM–3 PM	≥ 3/5	5,284	3.19
WS^{morn} (m/sec)	Morning mean	7 AM–10 AM	≥ 3/5	5,225	4.27
RH (%)	Morning mean	8 AM–11 AM	≥ 3/4	5,363	1.75
RF (1 or 0)	Daily factor	12 AM–11 PM	—	5,459	0

[a] If the exclusion criterion (i.e., fraction of hours in the specified period that are missing) is violated on a certain day, the species is recorded as not available for that day.

[b] There were 5,459 total days.

[c] — = only one measurement needed per day; RF (1 or 0) = rainfall (1 if rain, 0 if no rain); RH (%) = relative humidity; SR (W/m^2) = shortwave radiation; SR^M (W/m^2) = maximum shortwave radiation; T^m (°C) = mean temperature; T^M (°C) = maximum temperature; WS (m/sec) = wind speed.

Airport, which is located about 15 kilometers south of JST. This was used for all of 1999 for wind speed and temperature. Rainfall data are from the airport.

Uncertainty in Air Quality Data

Each of the air pollution concentration and meteorological measurements has associated uncertainty. Information on measurements made at SEARCH sites is described in detail by Hansen and colleagues (2003) and Blanchard and colleagues (2013). For the purpose of this analysis, the uncertainties of these measurements were found to be small compared with uncertainties in other inputs to the modeling.

Meteorological Detrending

The first of two empirical methods employed in this work involved meteorological detrending. The goal of detrending is to remove meteorological effects on air pollution (i.e., effects that are not linked to emissions) in order to identify the effects of emission changes at various time scales (Camalier et al. 2007; Cox and Chu 1993; Kuebler et al. 2001). Once the portion of the signal that corresponds to meteorological fluctuations is removed, the resulting time series is compared to the observations to determine the variability in the signal that is attributable to meteorological variability versus other factors (e.g., emissions changes). The method, including evaluation, is described in detail by Henneman and colleagues (2015), and in brief here.

Long-term time series of pollutant concentrations $C(t)$ were decomposed into their components:

$$\begin{aligned} ln\left[C(t)\right] &= C^{LT}(t) + C^{S}(t) \\ &+ C^{WH}(t) + C^{STM}(t) + C^{WN}(t). \end{aligned} \tag{3}$$

Superscripts refer to long-term (*LT*), seasonal (*S*), weekday–holiday (*WH*), short-term meteorology (*STM*), and white noise (*WN*). A log transform was used in the decomposition above to ensure that model residuals (defined later) follow assumptions of normality and homoscedasticity (Eskridge et al. 1997; Hogrefe and Rao 2000). Both the *LT* and *S* portions were quantified using a Kolmogorov-Zurbenko ($KZ_{m,p}$) filter, which is a multipass moving average filter that has been used in prior air quality data analyses (Kuebler et al. 2001; Rao and Zurbenko 1994; Zurbenko 1991):

$$Y_i = \frac{1}{m}\sum_{j=-k}^{k} X_{i+j}\ , \tag{4}$$

where Y is the filtered time series, X is data to be filtered, k is the number of iterations of a moving average with window length m, which is equal to $2k + 1$, p is the number of passes, and i is the data point reference. Kolmogorov-Zurbenko filters are useful for manipulating observation data because of their versatility in cutoff frequency and ability to handle missing data (Eskridge et al. 1997). Version 3.0.0 of the *kza* package in R was used for this analysis (Close and Zurbenko 2013). *LT* was calculated using a $KZ_{365,3}$ filter and subtracted from the original time series. Next, a $KZ_{15,5}$ filter captured annual variability, and the output was averaged by day-of-year to produce the *S* signal, which was also subtracted. These two steps were applied to observations of meteorological variables as well (except for rainfall, which was represented as a daily factor variable).

After the *LT* and *S* signals were removed from the concentration and meteorology time series, all that remained were values with variability on time scales of less than about three months (or Δs). The concentration Δ terms were regressed against ΔT^m, $(\Delta T^m)^2$, $(\Delta T^m)^3$, ΔWS, ΔRH, RF, $T^M*\Delta T^m$, $T^M*\Delta RH$, ΔT^m_{-1}, ΔWS_{-1}, ΔRH_{-1}, RF_{-1}, ΔT^m_{-2}, ΔWS_{-2}, ΔRH_{-2}, RF_{-2}, weekday indicators, weekday indicators multiplied by maximum temperature, month indicators, and holiday indicators. WS is wind speed, T is temperature, RH is relative humidity, RF is rainfall, m is mean and M is maximum. Subscripts of −1 and −2 represent lags of 1 and 2 days. The portion of the signal attributed to weekday and holiday variables is the *WH* component of Equation 3, the portion attributed to meteorological variables is the *STM* component, and the model residuals are attributed to *WN*. The *STM* signal was subtracted from the original time series of each pollutant, yielding a detrended time series independent of influence from short-term meteorological variability.

Because of O_3's application as a marker for photo-oxidants in the atmosphere, the detrended O_3 components can be used to measure the emissions-independent photo-oxidative state of the atmosphere (PS^*). PS^* was calculated by summing the *S* and *STM* terms from the O_3 detrending and subtracting the mean. This term is important in the empirical models for capturing emissions–air quality relationships in cases where emitted pollutants undergo chemical transformations in the atmosphere, such as emitted NO_x and VOCs producing O_3 in the atmosphere. The code is available at *https://github.com/lhenneman/searchAQ*.

Empirical Air Quality Modeling

While meteorological detrending allows for detailed control of the meteorological impact on observed air pollution concentrations, it is of interest to directly link emissions to ambient air quality. Linear regression models with daily concentration as the response and emissions and meteorological indicators as the covariates were developed for each of the ten species of interest (Table 4). The method extends the monthly $PM_{2.5}$ modeling using EGU emissions taken by Harrington and colleagues (2012) by using daily values for multiple pollutants and including mobile emissions.

An original list of the covariates for each species model was selected based on literature results in pollutant sensitivities studies (Table 4, first column) (Blanchard and Hidy 2005; Cohan et al. 2005; Liao et al. 2008; Seinfeld and Pandis 2006; Xing et al. 2011). These include emissions terms (*EGU*, *REG, MOB*, and *IND*) discussed above, including their interaction with PS^* and interactions between NO_x emissions terms and the ammonium nitrate dissociation constant (k_{NO_3}) from Mozurkewich (1993). Terms for interactions of emissions with k_{NO_3} were included to account for enhanced NO_3^- at only the coldest temperatures, and more details of its calculation can be found in the study by Henneman and colleagues (2017a). Interactions between pollutant species were generally avoided, except for the case of mobile NO_x and VOC emissions. Four meteorology indicators were chosen based on the results of the detrending and included in each model: wind speed, temperature, relative humidity, and rainfall. Each of these was scaled by its mean and normalized by its standard deviation. Meteorological terms in the empirical model are distinct from the STM component in the detrending model, as these terms control for variability in pollutant concentrations that is not attributable directly to emissions in the model, whereas the STM component is estimated independently of emissions. The original model for each species is:

$$\begin{aligned} C_i = {} & \beta_0 + \beta_E(\boldsymbol{E}) + \beta_{E\times PS^*}(\boldsymbol{E}\times PS^*) \\ & + \beta_{E\times k_{NO_3}}(\boldsymbol{E}\times k_{NO_3}) + \beta_M(\boldsymbol{M}) + \in_{\boldsymbol{i}}, \end{aligned} \tag{5}$$

Table 4. Covariates Included in Each Species Model, Denoted by Dots[a]

	O_3	NO_2	CO	SO_2	$PM_{2.5}$	SO_4^{2-}	NH_4^+	NO_3^-	OC	EC
$R2$	0.67	0.45	0.37	0.25	0.42	0.45	0.39	0.45	0.26	0.29
Intercept	⊙	⊙	⊙	•	⊙	⊙	•	⊙	•	⊙
$E_{EGU}^{NO_x}$	•	⊙					⊙	⊙		
$E_{EGU}^{NO_x} \times PS$ *		•			•					
$E_{EGU}^{NO_x} \times k_{NO_3}$								•		
$E_{REG}^{NO_x}$	⊙	•	⊙			⊙	•	⊙		
$E_{REG}^{NO_x} \times PS$ *	•	•	⊙		•	⊙			•	
$E_{REG}^{NO_x} \times k_{NO_3}$								•		
$E_{EGU}^{SO_2}$				⊙		•				
$E_{EGU}^{SO_2} \times PS$ *				•	⊙	⊙		⊙		
$E_{REG}^{SO_2}$				⊙	⊙	⊙	⊙			
$E_{REG}^{SO_2} \times PS$ *				⊙	⊙	⊙	⊙	⊙	⊙	
$E_{MOB}^{PM_{2.5}}$					⊙					
$E_{MOB}^{PM_{2.5}} \times PS$ *										
$E_{MOB}^{NO_x}$	•	⊙					⊙			
$E_{MOB}^{NO_x} \times PS$ *	⊙	⊙								
$E_{MOB}^{NO_x} \times k_{NO_3}$								⊙		

Table continues next page

[a] Circled dots represent statistically significant (P <0.05) regression parameters.

where i is the ambient species, the β's are vectors of the desired regression coefficients, $\boldsymbol{E}$ is the matrix of emission variables, $\boldsymbol{M}$ is the matrix of meteorology variables, and $\in$ is the vector of model residuals. Covariates were removed one at a time in a backward selection, maintaining those that were statistically significant ($P < 0.05$). In some cases, covariates that were deemed physically relevant were kept in the model even if they were not statistically significant. An example of this is EGU NO_x emissions in the O_3 model. Meteorological covariates were kept in all models.

Two important aspects of the overall analysis stem from the empirical modeling. The first is the source-specific ambient air quality sensitivities. These are calculated by summing the total contribution of each source to the modeled concentration. For instance, the $PM_{2.5}$ model included mobile emissions terms $E_{MOB}^{PM_{2.5}}$ and $E_{MOB}^{VOC} \times PS^*$. The total mobile sensitivity, therefore, is the sum of the products of these two emissions terms and their respective regression coefficients. The terms *sensitivity*

Table 4 (*Continued*). Covariates Included in Each Species Model, Denoted by Dots[a]

	O_3	NO_2	CO	SO_2	$PM_{2.5}$	SO_4^{2-}	NH_4^+	NO_3^-	OC	EC
E_{MOB}^{VOC}							$\odot$			
$E_{MOB}^{VOC} \times PS*$					$\odot$					
$E_{MOB}^{SO_2}$				$\odot$			$\odot$			
$E_{MOB}^{SO_2} \times PS*$							$\odot$			
E_{MOB}^{CO}			$\odot$							
$E_{MOB}^{CO} \times PS*$			$\odot$							
E_{MOB}^{EC}										$\odot$
$E_{MOB}^{EC} \times PS*$										$\odot$
E_{MOB}^{OC}									$\odot$	
$E_{MOB}^{OC} \times PS*$									$\odot$	
$E_{MOB}^{NO_x} \times E_{MOB}^{VOC} \times PS*$	$\odot$									
$E_{IND}^{PM_{2.5}}$									$\odot$	
WS	$\bullet$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$
Temperature	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$
RH	$\odot$	$\odot$	$\odot$	$\odot$	$\bullet$	$\odot$	$\odot$	$\bullet$	$\odot$	$\odot$
RF	$\bullet$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\odot$	$\bullet$	$\odot$

[a] Circled dots represent statistically significant ($P < 0.05$) regression parameters.

and *contribution* are synonymous when considering a 100% change in emissions, and thus can be the same in that context.

The second outcome of the empirical modeling is daily counterfactual time series. These are estimated for each of the scenarios in Table 1 by replacing the actual emissions in each of the species models with counterfactual emissions specific to each scenario. In the EGU_{ALL} scenario, for example, only EGU emissions in each species model are replaced by their counterfactual.

Chemical Transport Modeling

CTMs use numerical methods to estimate air pollution transport and chemistry across a gridded domain. They provide a number of benefits, including detailed spatial and temporal coverage and the ability to estimate concentrations and sensitivities to model inputs. CTMs are used for both research (e.g., Souri et al. 2016; Wang et al. 2010; Xie et al. 2011) and regulatory (e.g., Georgia EPD 2009; U.S. EPA 2005a,b) applications.

Simulations were performed for the eastern continental United States (details in Henneman et al. 2017b). The U.S. EPA Sparse Matrix Operator Kernel Emissions (SMOKE) platform (version 3.5.1) was used to prepare emissions using the U.S. EPA National Emissions Inventory 2002 and 2011 for the 2001–2002 and 2011–2012 model years, respectively (CMAS Center 2013, U.S. EPA 2014a). The Weather Research Forecast (version 3.6.1) (Skamarock et al. 2008) was used to generate meteorological fields. Simulations of concentration fields were conducted using the CMAQ model (version 5.0.2) (Byun and Schere 2006; U.S. EPA 2014b), with the inline photolysis option and the Carbon Bond (version 2005) chemical mechanism. The simulations were performed over the eastern United States using a grid of 12 km horizontal resolution and 13 vertical layers. The U.S. EPA NEI 2002 and 2011 platforms included biogenic and dust emissions, so inline biogenics and dust were disabled in CMAQ, as well as inline lightning. A fixed profile was used as boundary condition.

The CMAQ decoupled direct method was implemented to examine first-order sensitivities of pollutant concentrations to emissions, which provided insight into the response of pollutants to emission controls. On-road mobile and EGU are the two emission sources focused on here. Sensitivity of concentrations to VOCs, NO_x, primary PM, and the sum of these three pollutants was calculated for mobile and EGU.

In addition to a base run of four years (2001, 2002, 2011, and 2012), we conducted cross-simulations to separate impacts of emission and meteorology. We used 2001 emissions in a simulation using 2011 meteorology (01E:11M), and likewise, 2011 emissions were combined with 2001 meteorology (11E:01M). By comparing with base cases (01E:01M and 11E:11M), we could separate impacts of emissions (11E:01M − 01E:01M) and meteorology (01E:11M − 01E:01M).

Uncertainty in Counterfactual Concentrations

Uncertainty is inherent in each of the steps discussed so far. This section discusses how estimates of uncertainty at each step are estimated and propagated through the methods described above, and more details are given in Appendix B (available on the HEI website). The models for counterfactual concentrations have two major sources of uncertainty: the sensitivities of concentrations to emissions and the estimate of emissions changes from actual to counterfactual (ΔE). It is important in this context to highlight the difference between uncertainty in emissions estimates and uncertainty in the change in emissions due to specific regulations; the latter fits into the scope of this work and is assumed to be larger than the former. Further, capturing the actual emissions magnitude is less important in the type of statistical modeling here than capturing trends at various time scales, from daily to multiyear. In deterministic models, absolute magnitudes of emissions are important because they are required for accurate representation in the interaction chemistry equations in the model.

Counterfactual emissions cannot be evaluated based on observations. Uncertainty, therefore, is assigned using information from the averaging applied to calculate the baseline emissions factor (ER_{y^*}). The uncertainty in ΔE_{EGU} is estimated as a normal distribution with mean and standard deviation of the sampled (ER_{y^*}). Since MOVES emissions factors are based on multiple models and not on measured data, the uncertainty in ΔE_{MOB} is estimated using a ±50% uniform distribution around the daily ΔE_{MOB}, similar to the approach taken in a study by Napelenok and colleagues (2011).

The second major source of uncertainty — the empirical sensitivities of concentration to emissions — stems from uncertainty in the modeled emissions, selection bias in the covariates in the model, and error in the statistical model. To estimate this uncertainty, distributions of each regression parameter in Equation 5 were sampled simultaneously using information in the variance–covariance matrix of the regression.

The two distribution groups, the first from ΔE and the second from the sensitivities, were sampled 5,000 times and used in each of the species models in a Monte Carlo approach to estimate 5,000 alternative counterfactual outcomes for EGUMOB. The approach allows for uncertainty distributions to be estimated for counterfactuals relating to any combinations of regulatory programs.

Aerosol Acidity

Previous researchers have investigated aerosol acidity levels and discussed how they may change concurrently with dramatic changes in emissions (Lipfert and Wyzga 1993; Weber et al. 2016). Acidity (pH) is an important aerosol property and has been related to health impacts (Kleinman et al. 1989; Utell 1985). In the present study, aerosol acidity was investigated using the ISORROPIA II thermodynamic equilibrium model (Ansari and Pandis 1999; Nenes 1998), alongside with CMAQ estimates (see Appendix E for details). To investigate both spatial and temporal variability in empirical pH, ISORROPIA was applied to data at the four urban–rural pairs of SEARCH sites: Georgia (Atlanta-Jefferson Street–Yorkville), Alabama (Birmingham-BHM–Centerville), Mississippi (Gulfport–Oak Grove), and Florida (Pensacola–outlying landing field #8). The model used season-averaged 1-in-3-day sampled $SO_4{}^{2-}$, $NO_3{}^{-}$, $NH_4{}^{+}$, base cations (Mg^{2+}, Ca^{2+}, K^+, Na^+), and chloride ion (Cl^-), which were available from 2008 to 2015 (with some exceptions; see Appendix Table E.S2).

HEALTH IMPACTS

In a typical pre–post accountability study with health endpoints, the exposure of interest is a binary variable representing the presence or absence of an intervention; a measure of effect is estimated for this variable, controlling for potential confounders. Since this study assessed the impact of multiple overlapping pollution-control policies that were gradually implemented over time, we utilized a counterfactual approach that incorporated daily counterfactual ambient pollutant estimates over a 15-year period. Through these counterfactual air quality time series, we estimated counterfactual ED visits in the absence of pollution-control policies and consequently the number and percentage of ED visits that would have otherwise been expected without these policies but did not occur.

A key objective of this study was to contrast daily observed levels of pollutants with their corresponding counterfactual levels to estimate the number and percentage of cardiorespiratory ED visits that were prevented due to pollution-control policies. Figure 3B summarizes the main steps in the health impacts analysis. First, we developed multipollutant health effects models to assess associations between daily observed ambient pollutant levels and daily counts of ED visits for several cardiorespiratory outcomes, controlling for a number of time-varying covariates. Second, from these models we acquired parameter estimates for the associations between pollutants and ED visits. Third, we applied daily observed and counterfactual ambient pollutant levels to these parameter estimates to produce sets of daily risk ratios. These risk ratios describe the observed risk of ED visits compared with the counterfactual risk in the absence of pollution-control policies. Fourth, we incorporated the daily counts of ED visits with these daily risk ratios to estimate the difference in the daily number of ED visits prevented. Finally, these daily estimates of ED visits prevented were aggregated to estimate ED visits prevented by year and over the entire study period. Below, we describe the methods used to estimate the reduction in ED visits prevented in the 5-county Atlanta metropolitan area from 1999–2013.

Hospital ED Data Set

Data were collected on ED visits for the people living in the Atlanta, Georgia, metropolitan area between January 1, 1999, and December 31, 2013. Computerized billing records for patient-level data were pulled from 42 nonfederal acute care hospitals with EDs that serviced the 20-county Atlanta area. A single patient-level data set was created by combining data obtained directly from individual participating hospitals for the period 1999–2004 with a precombined data set from the Georgia Hospital Association for the period 2005–2013. A comparison of data received directly from individual hospitals versus through the Georgia Hospital Association for the 2002–2004 period indicated minimal differences by hospital in visits captured between the data sources (data not shown).

Patient variables included date of admission, the primary International Classification of Diseases 9th Revision (ICD-9) diagnostic code, date of birth, sex, race, and 5-digit residential ZIP code. ED visits were included if the patient ZIP code was located wholly or partially within the 5 primary urban counties of metropolitan Atlanta (Fulton, DeKalb, Gwinnett, Cobb, and Clayton). A similar data set was defined for patients from the larger 20-county Atlanta metropolitan area, which includes these additional 15 Georgia counties: Barrow, Bartow, Carroll, Cherokee, Coweta, Douglas, Fayette, Forsyth, Henry, Newton, Paulding, Pickens, Rockdale, Spalding, and Walton. The majority of the population of the 20-county area (estimated population of 5.4 million in 2013) resided in the 5-county urban core area (estimated population of 3.5 million in 2013).

ED data were thoroughly cleaned prior to analysis. Data cleaning included extensive discussions with hospital personnel regarding resolving ambiguities, standardizing variables, identifying potential inconsistencies, and determining periods of invalid data for each hospital. Hospital indicators were generated to distinguish periods of available and usable data for each hospital.

Health Impact Regression Modeling

We used Poisson generalized linear regression models accounting for overdispersion in order to estimate the joint effect of multiple pollutants on ED visits in a time-series framework. Health outcomes, multipollutant parameterizations, and model covariates are described in the following sections.

Selection of Outcomes of Interest

We created combined categories for cardiovascular and respiratory diseases using subcategories of outcomes shown to be associated with pollution levels in previous studies using the same Atlanta ED data (Krall et al. 2016; Metzger et al. 2004; Peel et al. 2005, 2007; Sarnat et al. 2013). All outcomes were defined using the primary ICD-9 code on patients' ED visit records. The RD outcome group included ED visits for upper-respiratory infection (ICD-9 codes: 460–465, 466.0, 477), bronchiolitis (ICD-9 codes: 466.1, 466.11, 466.19), pneumonia (ICD-9 codes: 480–486), chronic-obstructive pulmonary disease (ICD-9 codes: 491, 492, 496), and asthma (ICD-9 codes: 493, 786.07). The CVD outcome group included ED visits for ischemic heart disease (ICD-9 codes: 410–414), cardiac dysrhythmia (ICD-9 code: 427), CHF (ICD-9 code: 428), and peripheral and cerebrovascular disease (ICD-9 codes: 433–437, 440, 443–445, 451–453). We also considered asthma and CHF separately as two specific outcomes of interest within the combined outcome groups.

As a check for the adequacy of the health impact model for control of confounding, we also assessed ED visits for finger wounds (ICD-9 code: 883) as a health outcome. Since finger wounds would not be expected to be associated with air pollution, a nonzero estimate of the impact of pollution-control policies on ED visits for finger wounds could be indicative of residual confounding.

Multipollutant Model Parameterization

Multipollutant models were created to assess the acute health impact of air pollution by estimating the joint effect of daily ambient pollutant levels on daily cardiorespiratory ED visits, controlling for potential temporal confounders. Measured and counterfactual pollutant estimates were obtained as previously described for nine pollutants of interest: O_3, NO_2, SO_2, CO, $PM_{2.5}$, and the $PM_{2.5}$ components $SO_4{}^{2-}$, $NO_3{}^{-}$, OC, and EC. The purpose of this study was not to evaluate the associations between individual pollutants and health outcomes; rather, it was to assess how changes to broader air quality profiles due to pollution-control policies may have reduced adverse health outcomes. Therefore, we used multipollutant analyses to determine the joint effect of changes in ambient pollutant levels to best account for multipollutant covariation.

We assessed health impacts using four models with different multipollutant formulations. The primary model assessed the health impacts of joint reductions of seven pollutants: $PM_{2.5}$, O_3, CO, SO_2, NO_2, OC, and $NO_3{}^{-}$. In this 7-pollutant model, EC and $SO_4{}^{2-}$ were not included: daily EC levels were highly correlated with OC (r = 0.80), while $SO_4{}^{2-}$ was highly correlated with $PM_{2.5}$ (r = 0.79), and we removed the pollutant that did not contribute to the model. The second approach was to use all nine pollutants in the same model, which would assume that all pollutants were predictive of health impacts even when included in this joint model. The third approach was to look at the health impacts through changes only in $PM_{2.5}$, since this pollutant measure is a mixture affected by a number of different sources and was strongly associated with cardiorespiratory outcomes in our data. Finally, we assessed the health impacts of joint reductions on the five U.S. EPA criteria pollutants included in this study: $PM_{2.5}$, O_3, CO, SO_2, and NO_2.

The temporal relationships between pollutant levels and health outcomes are unlikely to be the same for all outcomes. To maintain consistent methodology with previous research that showed more delayed effects of respiratory outcomes, 3-day moving averages (average of pollutant levels same-day, 1 day prior, and 2 days prior [or lag 0–2]) were chosen a priori as the relevant exposure values for RD and asthma ED visits (Sarnat et al. 2013). For CVD and CHF ED visits, same-day pollutant values (lag 0) were used. We preferred stating an a priori lag structure to mitigate issues of multiple comparisons and the potential for data fishing for the strongest result.

Associations between levels of a specific pollutant and a health outcome may change based on the levels of other pollutants, possibly due to the effects of atmospheric chemistry or due to the synergistic effects of exposure to elevated levels of multiple pollutants. We considered first-order interaction terms between pollutants as potential predictor variables, deciding a priori to include either all pollutant interactions or none of them (as opposed to choosing only those that met some level of statistical significance). Inclusion of these pollutant-interaction terms impacted our estimates, so interaction terms were used in our primary analyses.

We also considered the possibility that dose–response relationships between pollutants and health outcomes may not be linear. To assess this premise, we compared models that included cubic polynomial terms for each pollutant with models that only included linear pollutant terms. Inclusion of the cubic polynomial terms impacted our estimates, so they were retained in our primary analyses. In sensitivity analyses, we also report results from models that did not allow for pollutant interaction, models that did not allow for nonlinearity, and models that did not allow for either interaction or nonlinearity.

Model Covariates

Prior studies have analyzed the association between ambient pollutants and ED visits using the same Atlanta ED data (Darrow et al. 2009; Gass et al. 2015; Krall et al. 2016; Metzger et al. 2004; Peel et al. 2005, 2007; Sarnat et al. 2010, 2013; Strickland et al. 2010, 2011; Winquist et al. 2012a, 2014, 2016). These studies identified important covariates and model parameterizations that were necessary for providing optimal control of potential temporal confounders. All covariates described below were included a priori based on findings from the time-series models used in these previous analyses.

Forty-two hospitals contributed ED data for this study; however, data from each hospital were not necessarily available over the entire study period of January 1, 1999, to December 31, 2013, for various reasons (e.g., hospitals opening or closing, mergers, or lack of data availability or integrity for certain periods) (Appendix Table C.4). Therefore, we included an indicator variable for each hospital taking the value 1 if the hospital contributed data on a given day and 0 if the hospital did not contribute.

Meteorological variables can be an important confounder of the association between air pollution and morbidity, and they can demonstrate nonlinear associations with both exposure and outcome (Schwartz and Marcus 1990). Meteorological measurements from Hartsfield-Jackson International Airport were used for the health impact models, as they were more representative of the overall metropolitan area. Cubic polynomials were included for the 3-day moving average of dew point (Dew) (lag 0–2). Maximum temperature (TempMax) and minimum temperature (TempMin) are important, but including them both in the model over the same periods would introduce excess collinearity, so we used staggered temperature covariates as had been done previously (Winquist et al. 2014). Since associations between temperature and health outcomes were more immediate for maximum daily temperature, the model included cubic polynomials for same-day maximum temperature as well as lagged minimum temperature (lag 1–2). We also included interaction terms between the same-day maximum-temperature cubic polynomials and season (defined as winter = December–February, spring = March–May, summer = June–August, fall = September–November).

We utilized a time-stratified formulation to control for long-term as well as seasonal trends. The underlying idea of this approach is that if the study period were restricted to a given month, long-term temporal trends should not be a source of error. One might conceptualize the time-stratified approach as replicating a study, month after month, and then pooling the results of each month-long study. We controlled for year, month, and weekday (with holidays separate) all as categorical variables, as well as the interaction terms year × month and month × weekday. This was not a true case–crossover formulation, as we did not include interaction terms for year × weekday and the three-way interaction for year × month × weekday. There was little evidence or reason to believe that the effect of day-of-week differed by year, and inclusion of these additional terms drastically increased the number of model parameters and affected model convergence. Indicator variables were also included for other dates that may have unique pollutant or ED profiles (day after Thanksgiving, day after Christmas, and dates of Christmas/Thanksgiving/Veteran's Day/New Year's Day when different from date of a federal holiday).

The inclusion of the year, month, and year × month variables effectively sets up a different baseline for cardiorespiratory ED visits for each of the 180 months of the 15-year study, accounting for temporal confounders not explicitly included in the model. These confounders could include long-term changes such as population growth, demographic changes, trends in health-seeking behavior, changes in average time spent outdoors, or any other unidentified long-term trends. This model formulation also controls for potential seasonal confounders that may affect both ambient pollutant levels and cardiorespiratory ED visits. The weekday, month × weekday, and holiday-related variables add additional control, as human behaviors such as driving patterns, exposure to outdoor air pollution, and health-seeking behaviors may be influenced by these factors.

The overall model had the following form:

$$
\begin{aligned}
&Log\left[E\left(Y_{x,t}\right)\right] = \alpha_x \qquad (6)\\
&+\sum_{j=1}^{i}\left[\begin{array}{l}\beta_{x,j,1}\left(pollutant_{j,t}\right)^1 + \beta_{x,j,2}\left(pollutant_{j,t}\right)^2\\ +\beta_{x,j,3}\left(pollutant_{j,t}\right)^3\end{array}\right]\\
&+\sum_{k=1}^{i}\sum_{l=k+1}^{i}\varphi_{x,k,l}\left(pollutant_{k,t}\times pollutant_{l,t}\right)\\
&+\sum_{m=1}^{42}\gamma_{x,m}hospital_{m,t}\\
&+\sum_{n=1}^{8}\delta_{x,n}weekday_{n,t}+\sum_{o=1}^{6}\zeta_{x,o}holiday_{o,t}+\sum_{p=1}^{15}\theta_{x,p}\left(year_{p,t}\right)\\
&+\sum_{q=1}^{12}\chi_{x,q}\left(month_{q,t}\right)+\sum_{p=1}^{15}\sum_{q=1}^{12}\psi_{x,p,q}\left(year_{p,t}\times month_{q,t}\right)\\
&+\sum_{n=1}^{8}\sum_{q=1}^{12}\vartheta_{x,n,q}\left(weekday_{n,t}\times month_{q,t}\right)+\eta_{x,1}\left(TempMax_t\right)\\
&+\eta_{x,2}\left(TempMax_t\right)^2+\eta_{x,3}\left(TempMax_t\right)^3\\
&+\eta_{x,4}\left(TempMin_{t,lag1-2}\right)\\
&+\eta_{x,5}\left(TempMin_{t,lag1-2}\right)^2+\eta_{x,6}\left(TempMin_{t,lag1-2}\right)^3\\
&+\eta_{x,6}\left(Dew_{t,lag0-2}\right)\\
&+\eta_{x,8}\left(Dew_{t,lag0-2}\right)^2+\eta_{x,9}\left(Dew_{t,lag0-2}\right)^3\\
&+\sum_{r=1}^{4}\lambda_{x,r}\left(season_{r,t}\times TempMax_t\right)\\
&+\sum_{r=1}^{4}\xi_{x,r}\left(season_{r,t}\times TempMax_t^{\,2}\right)\\
&+\sum_{r=1}^{4}\tau_{x,r}\left(season_{r,t}\times TempMax_t^{\,3}\right),
\end{aligned}
$$

where $Y_{x,t}$ is the daily count of ED visits for outcome x on day t in a model with i pollutants. Other variables are for the pollutants (j, k, l), hospitals (n), months (a), years (p), and seasons (r). The Greek letters (α, β, φ, γ, δ, ζ, θ, χ, ψ, ϑ,

η, λ, ξ, τ) are the regression coefficients for the different factors affecting pollutant concentrations. Variable definitions are as follows:

- Pollutant: daily pollutant level
- Hospital: Indicator variable taking the value 1 if the hospital contributed data on a given day and 0 if the hospital did not contribute
- Weekday: eight-category variable with a separate category for federal holidays.
- Holiday: separate non-federal holidays of interest (day after Thanksgiving, day after Christmas, and dates of Christmas/Thanksgiving/Veteran's Day/New Year's Day when different from date of a federal holiday)
- Year: categorical variable for year
- Month: categorical variable for month of year
- TempMax: daily maximum temperature
- TempMin: daily minimum temperature
- Dew: daily mean dew point
- Season: categorical variable for season (winter = December–February, spring = March–May, summer = June–August, fall = September–November)

Period for Fitting of Parameter Estimates

Initial analyses showed that there were stronger observed associations between pollutants and ED visits in the first half of the study period (roughly 1999–2005) compared with the latter half. If the change in the concentration–response parameters over the course of the study period occurred, in large part, as a consequence of pollution-control interventions, then it would introduce bias to estimate the counterfactual number of ED visits in the absence of pollution-control programs with parameters whose values were affected by the pollution-control programs. Conversely, if these changes in estimated associations between pollutants and ED visits were due to extraneous factors (e.g., decreased population susceptibility to air pollution health effects) unrelated to pollution-control policies, then using models fit over the entire study period may be more appropriate. Peters and colleagues (2009) observed changing concentration–response estimates between air pollutants and mortality in Erfurt, Germany; they hypothesized that this could be a result of changing source emissions profiles, which in turn could have been a result of emissions-reducing actions. Since we are uncertain about the cause of these changing associations, any of these possible explanations should be considered purely speculative. For these reasons, we decided to fit health models using data from 1999–2005 as well as models using data for the entire 1999–2013 study period. Results from both methods are presented.

Generating Daily Risk Ratios

The counterfactual model formulation allowed us to estimate outcomes if only the pollutant levels changed but all other factors (e.g., meteorology, temporal trends) remained the same. Thus, in order to estimate outcomes under counterfactual scenarios, we applied the counterfactual pollutant levels while keeping all other parameter values constant.

For each pollution-control scenario, we took the difference between daily counterfactual and daily observed pollutant levels and multiplied that by the appropriate parameter coefficient. This was performed for all linear, quadratic, and cubic pollutant terms, as well as pollutant interaction terms. These values were all summed up and then exponentiated to produce daily risk ratios for each scenario for each outcome:

$$RR_t = exp\begin{bmatrix} \beta_1 \left(PV_{1,t,obs} - PV_{1,t,cf}\right) \\ + \beta_2 \left(PV_{2,t,obs} - PV_{2,t,cf}\right) \\ + \ldots + \beta_n \left(PV_{n,t,obs} - PV_{n,t,cf}\right) \end{bmatrix}, \qquad (7)$$

where RR_t represents the daily risk ratio on day t; $PV_{x,t}$ represents pollutant values (which include raw measured pollutant levels as well as squared and cubed values and two-way products) for pollutant value x on day t; cf is the counterfactual estimate of that pollutant value; obs is the observed value; and there are n total pollutant terms. The β's are regression coefficients for the relationship between log daily ED visits and pollutant values. The risk ratio represents the daily observed risk of ED visits compared with the daily risk of ED visits in counterfactual scenarios: risk ratios below 1 describe protective effects of pollution-control policies.

Estimating ED Visits Prevented by Pollution-Control Policies

To obtain estimates for daily counts of counterfactual ED visits, we divided the daily observed number of ED visits by these daily risk ratios; these daily estimates were then subtracted from the observed daily ED visits in order to produce estimates of the daily number of ED visits prevented. These estimated daily numbers of ED visits prevented were then aggregated to produce estimates for ED visits prevented by season, by year, and for the entire study period. We added together the observed ED visits and the estimates of prevented ED visits to get estimates of all ED visits that would have occurred in the absence of pollution-control policies; the estimated number of ED visits prevented was divided by these counterfactual estimates of ED visits to estimate the percentage of ED visits prevented. These calculations were conducted for every health outcome and every combination of tested model parameterizations. The series of steps utilized in the health impact analysis is depicted in Figure 3B.

Some of the selected air pollution policies did not take effect until partway through the study period, and not all selected policies were fully realized immediately, rather they involved progressively stricter emissions controls or increased implementation over time. Indeed, the automobile fleets continue to reduce per-vehicle-mile emissions. The effect of the air pollution policies on air quality therefore was gradual over time. Air pollution reduction was modest in the early years of the study but substantial in the final years of the study period when the policies were more fully implemented and realized. As such, we would expect the health impacts to be greater in the later years. One of the advantages of our counterfactual approach compared with a pre–post analysis is the way we can account for the gradual implementation of pollution policies. Our primary interest is on the full effect of the policies rather than the particular way they were gradually implemented in Atlanta, and therefore our focus is on ED visits prevented over the final two years of the study (2012–2013), though the impacts of the combined set of policies are also calculated as a function of time. The 1999–2013 health impact estimates and the 2012–2013 health impact estimates are addressing two different questions. One question is: What is the health impact over the period that the health policies gradually became implemented? The other question is: What is the health impact during the period when the health policies were most fully implemented and the benefits realized? If pollution policies are sustained in future years, the second question may be more relevant, but both questions are interesting and addressed here.

Accounting for Uncertainty

Typically, when accounting for uncertainty of air pollution health effects (i.e., constructing interval estimates), that uncertainty results solely from the uncertainty in the estimation of the model parameters representing the health effects. For the health analyses here, we consider two broad layers of uncertainty: (1) the uncertainty in the estimation of the model parameters representing the health effects; and (2) the uncertainty in the estimation of the counterfactual daily time series for each pollutant. We used Monte Carlo simulations to account for the overall uncertainty in the health analyses.

We started with 5,000 sets of estimates of daily counterfactual time series of ambient pollutant levels for every pollutant in the analysis; the methods have been described earlier and incorporate uncertainty in both emissions and the relationship between emissions and ambient pollutant levels. To account for the uncertainty in the estimated health model parameters, we used the observed parameter estimates and their estimated variance–covariance structure to generate 5,000 sets of estimated coefficient values for each health outcome, assuming a multivariate normal distribution.

The 5,000 sets of time series of ambient pollutant levels were linked with the 5,000 sets of coefficient values in order to generate 5,000 sets of daily risk ratios. Daily counts of ED visits were then incorporated in order to produce 5,000 sets of daily estimates of ED visits prevented, which were then aggregated to produce 5,000 total estimates of ED visits prevented through pollution-control policies. We took the 2.5th and 97.5th percentiles to represent the 95% interval estimate, which incorporates the uncertainty of both the estimation of the model parameters representing the health effects and the uncertainty in the estimation of the counterfactual air pollution time series for each pollutant in the health model.

Statistical Programs Utilized

All analyses in the health impact modeling were performed through SAS, version 9.3 (SAS Institute, Cary, NC) and R, version 3.01 (The R Foundation for Statistical Computing 2013) using the following packages: data.table (version 1.9.6), MASS (version 7.3-47), and Cairo (version 1.5-9) (Dowle and Srinivasan 2017; Urbanek and Horner 2015; Venables and Ripley 2002). Descriptive analyses, data management, and preliminary analyses were performed in SAS; additional data management, final analyses, and graphical output were performed in R (R Development Core Team 2012).

RESULTS

EMISSIONS

Between 1995 and 2013, ANAA EGU NO_x and SO_2 emissions decreased, respectively, by 86% and 85%, and between 1997 and 2013, REG EGU emissions decreased by 82% and 83%, respectively (Figure 4, Table 2). Modeled mobile emissions saw large declines of between 61% and 93% for NO_x, SO_2, $PM_{2.5}$, CO, VOCs, EC, and OC (Table 2). While EGU emissions tended to experience large step changes when controls were installed and plants shut down or started up, mobile emissions changes were steadier, as they are driven by fleet turnover replacing older polluting vehicles with new

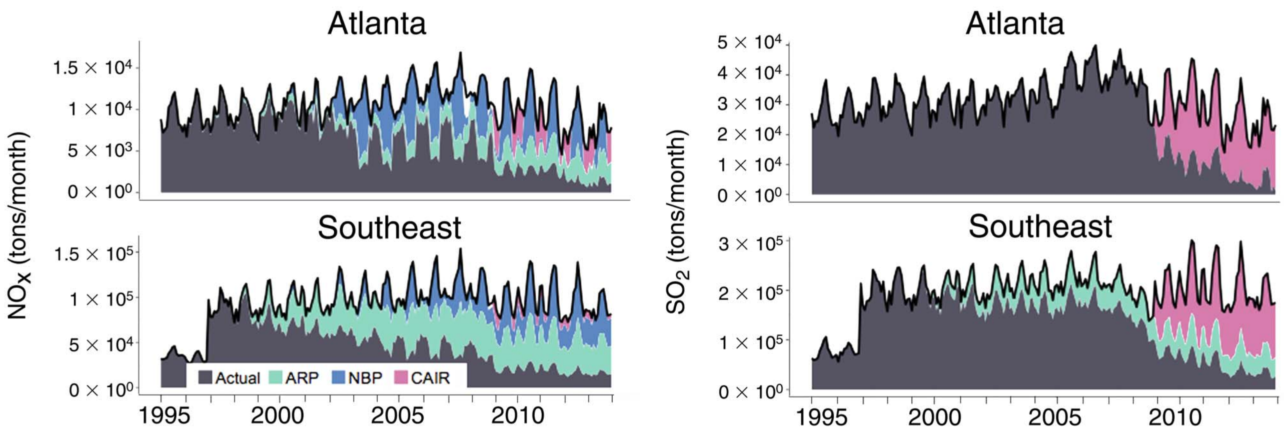


Figure 4. Changes in EGU emissions attributable to specific regulatory actions for NO_x and SO_2 in the ANAA and Southeast Region. ARP = Acid Rain Program; CAIR = Clean Air Interstate Rule; NBP = NO_x Budget trading Program.

cleaner ones. A notable exception is mobile SO_2 emissions, which fell markedly each time gasoline and diesel fuel sulfur limits were tightened in Georgia beginning in 1999.

One pattern of interest in EGU emissions is the summer/winter step functions in years 2000–2008 in the ANAA emissions, and to a lesser extent in the REG emissions. This pattern (and similar step functions) is helpful in determining the effectiveness of controls and regulatory programs. Similarly, SO_2 emissions declined rapidly between 2008 and 2010, when states were preparing for CAIR.

Utility Emissions Responses to Specific Regulations

A comparison between observed EGU (and REG) emissions and the EGU_{ALL} (and REG_{ALL}) scenario shows increasing emissions reductions over time. In the EGU_{ALL} scenario, which corresponds to the bold black line in the plots in Figure 4, counterfactual emissions maintain the expected seasonal pattern (summertime peaks). Counterfactual emissions peaked in 2007, when demand was highest. After 2007, the Great Recession reduced demand. Demand levels have not reached pre-recession levels since, though this may be due partly to increased efficiency in multiple aspects of the electricity industry (production, distribution, and use) (Craig 2016; Levy et al. 2016).

Reductions attributable to specific programs have impacted local and regional emissions differently — individual programs show reductions of varying fractions of the counterfactual. For example, large SO_2 emissions reductions attributable to the ARP were estimated for regional EGUs, but no corresponding reductions were estimated for the ANAA. It is possible, however, that the limited emissions data before 1997 impact this calculation somewhat. After 2008, CAIR had an immediate, large impact on both regional and ANAA SO_2 emissions. CAIR (and the associated $GRAQC_{jjj}$) is the only policy scenario that impacted ANAA SO_2 emissions.

A focus of the NBP and related programs was the reduction of NO_x emissions in summertime, and the results in Figure 4 from 2000–2008 reflect this. Large utilities in the ANAA addressed these regulations by applying controls to large plants that could be switched on during the summer time (an example of this is the selective catalytic reduction devices) and off during the winter (see Table A2-1 in Additional Materials 1). Other approaches included making adjustments to the boilers (which would have remained in effect in the winter time) to reduce NO_x emissions and trading load across units to those that were more efficient, although it is difficult to evaluate if this occurred.

CAIR and the related state-run programs had the effect of extending the NBP-related summertime NO_x emissions reductions to the winter. Since demand during the wintertime in the Southeast is generally lower than in the summer, NO_x reductions attributable to CAIR and related programs are small. SO_2 reductions are the main benefit from these programs, as the largest changes in SO_2 emissions factors occurred after CAIR was promulgated.

Two short periods, in the fall of 2007 and 2008, show lower than expected ANAA EGU NO_x emissions based on the assessment of regulatory programs. It is possible that during these periods one or more plants kept their selective catalytic reactors turned on, but it is not immediately obvious why they would do this (since controls would have to be shown to be necessary to the public service commission for cost recovery). For these reasons, we do not assign emission reductions during these periods to specific regulatory programs.

Mobile Emissions Responses to Specific Regulations

Results of the mobile emissions modeling show large reductions in emissions of NO_x, VOCs, and $PM_{2.5}$ over the period of interest (Figure 5). Modeled ANAA mobile NO_x, VOC, CO, and $PM_{2.5}$ emissions decreased by 78%, 62%, 67%, and 67%, respectively, between 1993 and 2013 (Table 2).

Of the three specific programs investigated, IM programs had the smallest effect on emissions across all years. The impact of IM controls diminished slightly over the period of interest. Gasoline programs had relatively larger impacts on NO_x, VOC, and CO, while diesel programs impacted $PM_{2.5}$ emissions the most. Benefits from the diesel programs did not begin until after 2006, when the Heavy Duty Diesel Rule came into effect.

The sum of the mobile policies does not describe the emissions changes estimated in the MOB_{ALL} scenario because of other policies that were implemented between 1993 and 2013 that are not being examined here. These include the standards established under the 1990 Clean Air Amendments themselves and the Low Emissions Vehicle Program implemented in the late 1990s (U.S. EPA 2011). There are other reasons car makers would change engines that may not be regulation-driven as well (e.g., performance measures).

An important aspect of empirical modeling performed for this work is to assess MOVES' ability to model the emissions of traffic-related pollutants correctly, including capturing the ratios of emitted species (see Appendix D, available on the HEI website). We evaluated the potential bias in mobile NO_x emission estimates in multiple ways, including empirical trend analysis of NO_x using both ground- and satellite-based observations and air quality modeling evaluation of multiple models. The results are somewhat mixed across methods, and some suggest a bias. Briefly, long-term annual average trends of the satellite-based tropical NO_2 vertical column density and ground-level NO_2 concentrations follow the same trends as the NO_x emissions in Atlanta. Observations have a higher monthly-to-annual-average ratio in winter and a slower interannual decreasing trend compared with emissions (Appendix Figure D.2). Additionally, no peak point was found in emissions around 2005 (Appendix Figure D.3). The ratio-of-ratios method suggests a potential underestimate of emissions; ratios of ambient NO_x concentrations to those of other directly emitted species (CO and EC) are lower in magnitude

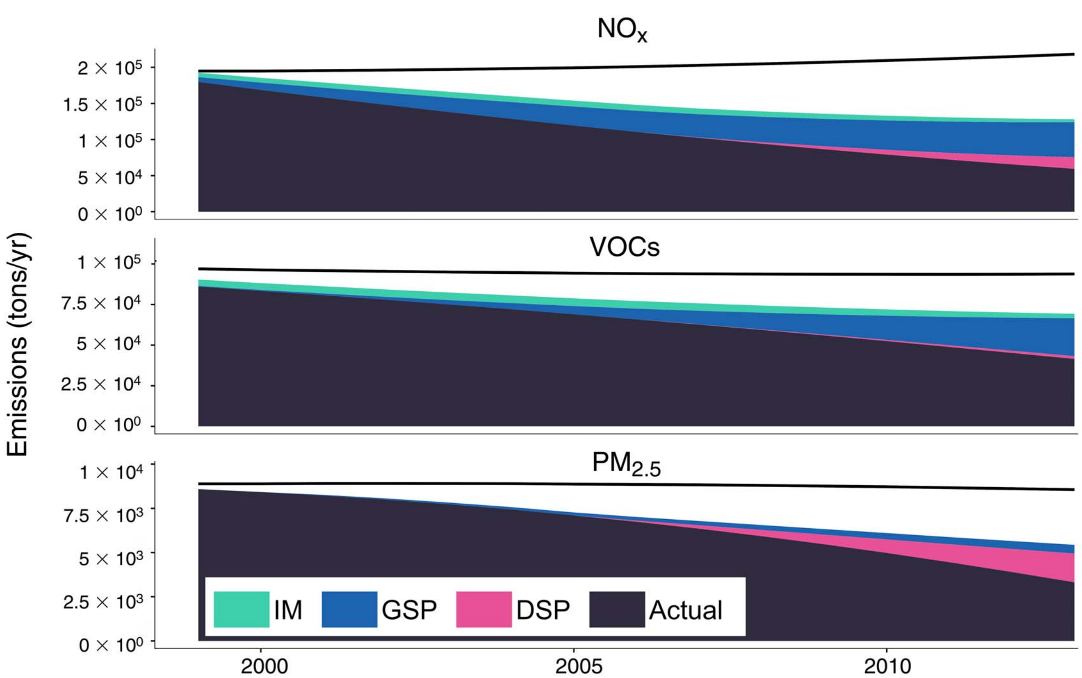


Figure 5. Changes in mobile emissions attributable to specific regulatory actions for NO_x, $PM_{2.5}$, and VOCs in the ANAA. The curved black line at the top of each plot is the scenario that includes all mobile regulations considered in the current study (MOB_{ALL}).

than are similar relationships in modeled emissions. In addition, we do find a high bias in our simulated NO_x concentrations in our air quality modeling at a finer resolution. However, the evidence here is not strong, and inconsistency exists between different approaches.

AIR QUALITY

Urban and rural areas across the Southeast have seen reductions in ambient air pollutant concentrations over the period of interest (Blanchard et al. 2010; Hidy and Blanchard 2015). Atlanta has not been an exception; observed annual averages of all pollutants included in the current analysis except O_3 decreased by 52%–91% between 1999 and 2013 (Table 5). O_3 is the exception (mean annual MD8h O_3 decreased 17.4% between 1999 and 2013) due to its complex chemistry and elevated background levels; however, the number of days with high O_3 levels decreased over the course of the study (Henneman et al. 2015).

Meteorological Detrending/Change in Concentrations over Time

The detrending method for each pollutant was evaluated against a related method detailed in a study by Kuebler and colleagues (2001) using 30 independent holdout tests, each performed by training the models for the signal components given in Equation 3 (i.e., *LT*, *S*, *WH*, *STM*, and *WN*) on 90% of the data. The method described here outperformed the method used by Kuebler and colleagues (2001) in terms of fit (R^2) and root mean square error for most pollutants (Henneman et al. 2015). The discussion focuses on detrending performed on observations in Atlanta; results from Birmingham are presented in Appendix A (available on the HEI website).

For O_3 and $PM_{2.5}$, *S*, *STM*, and *WN* fluctuations make up the majority of the variance. It was found that meteorological fluctuations have a much greater impact on day-to-day variability than annual or seasonal averages. Past detrending work has focused on seasonal or annual adjustments for regulatory purposes (the 8-hr standard in the United States is written for the fourth highest annual MDA8h) (Camalier et al. 2007; Kuebler et al. 2001; Rao and Zurbenko 1994). The focus of previous detrending work has been on O_3; the current study extends methods used in prior analyses to include multiple gaseous and particulate species (Figure 6).

Results of the detrending show the importance of meteorology on daily fluctuations of all species, as the mean meteorological contribution is, for various species, between 11% and 40%, and the standard deviation in *STM* explains between 11% and 46% of the total observed standard deviation (Appendix Table A.5). In general, *STM* as a fraction of observed (in both the mean and standard deviation) is higher in the wintertime, which can be largely explained by lower concentrations of most pollutants considered in the winter.

Table 5. Mean Changes in Ambient Pollutant Concentrations at JST between 1999 and 2013

	1999		2013		Change (%)	
Species	Observed	Detrended	Observed	Detrended	Observed	Detrended
Gaseous (ppb)						
O_3	46.5	45.9	38.4	39.5	17.4	13.9
NO_x	25.8	24.9	10.2	9.3	60.5	62.7
CO	382.9	381.7	183.8	181.7	52.0	52.4
SO_2	18.2	18.1	1.7	1.6	90.7	91.2
Particulate (µg/m³)						
$PM_{2.5}$	19.5	18.5	8.9	9.2	54.4	50.3
SO_4^{2-}	5.6	5.6	1.7	1.8	69.6	67.9
NH_4^+	2.9	2.9	0.8	0.8	72.4	72.4
NO_3^-	1.1	1.0	0.6	0.5	45.5	50.0
EC	2.1	1.9	0.7	0.7	66.7	63.2
OC	5.0	4.8	2.4	2.4	52.0	50.0

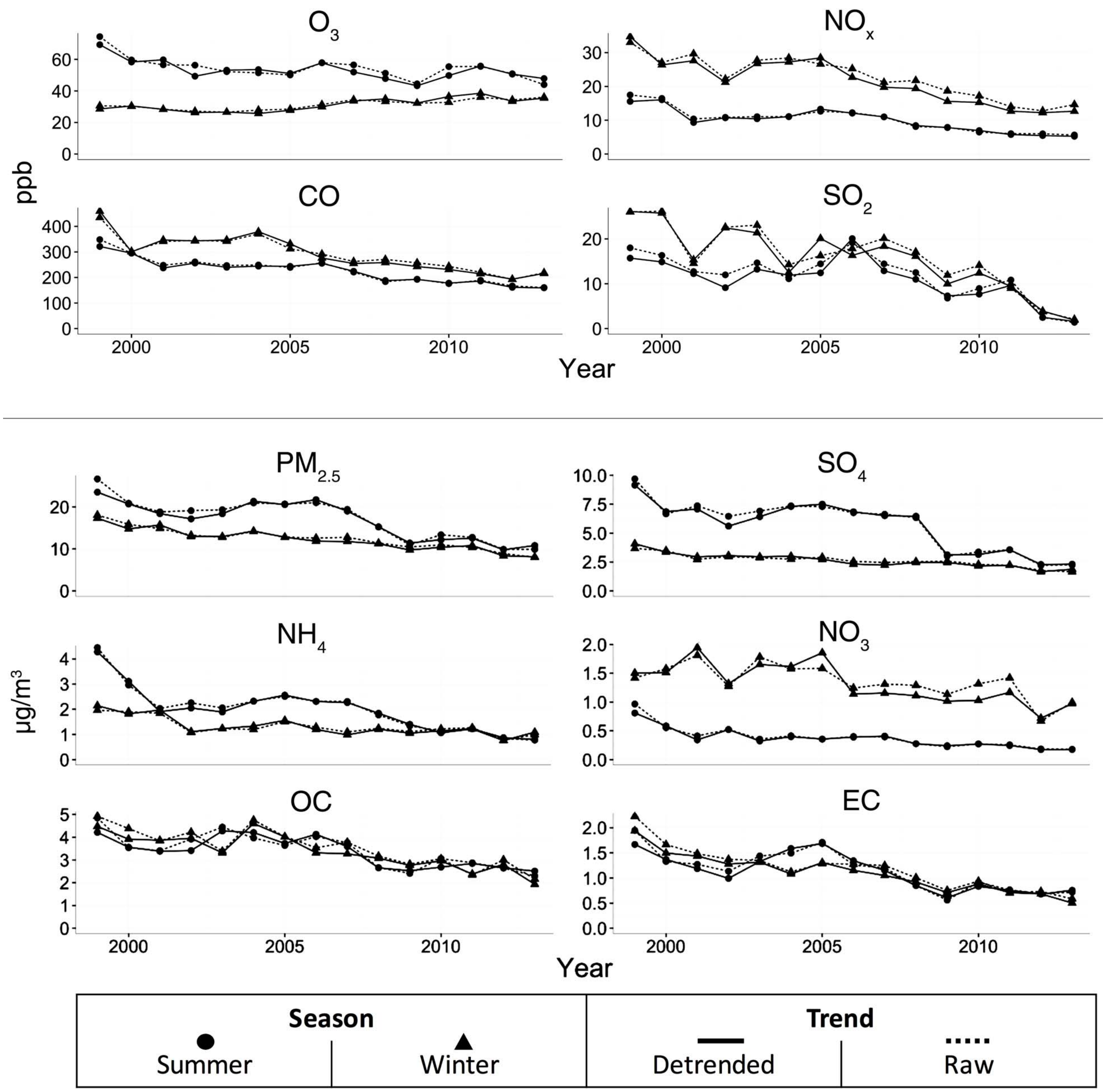


Figure 6. **Observed and detrended summer and winter mean gas (above the horizontal line) and aerosol (below) species concentrations at JST.**

A goal of detrending is to relate changes in observed pollutant levels with known changes in emissions. To this end, observed concentrations were adjusted by subtracting the *S* portion of their signal averaged over different periods (1, 2, 3, 4, 6, and 12 months), and changes in adjacent averaged concentrations were compared using *t* tests to determine if the changes were statistically significant. These changes were then compared to known periods of rapidly changing emissions (e.g., Plant McDonough's retrofit from coal to natural gas [Henneman et al. 2015]). Detrending decreases the averaging time necessary to observe a statistically significant change in air quality associated with emissions reductions for a number of pollutants, particularly NO_x, SO_2, and SO_4^{2-} (Henneman et al. 2015). Trends in O_3 and $PM_{2.5}$ are more difficult to assess at time scales under 1 year, because they vary on multiple time scales.

Empirical Emissions–Air Quality Relationships

The empirical models for each pollutant allow for the estimation of daily source-specific contributions to measured concentrations (Figure 7). These contributions (here, synonymous with sensitivities) take into account both the amount of emissions from a source and chemistry/meteorology (as parameterized by the statistical models) that lead to chemical reactions in the atmosphere. Source-specific contributions to measured concentrations at JST vary at multiple time scales (the monthly averaged plots hide the daily variability, which is considerable, but not as large as the monthly variability; Figure 7). In general, contributions from all sources decreased over the study period, though many show annual variability. Not all species are impacted by all three source categories (as determined by the model selection process). In most species, regional EGU emissions contribute more than local EGU emissions to observed ambient concentrations at JST. These data corroborate evidence presented by Harrington and colleagues (2012), who showed that peak $PM_{2.5}$ contributions come from power plants between 50 and 300 miles away from monitors in the eastern United States.

The *PS** terms in the model allows for much greater seasonality in emissions contributions than the emissions exert on their own. For example, seasonality in the mobile contribution to O_3 shows large positive contributions in the summertime and large negative contributions in the winter, even though mobile emissions do not vary nearly as much seasonally.

Contributions from different sources peak at different times of the year for various reasons. First, source emissions peak in different seasons. Mobile VOC emissions, for instance, tend to be higher in the winter than in the summer, and NO_x emissions peak in the summer. Biogenic

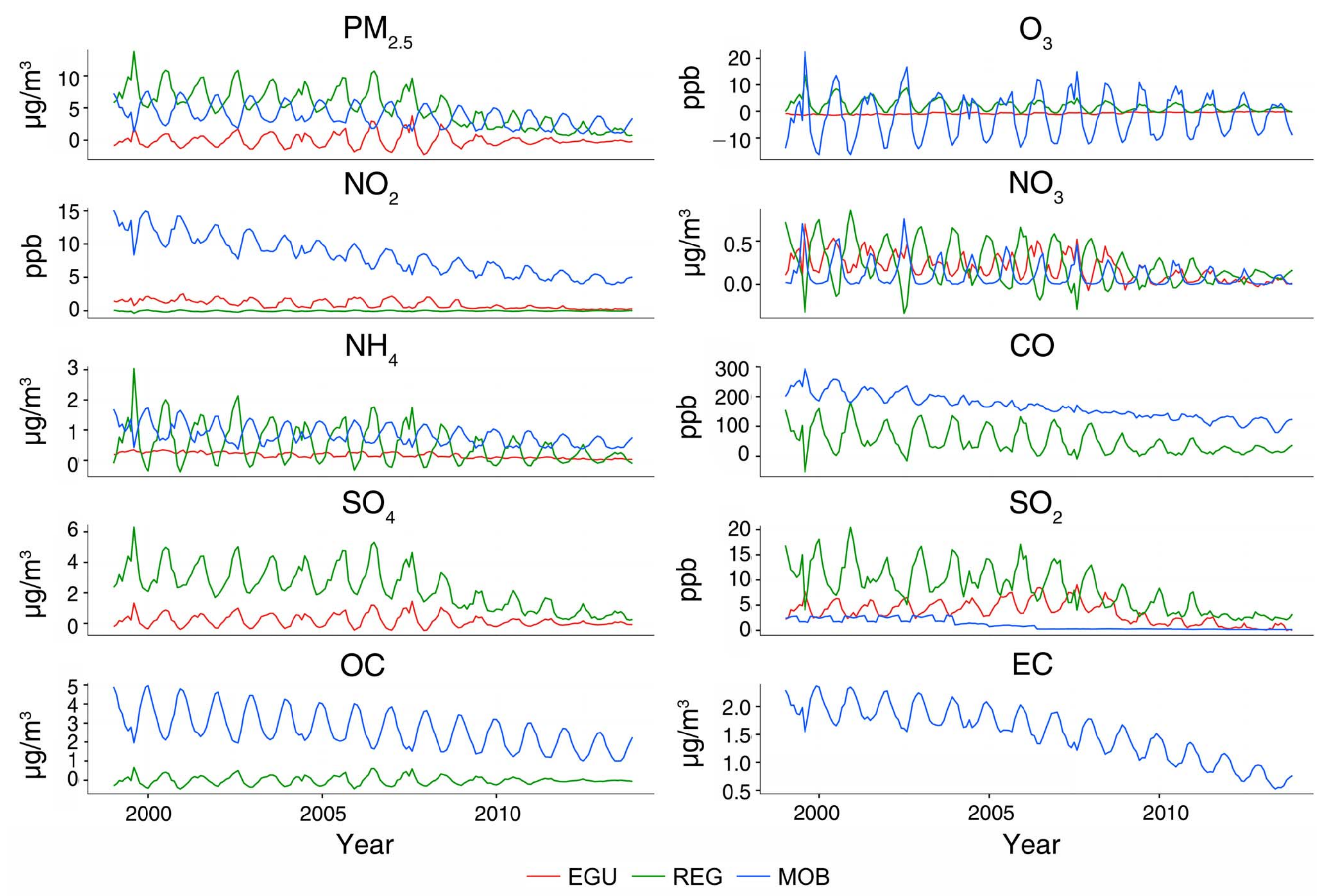


Figure 7. Contributions of three different sources (ANAA EGU, regional EGU, and mobile) to air pollutant concentrations estimated using terms in Equation 5.

VOC emissions peak in the summer. A second reason is the proximity of emissions to the monitoring site. Mobile sources, for instance, emit NO_x mainly as NO near ground level. Here, NO reacts with O_3 to form NO_2 and O_2, thus serving as a sink for O_3 (this reaction is particularly important in the winter, when radicals are less prevalent). NO_x emitted from power plant stacks has more time to process and react in the atmosphere (i.e., photolyze and create ozone). The importance of aging NO_x emissions is reflected in the difference between regional and EGU contributions to O_3 — the regional contribution is consistently more positive than that of the local EGUs (Figure 7).

CTM Operational Evaluation

CMAQ-modeled O_3 shows decreases in the summertime and increases in the winter, and $PM_{2.5}$ shows decreases in both seasons (Figures 8 and 9). CMAQ-modeled $PM_{2.5}$ and its species were evaluated using measurements from U.S. EPA's Air Quality System (U.S. EPA 2017) on a monthly basis. The evaluation uses normalized mean bias and error (NMB and NME) and the correlation coefficient (r) in comparisons with the limit values recommended by Emery and colleagues (2016).

Dynamic evaluation of CMAQ was conducted using JST data. Observed and modeled differences of daily concentrations were calculated between 2001 and 2011 and compared with each other by pairing them by time. While PM mass was measured every day throughout 2001 to 2011, PM species were sampled every three days after 2007. This leads to sparse data sets of PM species for each month, with only a few data points available. We then pooled data by season, that is, by three-month periods, to facilitate analysis for dynamic evaluation of both PM and PM species.

The change of concentration could be positive or negative, while the measured concentration is always positive. The definitions of NME and NMB are therefore slightly modified by taking the absolute value in the denominator to reconcile this difference: that is,

$$NMB = \frac{\sum(P_i - O_i)}{\sum|O_i|} \text{ and} \tag{8}$$

$$NME = \frac{\sum|P_i - O_i|}{\sum|O_i|}, \tag{9}$$

where P_i is the model predicted value, and O_i is the observed value. This modification does not impact operational evaluation and makes NMB and NME from dynamic evaluation more informative. The definition of r is not changed: that is,

$$r = \frac{\sum\left[\left(P_i - \overline{P}\right)\left(Q_i - \overline{Q}\right)\right]}{\sqrt{\sum\left(P_i - \overline{P}\right)^2}\sqrt{\left(Qi - \overline{Q}\right)^2}}, \tag{10}$$

where $\overline{P}$ and $\overline{Q}$ are the mean values of model simulation and observation, respectively. No cutoff was applied for evaluation in the present study.

CMAQ tends to overestimate $PM_{2.5}$ from October to March and underestimate from April to September, consistent with what was observed in a previous study (Simon et al. 2012). EC, OC, NO_3^-, and ammonium contribute to the cold-month overestimation, and OC, SO_4^{2-}, and ammonium contribute to the warm-month underestimation. (See Henneman et al. 2017b.)

CTM Emissions–Air Quality Relationships

CMAQ-decoupled direct method sensitivities for both O_3 and $PM_{2.5}$ show large reductions between the 2001–2002 period and the 2011–2012 period for on-road mobile and EGU sources. Results show seasonal and spatial variability (Henneman et al. 2017b).

Average monthly O_3 sensitivities to EGU emissions are large and positive in the summer of 2001, and almost equally negative in the winter. By 2011, sensitivities of 2.5 ppb or more (positive in the summer) are restricted to areas near large numbers of stationary sources, such as the Ohio River Valley and plants southwest of Atlanta. The pattern for reduction in $PM_{2.5}$ sensitivities is similar (though average sensitivities are positive throughout the year). There is a difference, however, in how these reductions were achieved: the majority of O_3 sensitivity reduction is attributable to lower NO_x emissions, whereas the majority of $PM_{2.5}$ sensitivity reduction is attributable to SO_2 emissions reductions (Henneman et al. 2017b).

Average July O_3 sensitivities to mobile sources are large and positive (on the order of 10 ppb) throughout much of the Southeast in 2001. In cities and around large roadways, the sensitivities are negative. In 2011, the pattern is generally the same, though highways and cities do not stand out as much. In January, sensitivities become less negative between 2001 and 2011. Mobile contributions to $PM_{2.5}$ decrease in summer and winter between 2001 and 2011. Wintertime $PM_{2.5}$ sensitivities to mobile emissions in 2001, which were smaller than summertime sensitivities, became larger than summer sensitivities in 2011. In the southeastern United States in 2011, only city sources are prevalent (rural on-road contributions drop to near zero) (Henneman et al. 2017b).

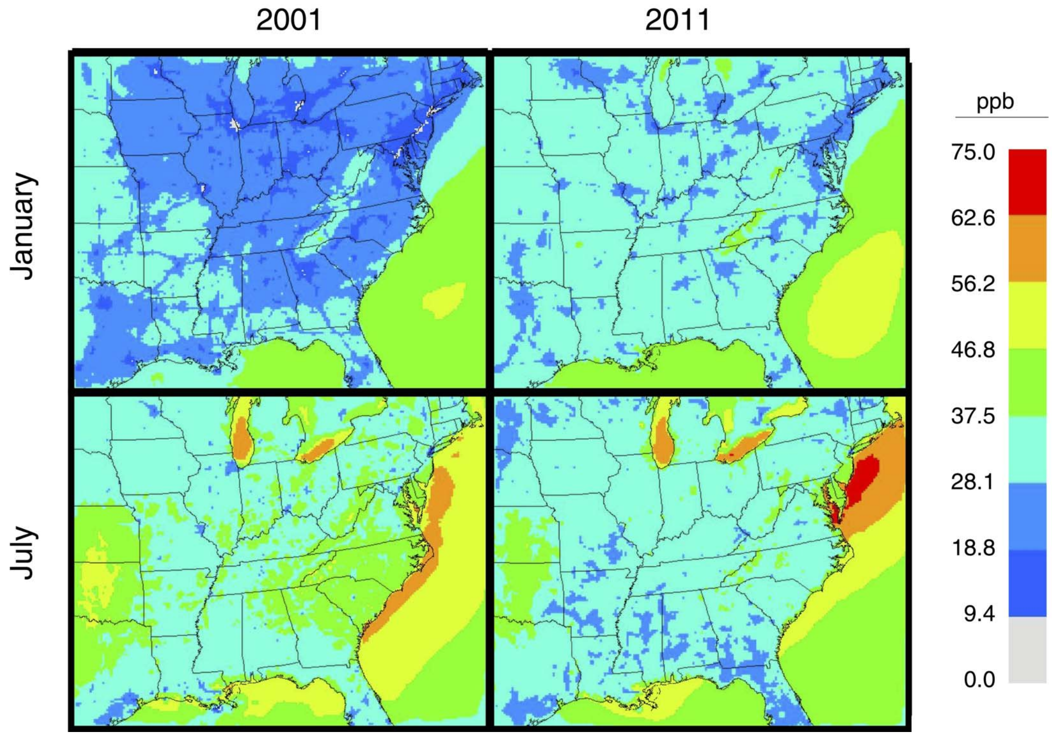


Figure 8. CMAQ-modeled O_3 concentrations in the eastern United States, showing a decrease in the summer between 2001 and 2011 and an increase in the winter.

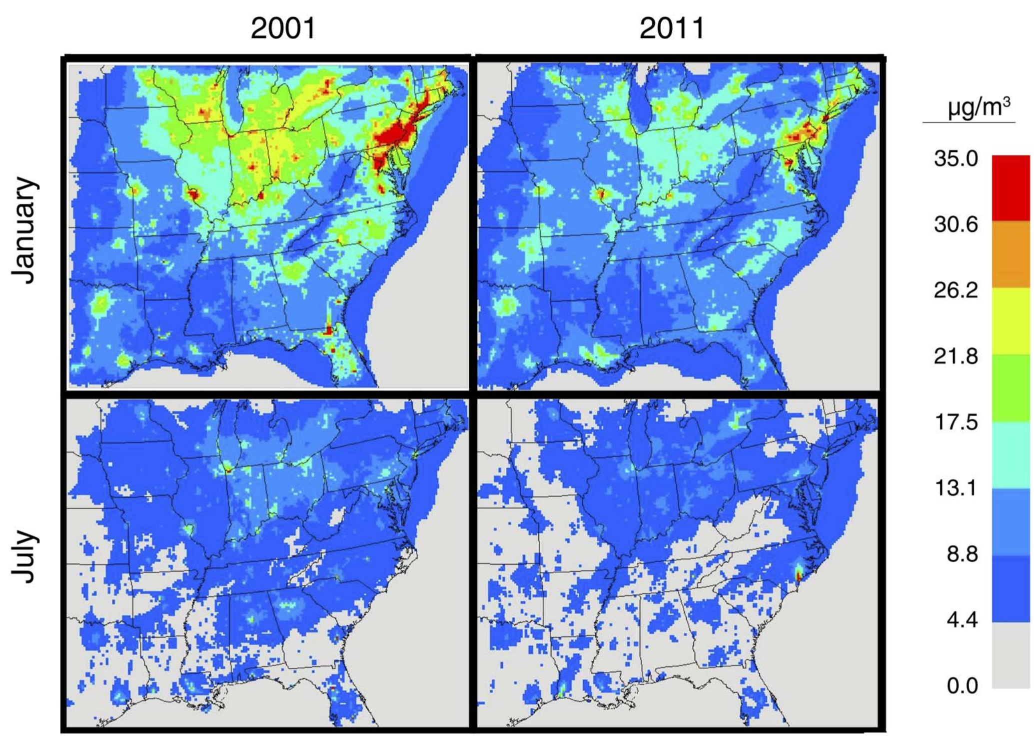


Figure 9. CMAQ-modeled $PM_{2.5}$ concentrations, showing a decrease across all seasons between 2001 and 2011.

CTM Evaluation

NMB, NME, and *r* were used to evaluate the modeled change of O_3, $PM_{2.5}$, and $PM_{2.5}$ species between 2001 and 2011 in comparison with the observed change on a seasonal basis (Henneman et al. 2017b). Modeled $PM_{2.5}$ concentration changes yield a NMB of −13% to 6.6% at JST, indicating that CMAQ captures the seasonally averaged change of PM from 2001 to 2011. In contrast, NME falls in the range of 53% (in fall) to 86% (in summer). This means CMAQ simulates PM concentration changes with a high error, likely due to the NME of $PM_{2.5}$ concentrations being consistently higher than 40% in 2001 and 2011. The *r* is higher than 0.6 in winter, spring, and fall, and is 0.35 in summer. The cause of the small bias and high error is likely the species-specific results; bias in modeled concentrations between species tend to cancel each other out.

Using switched emissions and meteorology between the decades, Henneman and colleagues (2017b) isolated meteorological versus emissions-driven impacts on concentrations from and sensitivities to mobile and EGU sources. Emissions changes had a greater impact than meteorology on O_3 concentration and sensitivity except at the higher end of the distribution of changes. Changes in $PM_{2.5}$ concentrations and sensitivities are relatively insensitive to meteorology, which corroborates findings in the detrending portion of this study. Further results and discussion of the operational, dynamic, and diagnostic evaluations are presented in a study by Henneman and colleagues (2017b).

Empirical and CTM Sensitivities Comparison and Evaluations

Empirical models were assessed by calculating the R^2 and comparing the root mean square error to the mean. Statistically significant relationships with emissions covariates, in addition to sufficiently high model fit parameters (R^2), provide evidence that the models effectively link emissions and air quality. A second evaluation of empirical sensitivities compared them with CMAQ-calculated sensitivities. For most of the pollutants, positive correlations were found between most CMAQ-modeled concentrations and observations and between empirical and CMAQ sensitivities in both the early (2000–2001) and later (2011–2012) periods (Appendix Figures A.10–A.19 and Appendix Tables A.7 and A.8; available on the HEI website). In this section, *EGU** refers to the combined impact of EGU and REG sources. The focus of this discussion is on the 2011–2012 period (since CMAQ performance was generally better for this period), but results from both are presented in Appendix A.

CMAQ-modeled species concentrations were positively correlated with observations in 2011–2012 for all species (see Table A.8). Correlation (R^2) values ranged from 0.09 (NH_4^+) to 0.69 (O_3), and slopes ranged from 0.038 (SO_2) to 1.5 (NO_3^-). The poor fit for SO_2 is expected due to the more elevated SO_2 concentrations' being strongly tied to large point sources and the related plume dynamics. Those plumes are concentrated and near the source have length scales smaller than can be resolved by the current model setup. Further, there is a strong dependence on wind speed and direction — if these are off slightly in the model, the plume impacts a different area.

One takeaway from the analysis is that the slopes between the empirical and CMAQ-simulated concentrations and sensitivities are typically between zero and one and that the intercepts are greater than zero. This means that, on average, concentrations simulated by CMAQ are biased high compared with observations on low ambient concentration days and biased low on high concentration days. This issue — a damped response to various factors, such as emissions or meteorology — has been observed previously (Bencala and Seinfeld 1979; Koo et al. 2015; Simon et al. 2012).

Two nitrogen-containing species (NO_2 and NO_3^-) have negative correlations for both *EGU** and mobile sensitivities. Part of the reason for this may be that CMAQ-estimated NO_2 is highly biased. Observed NO_2 tends to peak in the wintertime, when a lower boundary layer traps more NO_2 close to the surface. CMAQ-modeled NO_2, however, shows little annual variability in 2011–2012 (it shows more annual variability in 2001–2002, with greater peaks in the summertime). *EGU** contributions are small compared with mobile for both CMAQ and empirical models. The NO_2 result is important evidence that mobile emissions may be biased.

Total correlation (*r*) over four years of modeling at JST is between 0.4 (NO_2 and NH_4^+) and 0.82 (O_3) (Appendix Table A.7). NMB ranges from −34.2 (OC) to 324.2 (NO_2). Both NO_2 and CO have very high NMB and NME, which could signify a problem in the emissions inventory (discussed separately). However, CMAQ grid resolution may play a role; the grid cell that includes JST also includes large portions of the downtown Atlanta area, which is home to multiple large highways. These do not have as large of an impact on JST, but would drive the average grid cell concentration of primary pollutants up (Sarnat et al. In press).

NMB's for the majority of the sensitivity comparisons are negative (Appendix Table A.7); however, it should be recognized that the empirically derived sensitivities are not directly observed and are found to be uncertain. This reflects a well-reported tendency for CTMs to underestimate changes in ambient conditions for given changes in emissions. Of note is that NO_3^- has a low NMB for EGU sensitivities, a high NMB for MOB sensitivities, and small negative *r* values for both. Part of this may be explained by the fact that NO_3^- concentrations are small for most of the year, except during cold days in the winter.

A comparison of ozone sensitivities to ambient O_3 concentrations yields a near-linear relationship (see Figure 5 in Henneman et al. 2017a). Empirical and CMAQ-modeled sensitivities compare well across the range of observed O_3 concentrations, though modeled sensitivities are smaller at lower O_3 concentrations.

The comparison between empirical and CTM model results is informative because it juxtaposes sensitivities estimated using methods that are almost completely different (some emissions inputs are similar). A limitation of this comparison is that the emissions used for the empirical methods are not identical to those used by CMAQ.

Air Pollution Concentration Changes Attributable to Regulations

Counterfactual emissions were combined with empirical air quality models to estimate daily counterfactual concentration time series at JST in Atlanta. Monthly averaged differences between observed concentrations and counterfactuals for each scenario show that seasonal variability in emissions and sensitivities contributes to varying benefits over the year (Figures 10 and 11). In addition to seasonal benefits, the majority of the programs show increasing benefits over the period during which emissions programs reach their full implementation (Table 6). Benefits of the EGU scenario (Figure 10) are roughly the sum of the three EGU programs of interest, whereas benefits under the MOB scenario are not (Figure 11). This is because emissions reductions under the three mobile programs do not sum to emissions reductions under the MOB scenario due to the application of controls other than those studied here.

Emissions controls have led to both decreased $PM_{2.5}$ concentrations and changes in its composition (Table 6, Appendix Figure A.20). For instance, OC replaced SO_4^{2-} as the largest fraction of $PM_{2.5}$ at JST between 1999 and 2013. SO_4^{2-} (70%), NH_4^+ (73%), and EC (66%) saw the greatest reductions from 1999 to 2013; NO_3^- (47%) and OC (52%) saw more modest reductions. The counterfactual approach maintained the original concentration distributions of $PM_{2.5}$ species.

The correlation between species and the autocorrelation within species for both the actual and counterfactual concentrations are used to evaluate the approach used to create counterfactuals. Some changes are expected, since each species responds differently to emissions. Of the gaseous species, correlations are maintained in most comparisons, except CO and O_3, which become more negatively correlated (Appendix Table A.6). $PM_{2.5}$ remains similarly correlated to all other particulate species, though there are some differences between species. NH_4^+, in particular, tends to change correlations compared with many gaseous and particulate species. Overall, results of the correlation test provide evidence that, even though each species is

Table 6. Changes in Ambient Pollutant Concentrations at JST between 1999 and 2013, Comparing Observed to the Counterfactual EGUMOB Scenario

	1999			2013		
Species	Observed	EGUMOB	Change (%)[a]	Observed	EGUMOB	Change (%)[a]
Gaseous (ppb)						
O_3	46.5	46.8	1	38.4	39.7	3
NO_2	16.5	16.8	2	7.2	19.9	176
CO	383	433	13	183	383	108
SO_2	18.2	19.5	7	1.7	19.1	1,043
Particulate (µg/m³)						
$PM_{2.5}$	19.5	19.5	0	8.9	19.1	114
SO_4^{2-}	5.64	5.49	−3	1.67	4.81	176
NH_4^+	2.95	3.01	2	0.79	2.68	238
NO_3^-	1.05	1.21	15	0.56	1.32	133
EC	2.05	2.07	1	0.69	2.07	200
OC	4.96	5.05	2	2.38	3.82	61

[a] The Change (%) column compares the counterfactual to the observed in each year; positive numbers represent decreases in measured concentrations attributable to emissions reductions.

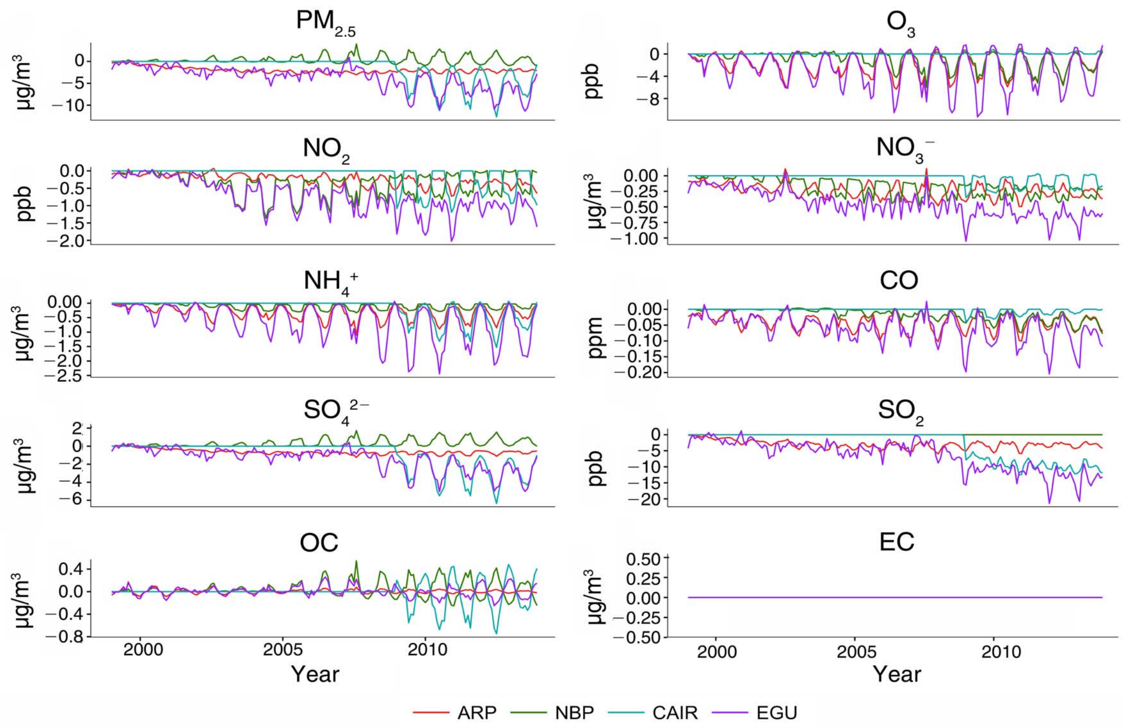


Figure 10. Monthly air pollution changes for each of the EGU emissions scenarios at JST. Negatives represent reduced air pollution concentrations as a results of a specific control policy.

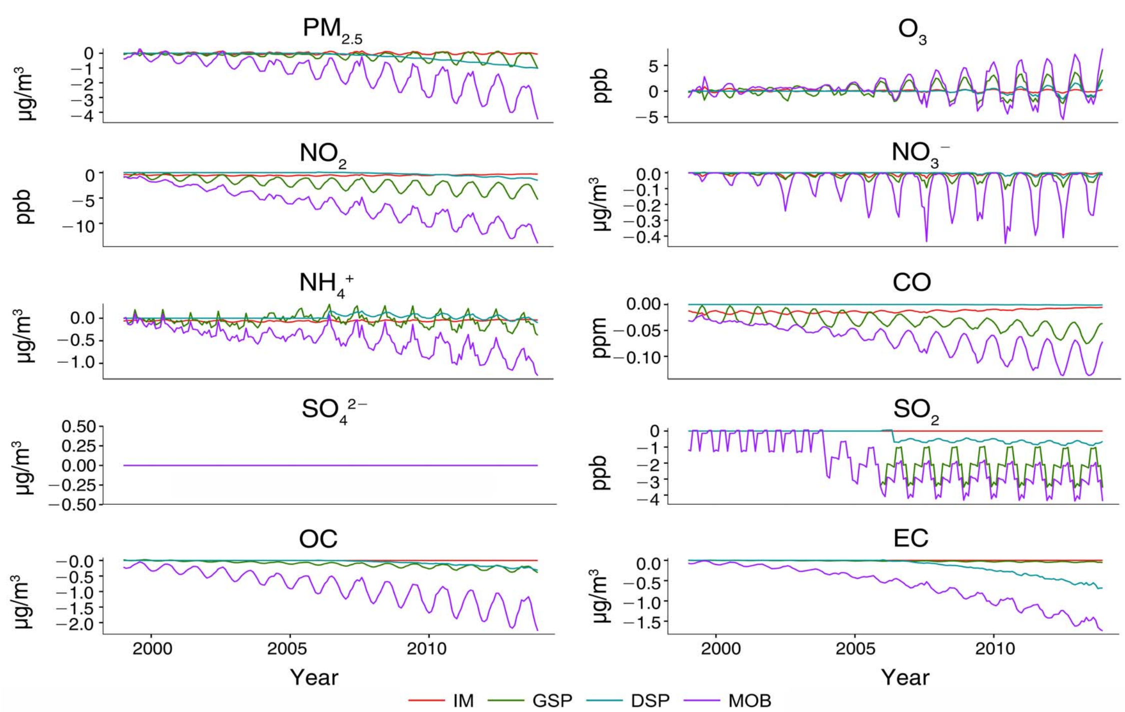


Figure 11. Monthly concentration deviations from the observed for each of the mobile emissions scenarios at JST. Negatives represent reduced air pollution concentrations compared with the counterfactual.

modeled separately, the majority of the species models capture the appropriate relationships between species.

For the majority of species, regulations are linked to reduced concentrations. Notable exceptions occur in species that have consistent negative sensitivities with certain source–species combinations, such as SO_4^{2-} and OC with NO_x. For SO_4^{2-}, for instance, the total sum of all programs caused a net reduction in concentration at JST (Figure 11). Benefits of the NBP, however, are negative. This can be attributed to the suppressing impact of NO_x on the formation of SO_4^{2-} aerosol (Brock et al. 2002; Muller et al. 2009). Mobile emissions were not found to relate strongly to SO_4^{2-} concentrations, so mobile programs did not cause any differences between counterfactual and observed.

Some counterfactual differences yielded changes that vary between seasons. O_3 is the most notable example of this, particularly in relation to mobile emissions (Figure 10). In the wintertime, NO_x titration has the effect of reducing O_3 concentrations. Since the majority of fresh NO_x emissions are NO, ground-level mobile emissions are more efficient at reducing O_3 than are elevated EGU emissions. Further, EGU emissions have more time in the atmosphere to react and convert to NO_2, which photolyzes to form O_3. This is manifested in the reductions modeled from EGU regulation counterfactual differences, which are negative except for during parts of each winter.

Uncertainty in Counterfactual Concentrations

Uncertainty results are presented here for two pollutants in the EGUMOB, EGU_{ALL}, and MOB_{ALL} scenarios (Appendix Figures B.2 to B.11). For the majority of the counterfactuals, contributors to total uncertainties, in order of increasing contribution, are statistical model parameters, EGU emissions changes, and mobile emissions changes. This order may change between species, as it depends on the number of emissions terms from each source that are included in each model. Species counterfactuals see year-round statistically significant differences compared with the observed, with the exception of O_3 (which sees statistically significant decreases in the summer, and increases in the winter). Under the EGU_{ALL} and MOB_{ALL} scenarios, many of the species see statistically significant changes under just one or the other. Most of the pollutants have statistically significant changes under either or both of the EGU_{ALL} and MOB_{ALL} scenarios.

Aerosol Acidity

Aerosol acidity trend analysis in the Southeast led to the important finding that, despite large reductions in acidic aerosol precursors, the pH of fine PM will remain low (see Appendix E for details). Observational data from eight SEARCH $PM_{2.5}$ monitoring sites and the thermodynamic model ISORROPIA II revealed that, despite substantial reductions in SO_4^{2-} levels at all sites (except Pensacola, Florida, which saw a slight increase), pH values remained low. Seasonal average pH over the study period ranged from 0.8 at Oak Grove, Mississippi, to 2.0 at Yorkville, Georgia, bringing the site average, based on trend intercept values, to 1.6. All sites except Centerville, Alabama, exhibited slight increases in pH. Yearly increases were estimated between 0.44% and 3.92% for all other sites except Oak Grove, Mississippi, the only site where the increase was substantially higher and statistically significant. Comparable trends were also observed with the CMAQ results, which also showed low pH values with little to no change in pH between the years 2001 and 2011 (Figure 12).

Total estimated yearly SO_4^{2-} reduction rates at the SEARCH sites range from −0.22 μg/m^3/yr (est: −7%/yr) at Outlying Landing Field #8, in Florida to −0.52 μg/m^3/yr (est: −11%) at Yorkville, Georgia. The observed reductions in SO_4^{2-} levels were accompanied by similar estimated reduction rates in ammonium (−5% to −11%) at the same sites. The downward trends of NH_4^+ and SO_4^{2-} were statistically significant at all sites except Gulfport and Oak Grove, Mississippi, and Pensacola, Florida, meaning that the reduction of SO_4^{2-} and ammonium was seen throughout the region. These results are consistent with similar observations found in another southeastern study by Saylor and colleagues (2015).

Despite reported reductions in NO_x emissions, little change in ammonium nitrate was observed at any of the sites; however, with low pH persisting throughout the region, conditions for NO_3^- formation were not favorable, nor are they likely to change with the slow rates of pH increase exhibited throughout the region.

This analysis suggests that there will be little change in pH compared with the substantial SO_4^{2-} reduction rates, and that has three immediate and important implications: (1) aerosol acidity will remain high, so health impacts associated with low pH will remain (though the exposure to low pH aerosol will be reduced on a mass basis); (2) NO_3^- substitution (i.e., the process by which a near-equal amount of NO_3^- aerosol is formed when the amount of SO_4^{2-} aerosol decreases) will be limited; and (3) aerosol chemistry that is driven by the presence of highly acidic aerosol will continue (e.g., formation of isoprene secondary organic aerosol [SOA] [Marais et al. 2016; Xu et al. 2015] and mobilization of metals such as copper or iron [Meskhidze et al. 2003; Weber et al. 2016]). This does not mean, however, that the pH of rain, or the acidic flux to the ground, is not responding to controls, as those are more responsive to the formation of sulfuric and nitric acid, which are both decreasing.

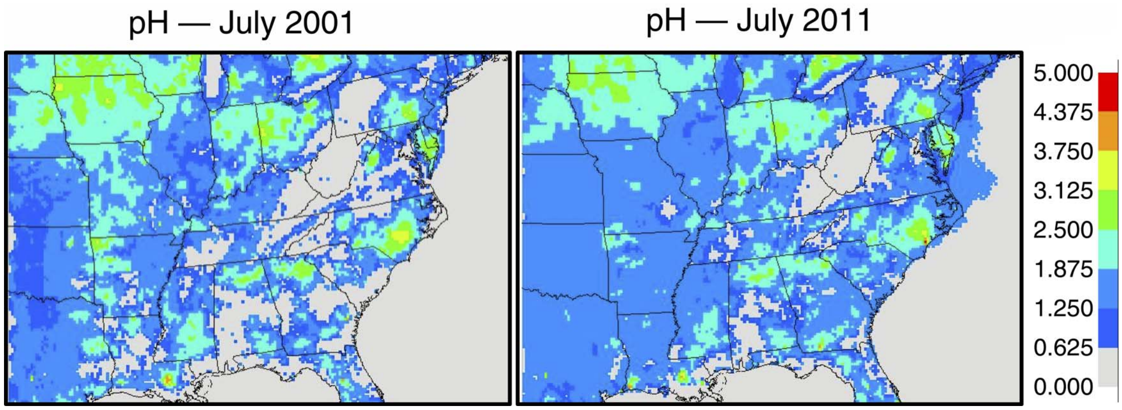


Figure 12. CMAQ-modeled pH concentrations. These were are relatively constant between summers in 2001 and 2011.

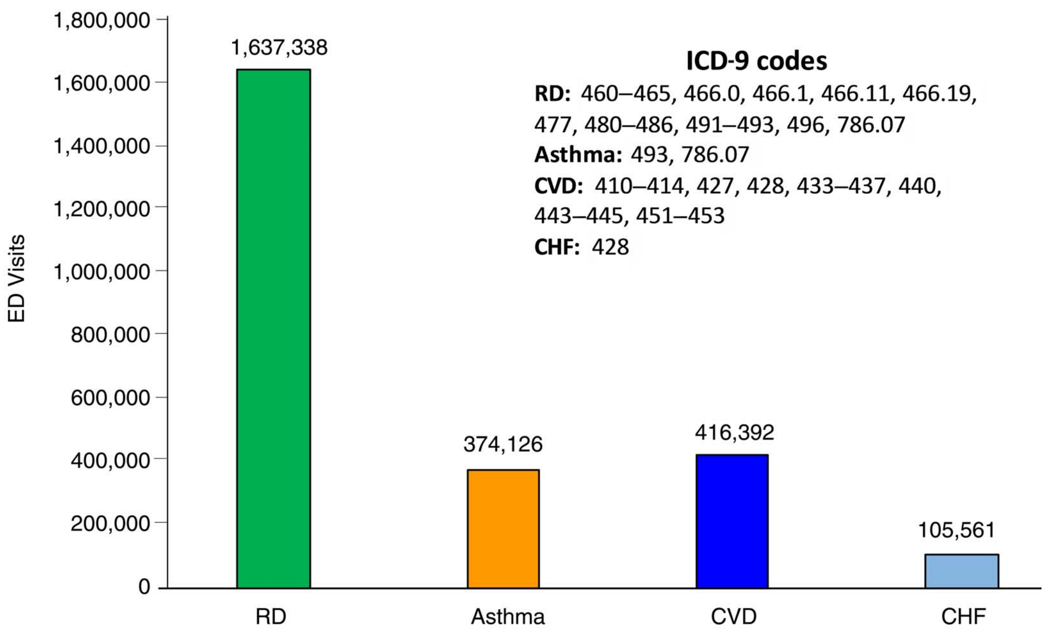


Figure 13. ED visits recorded by outcome in the 5-county Atlanta metropolitan area, 1999–2013. RD = respiratory disease, CVD = cardiovascular disease, CHF = congestive heart failure.

HEALTH IMPACTS

Hospital ED Descriptive Data

There were 16,191,785 total ED visits recorded in the 5-county Atlanta metropolitan area from 1999–2013, an average of roughly 1.08 million ED visits per year. There were 1,637,338 ED visits for RD; 374,126 ED visits for asthma; 416,392 ED visits for CVD; and 105,561 ED visits for CHF (Figure 13). For the control condition, finger wounds, there were 154,177 ED visits for this period; there was a mean daily count of 28 with a standard deviation (SD) of 7.5, and a slight downward trend over the period.

The number of total ED visits increased from 710,414 in 1999 to 1,237,541 in 2013, a 74.2% increase. The population of the 5-county Atlanta metropolitan area increased by 24.3% during the same period, suggesting that part of the increase in ED visits may be due to greater completeness of hospital data. From 1999 to 2013, RD ED visits increased by 59.5%, asthma ED visits increased by 76.0%, CVD ED visits increased by 140.9%, and CHF ED visits increased by 225.0% (Figure 14).

Overall there were 42 hospitals in the database, but the number of hospitals in the database at each point in time changed from 1999–2013 (Table 7). Over the 365 days in 1999, there was an average of 28.1 hospitals in our hospital database; accounting for missing data, an average of

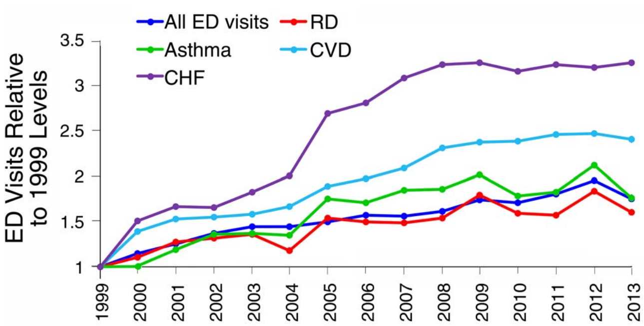


Figure 14. ED visits relative to 1999 levels, by year and category in the 5-county Atlanta metropolitan area, 1999–2013. RD = respiratory disease, CVD = cardiovascular disease, CHF = congestive heart failure.

27.9 hospitals contributed ED data per day. These numbers gradually increased to 37.9 hospitals in the database for 2006, with an average of 37.7 hospitals contributing ED data per day. Hospital numbers largely leveled off after 2006. The percentage of hospitals in the database contributing data ranged from 96.9% in 2005 to 100.0% from 2008–2010. The number of hospitals in the database and the percentage completeness of data from these hospitals did not appreciably change between 2004 (the last year of data being obtained directly from individual hospitals) and 2005 (the first year data were obtained from the Georgia Hospital Association).

ED Visits Prevented over Entire Study Period by Year

From the multipollutant Poisson generalized linear models, we obtained coefficient estimates for the associations between pollutant parameters and ED visits for RD, asthma, CVD, and CHF. We then incorporated the difference between observed and counterfactual pollutant levels to determine daily risk ratios, which were used to estimate the number of ED visits prevented by pollution-control policies. From 1999–2013, using concentration–response estimates derived from the 1999–2005 model, an estimated 52,717 RD ED visits in the 5-county Atlanta metropolitan area were prevented due to all selected pollution-control policies, which was 3.1% of all RD ED visits which would have occurred in the absence of these policies. Likewise, an estimated 38,038 asthma ED visits (9.2%), an estimated 3,057 CVD ED visits (0.7%), and an estimated 2,104 CHF ED visits (2.0%) were prevented due to all selected pollution-control policies.

Health impacts of pollution-control policies increased over the study period as additional policies were enacted and gradually became more fully realized (Figure 15). Using health impact models fit from 1999–2005, estimated RD ED visits prevented increased from 2.5% in 1999 to 6.1% in 2013 (3.1% overall), estimated asthma ED visits prevented increased from 2.0% in 1999 to 17.0% in 2013 (9.2% overall), estimated CVD ED visits prevented increased from 0.2% in 1999 to 2.5% in 2013 (0.7% overall), and estimated

Table 7. Average Daily Numbers of Hospitals in the Database and Contributing Data for Each Year of the Study[a]

Year	Hospitals in Database	Hospitals Contributing Data	Percent Contributing
1999	28.1	27.9	99.6
2000	32.7	31.9	97.7
2001	33.5	32.5	97.0
2002	34.4	33.4	97.1
2003	35.9	34.8	97.0
2004	37.0	36.0	97.3
2005	37.5	36.3	96.9
2006	37.9	37.7	99.3
2007	36.9	35.9	97.3
2008	37.0	37.0	100.0
2009	37.0	37.0	100.0
2010	37.0	37.0	100.0
2011	37.0	36.9	99.7
2012	37.0	36.8	99.6
2013	35.9	35.1	97.8

[a] Numbers reflect average of daily values over the course of each year.

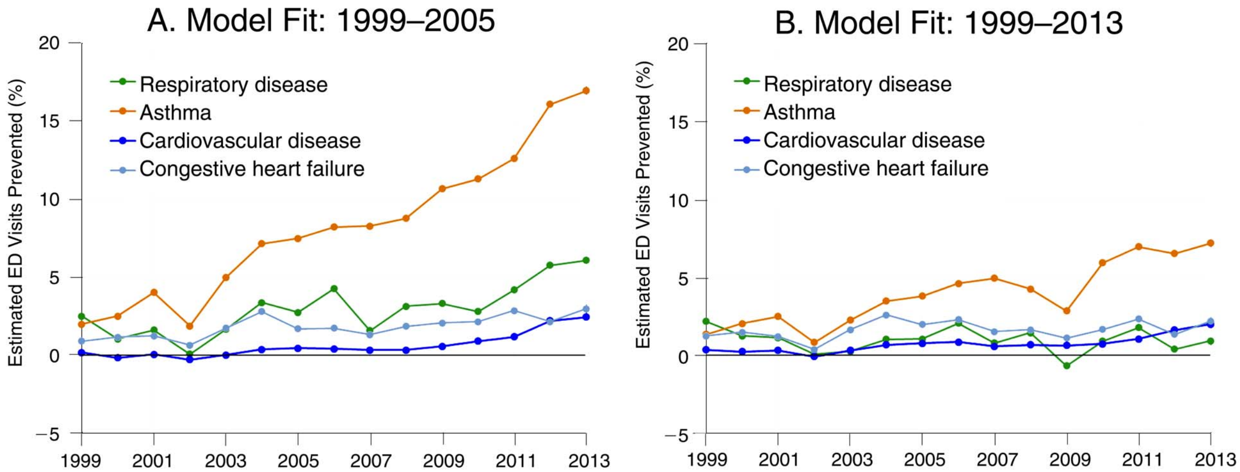


Figure 15. Estimated percentage of ED visits prevented by the set of pollution-control policies considered in this study by outcome and year in the 5-county Atlanta metropolitan area, 1999–2013. A: 1999–2005 model fit, and **B:** 1999–2013 model fit. Model formulation: 7-pollutant, all cubic polynomial and interaction terms included.

CHF ED visits prevented increased from 0.9% in 1999 to 3.0% in 2013 (2.0% overall). Using health impact models fit from 1999–2013, estimated asthma ED visits prevented increased from 1.3% in 1999 to 7.2% in 2013 (4.3% overall), and estimated CVD ED visits prevented increased from 0.3% in 1999 to 2.0% in 2013 (0.8% overall). ED visits for RD (0.9% estimated prevented overall) and CHF (1.7% estimated prevented overall) did not meaningfully increase over this period. To describe the full effect of all selected pollution-control policies, subsequent results are presented below for estimates of ED visits prevented over the last two years of the study (2012–2013), which capture the period of greatest impact of these policies.

ED Visits Prevented During Period of Fullest Implementation of Pollution-Control Policies

From 2012–2013, there were 260,018 recorded RD ED visits in the 5-county Atlanta metropolitan area. All selected pollution-control policies were estimated to prevent 16,315 visits, or 5.9% of all RD ED visits that would have occurred in the absence of these policies (95% interval estimate: −0.4% to 12.3%) using the health impact model fit from 1999–2005, or 0.6% (95% interval estimate: −3.7% to 6.1%) using the health impact model fit from 1999–2013 (Figure 16). There were 60,731 recorded asthma ED visits, and pollution-control policies were estimated to prevent 11,985 visits (16.5%; 95% interval estimate: 7.5% to 25.1%) using the health impact model fit from 1999–2005, or 6.8% (95% interval estimate: −0.3% to 13.9%) using the health impact model fit from 1999–2013. There were 69,910 recorded CVD ED visits, and pollution-control policies were estimated to prevent 1,662 visits (2.3%; 95% interval estimate: −1.8% to 6.2%) using the health impact model fit from 1999–2005, or 1.8% (95% interval estimate: −1.0% to 4.6%) using the health impact model fit from 1999–2013. There were 18,129 recorded CHF ED visits, and pollution-control policies were estimated to prevent 477 visits (2.6%; 95% interval estimate: −6.3% to 10.4%) using the health impact model fit from 1999–2005, or 1.8% (95% interval estimate: −3.5% to 7.7%) using the health impact model fit from 1999–2013.

From 2012–2013, there were 16,939 recorded ED visits for finger wounds in the 5-county Atlanta metropolitan

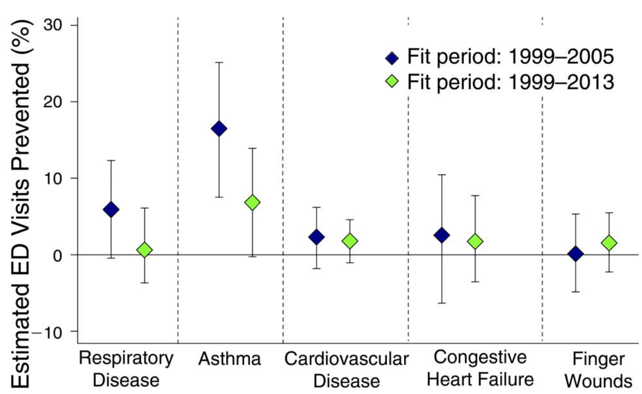


Figure 16. Estimated percentage of ED visits prevented by the set of pollution-control policies considered in this study by outcome and period of model fit, 5-county Atlanta metropolitan area, 2012–2013. Model formulation: 7-pollutant, all cubic polynomial and interaction terms included.

area, and all selected pollution-control policies were estimated to prevent 24 visits, or 0.1% of all finger wound ED visits that would have occurred in the absence of these policies (95% interval estimate: −4.8% to 5.3%) using the health impact model fit from 1999–2005, or 1.6% (95% interval estimate: −2.2% to 5.5%) using the health impact model fit from 1999–2013. Unlike the primary outcomes of interest, estimated effects were not higher using the health impact model fit from 1999–2005. These analyses, which showed no strong connection between pollution-control policies and finger wounds, did not reveal evidence of any uncontrolled confounding in the health impact model.

Exploration of the Sensitivity of Health Impact Results to Model Parameterizations

Estimated relative effects of pollution-control policies were generally slightly greater for the 5-county Atlanta metropolitan area than for the 20-county Atlanta metropolitan area using models fit for 1999–2005 (Figure 17A). Compared with an estimated 5.9% of RD ED visits prevented in the 5-county area, there were an estimated 5.1% of RD ED visits prevented in the 20-county area (95% interval estimate: −0.8% to 11.0%). Compared with an estimated 16.5% of asthma ED visits prevented in the 5-county area, there were an estimated 13.6% of asthma ED visits prevented in the 20-county area (95% interval estimate: 5.0% to 21.5%). Compared with an estimated 2.3% of CVD ED visits prevented in the 5-county area, there were an estimated 1.6% of CVD ED visits prevented in the 20-county area (95% interval estimate: −2.3% to 5.7%). Compared with an estimated 2.6% of CHF ED visits prevented in the 5-county area, there were an estimated 2.8% of CHF ED visits prevented in the 20-county area (95% interval estimate: −4.9% to 9.6%).

The same patterns held for models fit for 1999–2013 (Figure 17B). Compared with an estimated 0.6% of RD ED visits prevented in the 5-county area, there were an estimated 0.1% of RD ED visits prevented in the 20-county area (95% interval estimate: −4.2% to 5.3%). Compared with an estimated 6.8% of asthma ED visits prevented in the 5-county area, there were an estimated 4.7% of asthma ED visits prevented in the 20-county area (95% interval estimate: −1.7% to 11.0%). Compared with an estimated 1.8% of CVD ED visits prevented in the 5-county area, there were an estimated 1.1% of CVD ED visits prevented in the 20-county area (95% interval estimate: −1.4% to 3.5%). Compared with an estimated 1.8% of CHF ED visits prevented in the 5-county area, there were an estimated 1.5% of CHF ED visits prevented in the 20-county area (95% interval estimate: −3.1% to 6.5%).

Results of health impact models were compared based on parameterizations of pollutant variables: use of either a linear term or a cubic polynomial for each pollutant, and the inclusion or exclusion of first-order linear interaction terms for each pair of pollutants. For models fit from 1999–2005, for RD and asthma (Figure 18A), the lowest estimated percentage of ED visits prevented was for the model with only linear terms and no interaction terms, while the models with cubic polynomials and interaction terms had the highest estimates for percentage of ED visits prevented. These patterns were not as apparent for CVD and CHF, though the models with cubic polynomials and interaction terms had the highest estimated percentage of ED visits prevented for CHF and close to the highest estimated percentage of ED visits prevented for CVD. For models fit from 1999–2013 (Figure 18B), patterns were similar: models with only linear terms and no interaction terms had the lowest estimated percentage of ED visits prevented for each outcome, while models with cubic polynomials and interaction terms had either the highest or close to the highest estimates for ED visits prevented.

Results of health impact models were also compared for a 1-pollutant model (which estimated ED visits prevented due to the effect of policies on ambient $PM_{2.5}$), a 5-pollutant model (which included $PM_{2.5}$, CO, O_3, NO_2, and SO_2), a 7-pollutant model (which included $PM_{2.5}$, CO, O_3, NO_2, SO_2, OC, and NO_3^-), and a 9-pollutant model (which included $PM_{2.5}$, CO, O_3, NO_2, SO_2, OC, NO_3^-, EC, and SO_4^{2-}). For models fit from 1999–2005 (Figure 19A), for all outcomes, the highest estimated percentages of ED visits prevented were for the 7-pollutant models. For models fit from 1999–2013 (Figure 19B), the 7-pollutant models produced either the highest or close to the highest estimated percentage of ED visits prevented for each outcome. The fact that the 7-pollutant model produced consistently high estimates of ED visits prevented may suggest that the difference between results by model is not random variation, but rather that the 7-pollutant model formulation more fully captured the impact of pollution-control policies than the other models.

ED Visits Prevented for Specific Policy Scenarios

Estimated percentages of ED visits prevented, using models fit from 1999–2005, are shown in Figure 20, A–D for RD, asthma, CVD, and CHF. EGU policies were generally estimated to have a greater health impact than mobile policies for preventing RD ED visits, especially the ARP and the NBP. Those patterns were also similar for asthma ED visits. For CVD ED visits, the CAIR multipollutant program was estimated to have the largest impact of any set of policies. However, these patterns were not similar using models fit for 1999–2013 (Figure 20 E–H). For all outcomes, there was considerable overlap in the interval estimates for the majority of pollution-control scenarios.

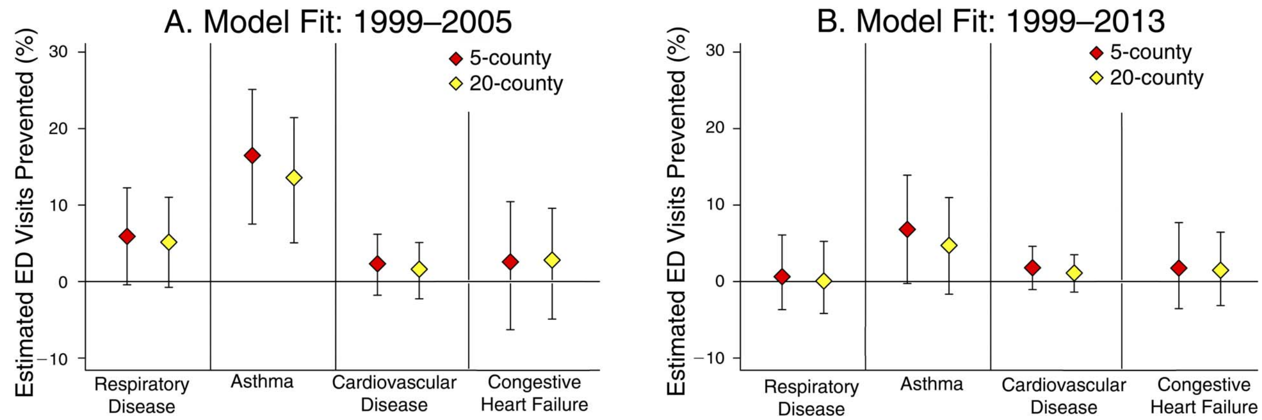


Figure 17. Estimated percentage of ED visits prevented by the set of pollution-control policies considered in this study by outcome, 2012–2013, comparing results for the 5-county Atlanta metropolitan area (red) with results for the 20-county Atlanta metropolitan area (yellow). A: 1999–2005 model fit, and **B:** 1999–2013 model fit. Model formulation: 7-pollutant, all cubic polynomial and interaction terms included.

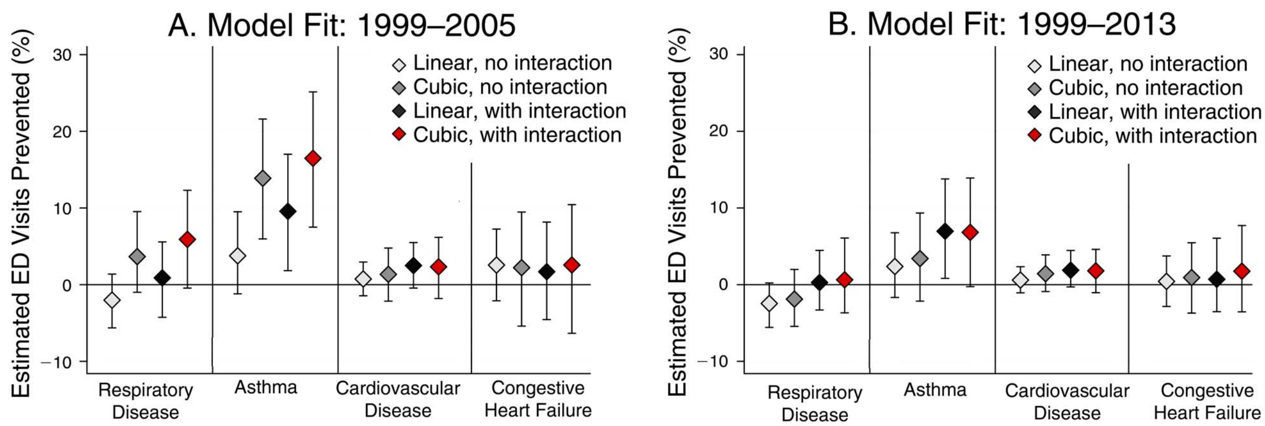


Figure 18. Estimated percentage of ED visits prevented by the set of pollution-control policies considered in this study by outcome, 2012–2013, for the 5-county Atlanta metropolitan area, comparing models with linear pollutant terms, models including cubic polynomial pollutant terms, and models including first-order interactions between linear pollutant terms. A: 1999–2005 model fit, and **B:** 1999–2013 model fit. Model formulation: 7-pollutant model.

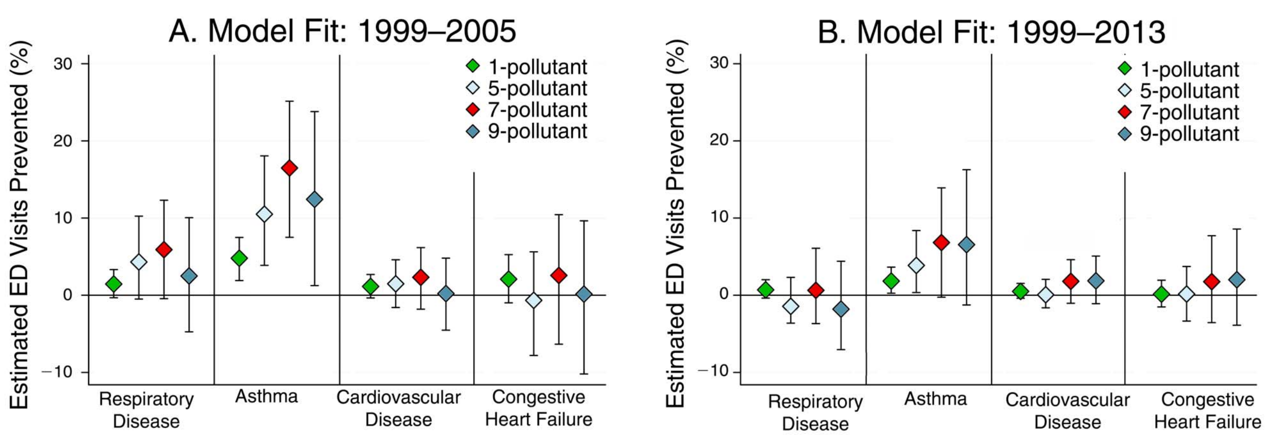


Figure 19. Estimated percentage of ED visits prevented by the set of pollution-control policies considered in this study by outcome, 2012–2013, for the 5-county Atlanta metropolitan area, comparing models with different multipollutant formulations. A: 1999–2005 model fit, and **B:** 1999–2013 model fit. Model formulation: all cubic polynomial and interaction terms included. The 1-pollutant model included $PM_{2.5}$; the 5-pollutant model included $PM_{2.5}$, CO, O_3, NO_2, and SO_2; the 7-pollutant model included $PM_{2.5}$, CO, O_3, NO_2, SO_2, OC, and NO_3^-; the 9-pollutant model included $PM_{2.5}$, CO, O_3, NO_2, SO_2, OC, NO_3^-, EC, and SO_4^{2-}.

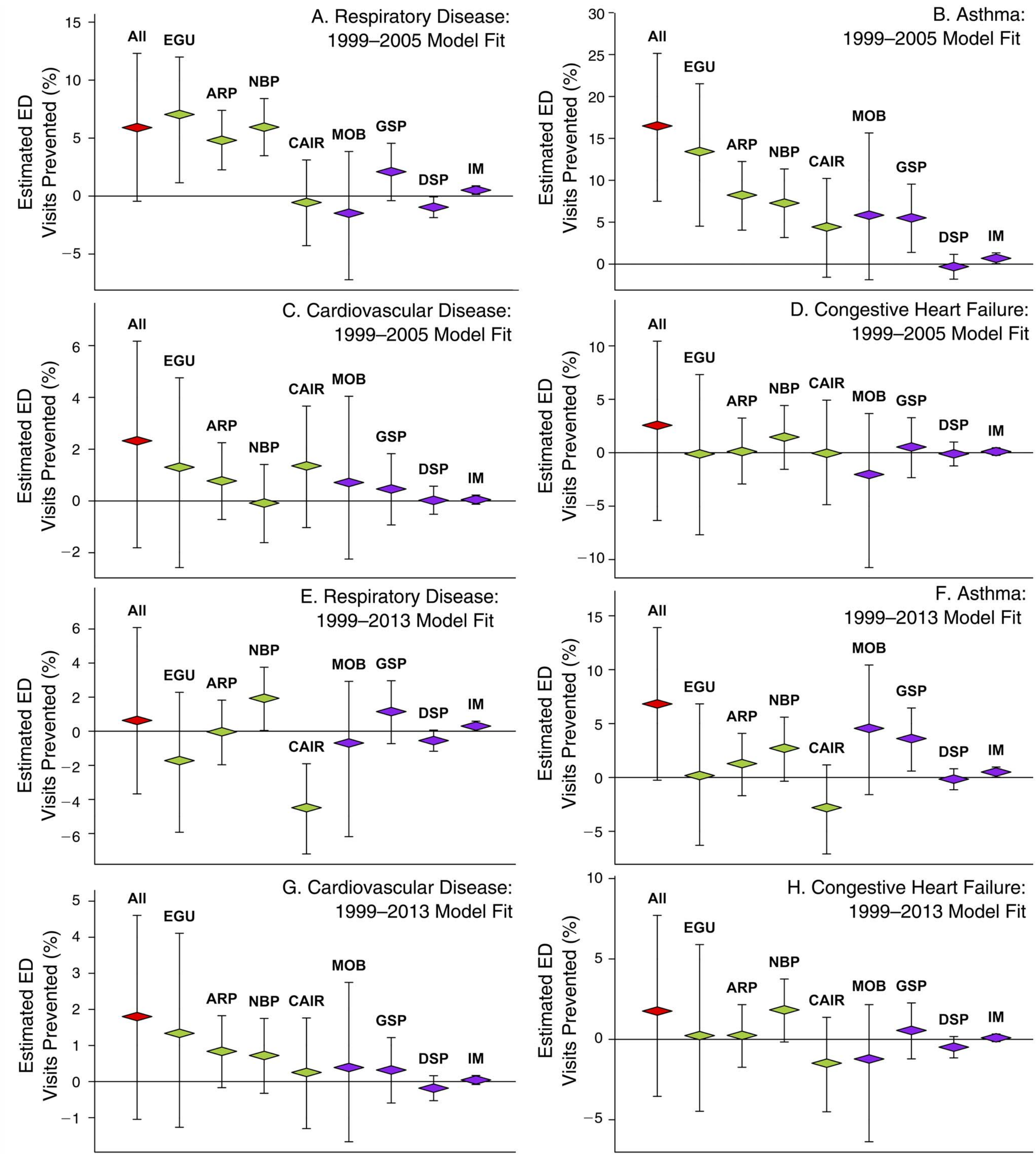


Figure 20. Estimated percentage of ED visits prevented by individual pollution-control actions, 2012–2013, for the 5-county Atlanta metropolitan area. A–D: 1999–2005 model fit, and **E–H:** 1999–2013 model fit, by cardiorespiratory outcome. Policies regulating EGU emissions are shown in green; policies regulating mobile emissions are shown in purple. The 95% confidence intervals with all uncertainty included were calculated for all pollution-control policies, all EGU policies, and all mobile policies; confidence intervals for the individual sets of policies reflect uncertainty in the health impact model. All = all pollution-control policies; EGU = all electricity generating unit policies; ARP = Acid Rain Program; NBP = NO_x Budget Trading program; CAIR = Clean Air Interstate Rule/multipollutant rule; MOB = all mobile policies; GSP = gasoline programs; DSP = diesel programs; IM = inspection and maintenance programs.

DISCUSSION AND CONCLUSIONS

Previous discussions of accountability methods and frameworks have concluded that there is no single approach that can attribute causality in changes along the chain of accountability to regulatory actions; each approach should be judged on its merits — in other words, ability to answer the question at hand while accounting for confounding factors and estimating uncertainties — and limitations (HEI 2010; Pope et al. 2012). To this end, the present study utilizes multiple methods to investigate relationships along the chain and compares the results in light of each method's strengths and limitations.

EMISSIONS

Uncertainties in emissions estimates are well recognized. An added complexity in conducting an accountability assessment is linking regulations to rules and the implementation of specific controls. Meetings with stakeholders proved to be an important aspect of this study in assessing this relationship. Stakeholders in the field emphasized that decisions made at the utility or facility level reflect multiple inputs, including regulations (one or more), generating capacity, fuel cost, and future projections of each of these. Definitively attributing emissions reductions to specific regulatory programs at the federal level via complex control programs, such as those investigated in this report, proved impossible using the methods involved here. One would have to both correctly link specific control decisions to a program or programs and account for how the ensuing external factors (e.g., fuel costs, economic growth, and other regulations) further affected operating decisions.

Regionally, the most effective regulation at reducing year-round NO_x emissions was the ARP (and related state programs). During summers beginning in 1999 (and particularly in 2003 and onward), our analysis attributes large reductions in NO_x to the NBP. CAIR is associated with smaller reductions, largely because ARP and NBP already required major reductions prior to CAIR's promulgation. The region-wide adoption of natural gas toward the end of the period of interest impacted NO_x emissions, but we could not attribute the degree of fuel switching exclusively to regulations or changes in fuel cost. The ARP is associated with moderate reductions in SO_2 emissions regionally, though CAIR is associated with the largest reductions in both the ANAA and regionally, occurring year 2009 and later.

Mobile emissions programs are associated with major reductions in emissions of all the modeled pollutants. GSP programs have reduced NO_x and VOC emissions the most over the period, and DSP programs have reduced $PM_{2.5}$ emissions. IM programs, as modeled by MOVES, have not had a substantial impact on emissions.

There is evidence from other areas of the country that current mobile-source NO_x emissions estimates may be biased high by as much as a factor of two. Frost (2016) suggested that CO emissions estimates may be biased as well. If this is true in the Southeast, or elsewhere, it has major implications on this work and air quality management generally. First, it would indicate that regulations targeting mobile-source emissions have been more effective than is evidenced by modeled emissions. Second, the amount of NO_x emissions to be further reduced would be lower, which has implications on how to assess potential future interventions and regulatory effectiveness. There are other ramifications (e.g., air quality model evaluation and the development of air quality sensitivities). We examined this potential bias in multiple ways, including through our air quality modeling evaluation, empirical trend analysis of NO_x trends using both ground and satellite-based observations, and a ratio-of-ratios method. Our results are suggestive of a bias, but the evidence here is not strong and can be obscured by the more complex chemistry of the large biogenic fluxes in the Southeast that, for example, lead to formation of organonitrates, which are not typically measured and are not observed from space. The ratio-of-ratios method is suggestive of a potential bias, as the ambient concentrations of other directly emitted species (CO and EC) have been reduced proportionally less in relationship to their emissions. We do find a high bias in simulated NO_x and CO concentrations in our air quality modeling. On the other hand, the satellite data and ground level observations follow similar annual trends as the NO_x emissions, which are also similar to those in previous studies (Blanchard et al. 2016; Pachon et al. 2012; Vijayaraghavan et al. 2014).

It is hard to overstate the importance of better understanding and quantifying mobile-source NO_x emissions, particularly in the light of the major reductions in utility NO_x emissions and the tightening ozone standard. Understanding how NO_x-limited the region is will be critical to identifying the most effective strategies to further reduce O_3 (and PM) and in quantifying the degree of further controls on various sources that is required. The empirical method used here to construct counterfactuals is less sensitive to systematic biases in emissions; the trend is most important.

AIR QUALITY

O_3 results using both empirical and chemical transport modeling methods suggest that mean annual O_3 is not impacted to a large degree by meteorology. Detrending adjustments for O_3 were larger for the summertime means, with seven of the 13 years being attributed adjustments of 3 ppb or more, with a maximum adjustment of 6.9 ppb in 2002 (Appendix Figure A.9; available on the HEI website). Mean annual counterfactual O_3 is consistently higher than observed, but still follows a cyclical multiyear pattern over the course of the period. This suggests that, while meteorological contributions are important on a day-to-day basis — particularly leading to high and low O_3 days — emissions drive peak and mean levels of O_3, as well as variability. The dip in concentrations in the years surrounding 2009 is of interest because of the recession, which began in 2008. Dynamic analysis using CMAQ and empirical modeling reaffirms that improvements in air quality (specifically, reductions in the highest O_3 levels) between 1999 and 2013 are attributable more to emissions reductions than to meteorological variability.

Blanchard and colleagues (2014) found, using a different detrending method, that detrending changed the long-term trend in MDA8h O_3 from insignificant to significant at JST and other SEARCH sites. In general, the annual O_3 adjustments from the detrending are smaller in this work than in previous detrending studies (Blanchard et al. 2014; Camalier et al. 2007). The present study looks at seasonal averages, not annual peaks, and finds that detrending does not impact the assessment of long-term trends in summer (decrease) and winter (increase). The current work is able to match statistically significant changes in air pollution concentrations with periods of emissions changes in utilities (Henneman et al. 2015).

Direct comparisons of concentrations and sensitivities between the empirical and CTM methods show that the two approaches estimate similar emissions–air quality relationships, with exceptions (in particular, NO_2 and NO_3^-). Both models demonstrate that NO_x controls have reduced summertime O_3 and increased wintertime O_3. SO_4^{2-} is closely linked to EGU SO_2 emissions, and has been reduced, particularly in response to ARP and CAIR-related controls. The two models agree best on O_3 and $PM_{2.5}$ concentrations and sensitivities, providing confidence in model results for two species that have proven difficult to reduce. Other gaseous and particulate species show more variability between the models, which asserts evidence that model users should place less credence in these results.

Operational and dynamic evaluations of CMAQ show that the model captures spatial and temporal variability in multiple air pollutants. One concern raised by the results is the persistent low bias of OC in the summer and high bias in the winter, which contributes a large fraction of total $PM_{2.5}$ bias. This fraction has increased as other species of PM (e.g., SO_4^{2-}) have been reduced. High bias of NO_2 and CO in Atlanta may be indicative of a bias in emissions or due to the grid size. O_3 modeling has a relatively low bias, but tends to be biased low on high O_3 days. Recent upgrades to CMAQ (versions 5.1 and 5.2) have been shown to reduce this bias (Appel 2018; Pleim 2016).

Aerosol acidity, and its change across the period of emissions reductions, evolved as a potentially important issue during this research project. Aerosol pH has previously been linked to ambient aerosol concentrations, composition, and toxicity (Kleinman et al. 1989; Weber et al. 2016). The notable finding is that aerosol pH remains low, despite major reductions in SO_2 and NO_x emissions, as studied here. Here, we specifically focused on the trends in aerosol pH in the Southeast using the data, model applications, and approaches developed for this study. We find that, indeed, aerosol pH in the Southeast remains low (pH's around 1–3, depending on methods used), and the aerosol acidity today is similar to what it was a decade ago, despite the major reductions in the two main acidifying species — sulfuric acid and nitric acid produced from gas phase reactions of SO_2 and NO_x. Further, aerosol pH is not expected to increase significantly in the near future, in spite of further NO_x and SO_2 controls. The result is due both to thermodynamic considerations and the availability of ammonia. Aerosol pH has health implications as acidic aerosols have been linked to health impacts (Kleinman et al. 1989; Utell 1985) and has been linked to the formation of SOA. On the other hand, the continued low pH is not an indication of a failure of the controls. There is a significant reduction in aerosol levels, and the deposition of acidifying and eutrophying species has been reduced. This improved understanding of aerosol pH, as well as the potential impact on SOA formation, is critical to assessing the impact of future emissions controls in biogenic-rich areas like the Southeast. A further issue is to examine the potential benefits of how reduced SO_4^{2-} and NO_3^- formation may impact the transport of ammonium, a reactive nitrogen species, which can lead to harmful ecological impacts downwind.

As mentioned above, it is impossible to validate the counterfactual concentrations time series directly. Instead, we have applied multiple tools to evaluate various aspects of the modeling process. These include comparisons of sensitivities across modeling platforms (empirical and CTM), assessments of cross-correlations between pollutant species, and quantitative measures of uncertainty that account for potential biases in counterfactual emissions and emissions–concentration sensitivities. The first two tests provide confidence that the sensitivities to emissions capture suitable relationships. The third test confirms that the differences between observed ambient concentrations and counterfactuals are statistically significant for the majority of pollutants in the EGUMOB, EGU, and MOB cases.

HEALTH IMPACT ANALYSIS

Strengths of Analysis

One major strength of this study is the use of a large hospital database consisting of 42 Atlanta area hospitals. We had access to patient-level ED data, which were converted into daily counts for a 15-year period for several cardiorespiratory outcomes. There were over 16 million ED visits recorded in this database from 1999–2013, and the considerable study size (both in ED visit counts and in length of the study) allowed for assessment of a variety of outcomes, model parameterizations, and pollution-control scenarios. In addition, the large suite of air quality variables that were continuously measured on a daily basis over the 15-year period allowed for assessment of different multipollutant model formulations. The Atlanta hospital and pollutant databases had been used in a number of previous studies that had established acute relationships between pollutants and ED visits, assessed potential biases, compared different analytical methods, and determined optimal control for time-varying confounders; lessons learned from these earlier studies contributed substantially to the improvement of the current study.

Decisions in Modeling Associations Between Pollutants and ED Visits

Numerous factors needed to be considered when deciding on the methodological approaches for the health impact models. In order to limit concerns over multiple comparisons or data fishing, we used a priori decisions for modeling decisions whenever it was suitable. We used set covariates, which had been determined through previous studies, including the particular formulation of meteorological terms. We used a priori lag structures, with lag 0–2 used for RD and asthma and lag 0 used for CVD and CHF, decisions also based on previous research. The 1-, 5-, and 9-pollutant models were chosen a priori, and the 7-pollutant model reduced concerns about collinearity. We did try models with or without cubic polynomials and interaction terms since we were not sure if those factors would affect model results. However, to avoid picking and choosing convenient results, we decided a priori to either use cubic polynomials for all pollutants or none, and to either include interactions between all pollutants or none.

There were other modeling decisions in which we were guided at least partially by results, but this was only done when these choices were appropriate and consistently applied. For example, pollutants were more predictive of ED visits in the first half of the study period. Factors, such as the greater variability in ambient pollutant levels — emission reductions have reduced pollutant concentration averages, peaks, and variability — potentially leading to more accurately measured associations with ED visits, may have contributed to this result. Similarly, after early testing showed that 7-pollutant models, models with cubic polynomials, and models with interaction terms consistently captured more of the health impact of pollution-control policies, we used those model parameterizations for the primary model results. These modeling choices resulted in consistently stronger results for all outcomes, suggesting that the difference may be due to a decrease in model misspecification as opposed to simply random noise. The 7-pollutant model incorporated the most information on overall air quality without the larger collinearity issues of the 9-pollutant model, and the cubic polynomials and interaction terms additionally added to a more refined model of the associations between pollutants and ED visits.

The 7-pollutant model with cubic polynomials had a total of 42 pollutant terms (3 for the cubic polynomial terms for each of the 7 pollutants, and 21 two-way interaction terms). There were downsides to using this complex model. For one, within this multipollutant framework it was difficult to determine the associations between individual pollutants and ED visits, or to assess the effects through reductions of a single pollutant on ED visits. However, these were never the goals of this particular study; there is already ample evidence connecting each of these pollutants to cardiorespiratory outcomes. Furthermore, trying to assess the effects on health outcomes through reductions of individual pollutants would invariably not capture the full impacts of pollution-control policies: even policies aimed at reducing emissions of individual pollutants would, by affecting atmospheric chemistry, result in changes to the overall air quality profile.

A general challenge when allowing for multipollutant interactions and nonlinearity in a model is in its interpretation. However, interpreting model parameters is less of an issue in this analysis compared with more conventional analyses. Our goal here is to simply estimate daily risk ratios for one specified contrast (counterfactual vs. actual levels) each day by estimating a joint effect. A simpler model would not meaningfully simplify our interpretation of the desired joint effect.

Counterfactual Study Design

Effects of interventions are often estimated using study designs that are framed as natural experiments. The term *natural experiment* is somewhat of a misnomer as, by definition, in an experiment the conditions under study are manipulated by the investigator rather than by natural forces. The implication of a natural experiment is that the results obtained are consistent with what would have been obtained in an actual experiment if one had been undertaken. For our study question, we could conceptualize a natural experiment utilizing a spatial and/or a temporal contrast. To utilize a spatial contrast we would need to find a city similar to Atlanta (e.g., similar in population size, population makeup, ED usage, meteorology, and traffic), except that this control city would not be subject to the air pollution-control policies that Atlanta had experienced during the study period (1999–2013). Thus, the pollution levels in the control city should approximate the counterfactual pollution levels during the intervention period; that is the pollution levels that would have occurred in Atlanta in the hypothetical absence of the air pollution-control policies.

Perhaps more plausible than finding such a control city would be to use a control period in Atlanta and perform a pre–post analysis. However, this approach to approximate the counterfactual is also problematic because of the length of the intervention period, Atlanta's changing population over time, the occurrence of seminal events (e.g., the Great Recession), and the gradual implementation of air pollution policies at different time points (e.g., engine emissions standards for new cars that were phased in over several years).

Rather than use a proxy city or proxy period to represent the counterfactual pollution levels in Atlanta, our approach was to incorporate changes in emissions inputs with meteorology and atmospheric chemistry to directly estimate daily counterfactual levels for selected pollutants in Atlanta during the study period accounting for model uncertainty. The commonality of these approaches is that valid causal inference depends on the accuracy of representing the counterfactual experience — either by design or by analysis.

In our initial analyses we observed stronger health associations in the first half of the study period (roughly 1999–2005) compared with the latter half. This temporal contrast of the concentration–response parameters led us to a broader consideration of the counterfactual experience. That is, we needed to consider not only how air pollutant levels were affected by specified pollution-control policies but also whether these policies affected the health concentration–response functions. If so, it would be a source of error to estimate the counterfactual number of ED visits using concentration–response parameters that were affected by air pollution-control policies. On the other hand, the change in the estimated health associations over time could be due to extraneous factors unrelated to pollution-control policies such as changing population susceptibility, model misspecification, or chance. This was a challenging issue and led us to present two sets of results; one using concentration–response parameters estimated from the 1999–2005 period and the other using concentration–response parameters estimated from the entire 1999–2013 study period.

Use of Single Central Monitor Data to Predict ED Visits Across the 5-County Area

We used results from a single central monitor to predict ED visits across the 5-county Atlanta metropolitan area. While pollutant data were available from other Atlanta-area monitors, the Jefferson Street location was the only site where measurements of all considered pollutants were colocated for the entire study period. Using data from other monitors would have involved additional layers of modeling for actual and counterfactual estimates, which would be an additional source of error.

The use of a single central monitor could be a potential study limitation, as pollutant levels measured at the monitor may differ considerably from pollutant levels experienced by the study population. Exposure measurement error for time-series analyses was assessed in a previous study in the Atlanta metropolitan area; this study found that the use of measurements from urban monitors (within 20 miles of the city center) located different distances from geographical subpopulations produced similar associations between pollutants and health outcomes (Sarnat et al. 2010). This suggested that even if measured pollutant levels differed from ambient pollution levels where individuals are located, daily trends in these measures were correlated enough so that measurements from a single central monitor could reproduce valid health associations. Separately, another analysis using the same Atlanta hospital data assessed the use of either unweighted averages of pollutant concentrations over several area monitors or population-weighted averages of these pollutant concentrations, compared with using data from a single central monitor; all air quality metrics resulted in similar associations between pollutants and pediatric asthma ED visits (Strickland et al. 2011). In a follow-up analysis that incorporated simulated measurement error, observed associations between pollutants and ED visits were generally biased toward the null, and this bias was greater when using single central monitor measurements compared with population-weighted average concentrations (Strickland et al. 2013).

A study using simulated time-series pollutant data and Poisson generalized linear models similar to those used in this study showed that associations between pollutants and health outcomes were all biased toward the null, though less for Berkson-type errors, which would result when daily pollutant measurements were close to the population-average exposure (Goldman et al. 2011). If the measurements differ meaningfully from population average exposures, this can create biased associations, with the direction most likely toward null effects (Zeger et al. 2000). The comparison of results from the 5-county analyses with results from the 20-county analyses supports this hypothesis. The population of the 20-county area includes people even further from the central monitor and whose individual exposure to ambient pollutant levels is likely quite different; these results show effects that trend toward the null. This suggests that exposure measurement error could be resulting in a bias toward the null for this study. If so, the true impacts of pollution-control policies may be greater than those estimated for the 5-county analysis.

ED Visits as the Outcome of Interest

Hospital ED visits represent serious adverse health outcomes: patients are suffering distress that is drastic and severe enough to seek immediate, potentially life-saving medical care. Such outcomes would be relatively uncommon compared with more moderate health effects of ambient air pollution, such as mild respiratory distress or minor irritation of the eyes and throat. Figure 21 depicts many potential effects of air pollution; an analysis focusing on ED visits would not capture less severe effects that occur in a larger number of people or less common effects like mortality attributed to ambient pollutants. The impact of pollution-control policies on ED visits cannot be generalized to these other outcomes; however, if an intervention prevents ED visits, it would also likely have additional health benefits.

Furthermore, this study assessed only impacts of pollution-control policies on cardiorespiratory outcomes; ambient air pollution has also been linked to other health problems such as urinary dysfunction, nervous system damage, digestive issues, and developmental disorders (Kampa and Castanas 2008). Finally, this study captured only acute effects of daily increases in pollutants. Long-term exposure to ambient air pollution can lead to cumulative harm and ultimately increased rates of mortality, especially from CVD, stroke, or lung cancer (Götschi et al. 2008; Laden et al. 2006; Pope et al. 2004; Raaschou-Nielsen et al. 2013; Stafoggia et al. 2014). While this study estimated that tens of thousands of ED visits in the Atlanta metropolitan area had been prevented by pollution-control policies, this result is only a fraction of the overall health impact of these policies.

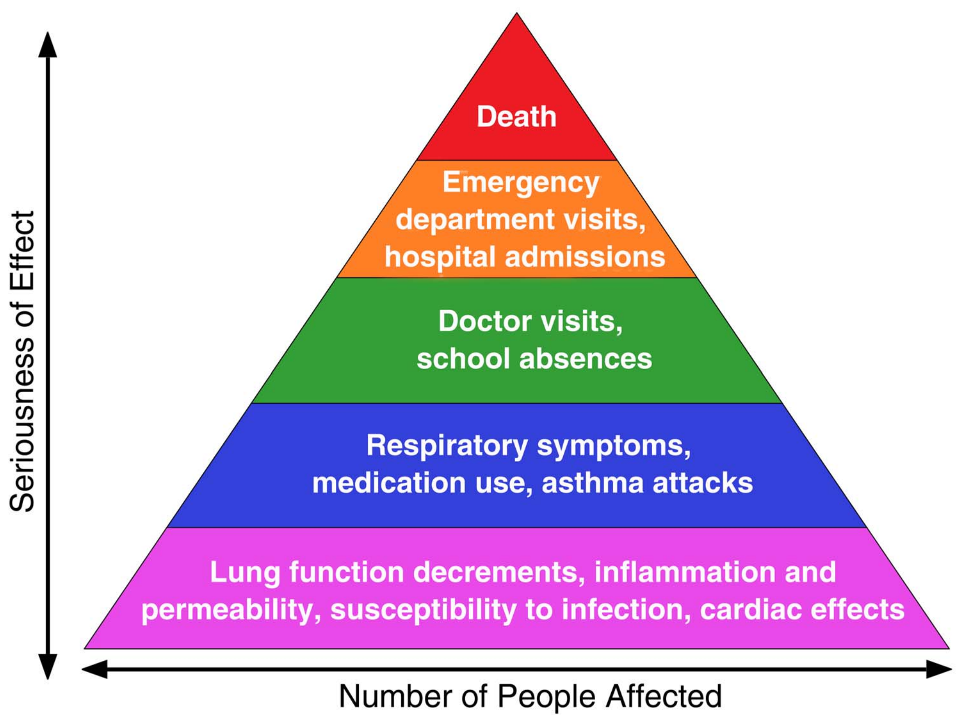


Figure 21. Pyramid of effects of air pollution. (Courtesy of the U.S. EPA 2016b.)

IMPLICATIONS OF FINDINGS

This work employs detailed analyses and multiple models applied along the chain of accountability to quantify effects of regulatory actions on emissions, air quality, and public health. At the first link (connecting regulations and emissions), stakeholder engagement played a key role in assessing how regulatory actions and programs were interpreted and implemented by those in charge of complying with and enforcing them. The discussions and modeling highlighted a difficulty inherent in all accountability studies in attributing reductions to specific regulations, that is, that emitters make decisions regarding controls, fuel, and other factors on a continuous time scale as regulations are put into place and that they make decisions based on a number of inputs and objectives. This confounds attributing changes in emissions exclusively to control programs. While acknowledging the difficulties in attributing emissions reductions to specific control programs, matching emissions reductions to known implementation dates allowed for conclusions to be made regarding the impacts of specific controls. We strongly recommend that accountability studies actively engage stakeholders with intimate knowledge of the specific region and controls examined.

In connecting emissions changes with air quality changes, the work shows how results from multiple methods of different types (i.e., empirically based statistical vs. deterministic) can build evidence for causal linkages. Empirical models that employ a nonlinear measure of atmospheric photochemical state (PS^*) to associate changes in ambient concentration with emissions changes allow for the estimation of nonlinear effects of emissions controls on a daily basis and are a novel aspect of this work. Both statistical and air quality model methods highlight the importance of high-quality long-term data records for accountability studies.

The project highlighted the importance of deterministic air quality models to the accountability field. Long-term emissions inventories and meteorological inputs allow for the comparison of multiple air pollution metrics over wide spatial areas across long time spans. Further, the use of CMAQ for a dynamic analysis and comparison with empirical methods shows the adaptability of CTMs to answer a number of different accountability-related questions.

Policy analysis and emissions modeling showed that the combined effect of all regulations over the study period resulted in reduced emissions of multiple pollutants from both EGU and mobile sources. The ARP, NBP, and related state programs had the largest effects on EGU NO_x emissions, and the ARP and CAIR had the largest effects on SO_2 emissions. For mobile sources, gasoline programs beginning in 2000 had the largest impacts on modeled mobile NO_x, and diesel programs had the largest impact on $PM_{2.5}$. Modeled mobile emissions, however, continue to have high associated uncertainties, which carry on to modeled emissions changes attributable to controls.

Various approaches used to evaluate potential bias in mobile emissions, including ambient observations, satellite monitoring, air quality modeling, and empirical methods, yielded mixed results on the question of whether modeled mobile NO_x emissions are biased, though there is enough information to suggest there is likely a positive bias to support further study of this issue. This may explain, in part, the lack of response sometimes seen in air quality modeling of long-term trends. Further, this would change the modeled response to emissions and modeling of future air quality in response to additional controls for air quality planning and other purposes.

Multiple methods relating air quality to emissions show that emissions reductions have led to reduced ambient concentrations of multiple pollutants that have previously been linked to negative health outcomes. Two statistical methods — meteorological detrending and empirical ambient concentration–emissions models — provide evidence that, while meteorological variability is important on a daily time scale, multiyear trends in ambient air pollution concentrations are driven by anthropogenic emissions (EGU and mobile, in particular) in the southeastern United States.

Accounting for uncertainty in mobile and EGU emissions changes and model parameters, emissions programs under the 1990 C led to statistically significant changes in all of the air pollutants assessed in this study. Response in O_3 was season dependent; decreases compared with the counterfactual were observed in the summertime, and increases were observed in the wintertime. Overall, emissions reductions resulted in a 3% decrease in mean O_3.

The lack of response of aerosol pH to SO_2 and NO_x controls is important in terms of understanding potential health impacts and the response of future air quality (and climate drivers) to emissions controls. Acid-catalyzed reactions that lead to SOA formation will continue; however, the amount of SOA formed should decrease (as has been observed here and by Blanchard et al. [2016] and Marais et al. [2016]) as the total aerosol is reduced, reducing the volume of aqueous droplets that act as reactors. The continued low pH is not indicative of a failure to reduce acid deposition. The total amount of acidifying components in the atmosphere is reduced, as is the flux of acidifying components to the soil and water. The impact of the reduced ammonium formation on downwind deposition of reactive nitrogen should be further studied.

Results of the CMAQ operational and dynamic evaluation highlight important areas in air quality models that need improvement. For instance, results imply that efforts are needed for CMAQ to better capture PM concentration changes in the summer. Continued negative bias in CMAQ SOA formation leads to increased uncertainty in our ability to estimate the impact of controls using air quality models. Further, comparisons with empirical methods show generally lower sensitivities in CMAQ and should be a continued focus of research.

Pollution-control policies, by lowering pollutant emissions and ambient pollution levels, were estimated to substantially reduce cardiorespiratory ED visits in the 5-county Atlanta metropolitan area. From 1999–2013, all selected pollution-control policies were estimated to prevent 3.1% of RD ED visits, 9.2% of asthma ED visits, 0.7% of CVD ED visits, and 2.0% of CHF ED visits using models fit from 1999–2005. These results do not reflect the full impact of pollution-control policies, but rather the impact over the period when these policies were gradually implemented. The period of the study finding the greatest regulatory impact on reducing air pollution levels occurred in the final years of the study. During 2012–2013, all selected pollution-control policies were estimated to reduce RD ED visits by 5.9%, asthma ED visits by 16.5%, CVD ED visits by 2.3%, and CHF ED visits by 2.6% using models fit for 1999–2005. For the 5-county Atlanta population of 3.5 million people, this constituted an estimated 8,157 RD ED visits prevented each year, 5,992 asthma ED visits prevented each year, 831 CVD ED visits prevented each year, and 239 CHF ED visits prevented each year. The results for the combined suite of pollution-control policies were more robust than those for individual policies, so the relative effectiveness of these policies should be compared with caution. The policies described in this study continue to be implemented, so similar quantities of dramatic health impacts should still be occurring every year. These reductions of cardiorespiratory ED visits, while considerable, likely capture only a small portion of the overall impact of pollution-control policies: prevention of moderate cardiorespiratory outcomes, effects on other organ systems, and reductions on chronic conditions stemming from cumulative air pollution exposure are not within the scope of this study.

The use of the chain of accountability to determine the effect of pollution-control policies on health effects depends on the argument that the links within the chain are truly causal. If any of the links in the chain of accountability were not causal, then the effect of policies on health outcomes would be zero. Although we have done our best to control for extraneous factors, this study is not intended to be the sole provider of evidence of the causality for these links. There is an extensive literature — both observational and experimental — which serves as the foundation for the plausibility of the causal relationships between pollution-control policies and pollutant emissions, between emissions and air quality, and between air quality and health outcomes. This study should be viewed in this context, with the existing literature serving as support for our interpretation of results.

Our study represents a vast undertaking that constitutes a significant step forward in terms of air pollution accountability studies. We combined proven and novel methodologies to link pollution-control policies to emissions levels, ambient pollutant levels, and health outcomes. The use of the counterfactual approach allowed for the investigation of different sets of overlapping policies that were implemented gradually over long periods of time. Yet, this study would not have been possible without the extensive pollutant and hospital data sets that provided daily data over a long-term, 15-year study period. The health data were aggregated from 42 different hospitals capturing daily counts of ED visits over different ranges of a large metropolitan area, and this substantial data set allowed for the partitioning of ED data to assess daily counts of several different health outcomes.

All these factors not only contributed to a thorough evaluation of the impacts of pollution-control policies on health outcomes, but also allowed us to conduct numerous sensitivity analyses in order to evaluate the effects of various methodological choices. These included (1) health impacts over different years; (2) health impacts over different geographical scales; (3) pollutant-health associations over different periods; (4) different numbers of pollutants included in health impact models; and (5) different formulations of the health-impact models. Investigation into the details of key regulatory actions allowed for the attribution of changes in emissions, ambient pollutant levels, and ultimately health outcomes to specific pollution-control policies. Accounting for uncertainty at each step in the study allowed for a greater ability to address potential concerns regarding the modeling of counterfactual data.

Testing different modeling choices also resulted in the confirmation or discovery of potentially important knowledge relevant to air pollution epidemiology. Multipollutant models with nonlinear and interaction terms may be more effective at capturing the full extent of the relationship between pollutant levels and cardiorespiratory outcomes. These observed relationships may change over time, though it can be difficult to determine if the changes are real or possibly due to factors such as data irregularity or a change in variance of pollution levels. When exposure is measured at a central monitor location, estimated health impacts may be stronger closer to this location due to reduced measurement error. Finally, uncertainty in associations between ambient pollutant levels and health outcomes is much greater than uncertainty in the associations between emissions and ambient pollutant levels.

This project has produced a wealth of information that is likely to be tremendously useful to further assessments of the impacts of pollution-control policies. We look forward to the results and methodological advances described in this study being utilized to inform and guide future air pollution regulatory actions.

ACKNOWLEDGMENTS

This material is based on work supported by the Health Effects Institute, the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903, the National Institute for Occupational Safety and Health Training Grant T03OH008609, and an endowment to Georgia Institute of Technology by Howard T. Tellepsen. The project leverages resources from the Southeastern Center for Air Pollution & Epidemiology (SCAPE) under U.S. EPA STAR Grant #R834799. The contents are solely the responsibility of the grantee and do not necessarily represent the official views of the sponsors.

Multiple individuals provided invaluable information, including current and former employees of Georgia Power, Southern Company, Atmospheric Research & Analysis, Inc., Georgia EPD, and U.S. EPA Region IV. We thank Eric Edgerton and Charlie Blanchard for their input on the use of SEARCH data and trend analyses. We also would like to thank Prof. Cory Zigler for discussions on various aspects of accountability research.

REFERENCES

Anderson DC, Loughner CP, Diskin G, Weinheimer A, Canty TP, Salawitch RJ, et al. 2014. Measured and modeled CO and NOy in DISCOVER-AQ: An evaluation of emissions and chemistry over the eastern U.S. Atmos Environ 96:78–87; doi:10.1016/j.atmosenv.2014.07.004.

Ansari AS, Pandis SN. 1999. An Analysis of Four Models Predicting the Partitioning of Semivolatile Inorganic Aerosol Components. Aerosol Sci Technol 31:129–153; doi:10.1080/027868299304200.

Appel KW, Napelenok S, Hogrefe C, Pouliot G, Foley KM, Roselle SJ, et al. 2018. Overview and evaluation of the Community Multiscale Air Quality (CMAQ) Modeling System Version 5.2. In: Air Pollution Modeling and its Application XXV. ITM 2016. Springer Proceedings in Complexity (Mensink C and Kallos G, eds). Cham, Switzerland:Springer.

Atkinson RW, Barratt B, Armstrong B, Anderson HR, Beevers SDD, Mudway ISS, et al. 2009. The impact of the congestion charging scheme on ambient air pollution concentrations in London. Atmos Environ 43:5493–5500; doi:10.1016/j.atmosenv.2009.07.023.

Balachandran S, Pachon JE, Hu Y, Lee D, Mulholland JA, Russell AG. 2012. Ensemble-trained source apportionment of fine particulate matter and method uncertainty analysis. Atmos Environ 61:387–394; doi:10.1016/j.atmosenv.2012.07.031.

Bates JT, Weber RJ, Abrams J, Verma V, Fang T, Klein M, et al. 2015. Reactive oxygen species generation linked to sources of atmospheric particulate matter and cardiorespiratory effects. Environ Sci Technol 49:13605–13612; doi:10.1021/acs.est.5b02967.

Bencala KE, Seinfeld JH. 1979. An air quality model performance assessment package. Atmos Environ 13:1181–1185; doi:10.1016/0004-6981(79)90043-X.

Bergin MS, West JJ, Keating TJ, Russell AG. 2005. Regional atmospheric pollution and transboundary air quality management. Annu Rev Environ Resour 30:1–37; doi:10.1146/annurev.energy.30.050504.144138.

Blanchard CL, Hidy GM. 2005. Effects of SO_2 and NO_x emission reductions on $PM_{2.5}$ mass concentrations in the southeastern United States. J Air Waste Manage Assoc 55:265–272; doi:10.1080/10473289.2005.10464624.

Blanchard CL, Hidy GM, Shaw S, Baumann K, Edgerton ES. 2016. Effects of emission reductions on organic aerosol in the southeastern United States. Atmos Chem Phys 16:215–238; doi:10.5194/acp-16-215-2016.

Blanchard CL, Hidy GM, Tanenbaum S. 2010. NMOC, ozone, and organic aerosol in the southeastern United States, 1999–2007: 2. Ozone trends and sensitivity to NMOC emissions in Atlanta, Georgia. Atmos Environ 44:4840–4849; doi:10.1016/j.atmosenv.2010.07.030.

Blanchard CL, Hidy GM, Tanenbaum S. 2014. Ozone in the southeastern United States: An observation-based model using measurements from the SEARCH network. Atmos Environ 88:192–200; doi:10.1016/j.atmosenv.2014.02.006.

Blanchard CL, Hidy GM, Tanenbaum S, Edgerton ES, Hartsell BE. 2013. The Southeastern Aerosol Research and

Characterization (SEARCH) study: Spatial variations and chemical climatology, 1999–2010. J Air Waste Manage Assoc 63:260–275; doi:10.1080/10962247.2012.749816.

Blanchard CL, Tanenbaum S, Hidy GM. 2012. Source contributions to atmospheric gases and particulate matter in the southeastern United States. Environ Sci Technol 46:5479–5488; doi:10.1021/es203568t.

Boersma KF, Eskes HJ, Brinksma EJ. 2004. Error Analysis for tropospheric NO_2 retrieval from space. J Geophys Res 109; doi:10.1029/2003JD003962.

Boersma KF, Eskes HJ, Dirksen RJ, van Der A RJ, Veefkind JP, Stammes P, et al. 2011. An improved tropospheric NO_2 column retrieval algorithm for the Ozone Monitoring Instrument. Atmos Meas Tech 4:1905–1928; doi:10.5194/amt-4-1905-2011.

Boogaard H, Janssen NAH, Fischer PH, Kos GPA, Weijers EP, Cassee FR, et al. 2012. Impact of low emission zones and local traffic policies on ambient air pollution concentrations. Sci Total Environ 435-436:132–140; doi:10.1016/j.scitotenv.2012.06.089.

Brock CA, Washenfelder RA, Trainer M, Ryerson TB, Wilson JC, Reeves JM, et al. 2002. Particle growth in the plumes of coal-fired power plants. J Geophys Res 107:AAC 9-1–AAC 9-14.

Brunekreef B, Holgate ST. 2002. Air pollution and health. Lancet 360:1233–1242; doi:10.1016/S0140-6736(02)11274-8.

Byun D, Schere KL. 2006. Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system. Appl Mech Rev 59:51; doi:10.1115/1.2128636.

Camalier L, Cox W, Dolwick P. 2007. The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. Atmos Environ 41:7127–7137; doi:10.1016/j.atmosenv.2007.04.061.

Chen Y, Zheng M, Edgerton ES, Ke L, Sheng G, Fu J. 2012. $PM_{2.5}$ source apportionment in the southeastern U.S.: Spatial and seasonal variations during 2001–2005. J Geophys Res 117(D8):304; doi:10.1029/2011JD016572.

Clancy L, Goodman P, Sinclair H, Dockery DW. 2002. Effect of air-pollution control on death rates in Dublin, Ireland: An intervention study. Lancet 360:1210–1214; doi:10.1016/S0140-6736(02)11281–11285.

Close B, Zurbenko I. 2013. kza: Kolmogorov-Zurbenko Adaptive Filters. R package version 3.0.0. Available: *http://CRAN.R-project.org/package=kza?*

CMAS Center 2013. SMOKE version 3.5.1. The institute for the Environment – The University of North Carolina at Chapel Hill. Available: *www.cmascenter.org/help/documentation.cfm?model=smoke&version=3.5.1*.

Cohan DS, Hakami A, Hu Y, Russell AG. 2005. Nonlinear response of ozone to emissions: Source apportionment and sensitivity analysis. Environ Sci Technol 39:6739–6748.

Cohen AJ, Ross Anderson H, Ostro B, Pandey KD, Krzyzanowski M, Künzli N, et al. 2005. The global burden of disease due to outdoor air pollution. J Toxicol Environ Heal Part A 68:1301–1307; doi:10.1080/15287390590936166.

Cox WM, Chu S-H. 1993. Meteorologically adjusted ozone trends in urban areas: A probabilistic approach. Atmos Environ 27:425–434; doi:10.1016/0957-1272(93)90019-3.

Craig CA. 2016. Energy consumption, energy efficiency, and consumer perceptions: A case study for the southeast United States. Appl Energy 165:660–669; doi:10.1016/j.apenergy.2015.12.069.

Cropper M, Fraas A, Morgenstern R. 2017. Looking backward to move regulations forward. Science 355:1375–1376.

Curtis L, Rea W, Smith-Willis P, Fenyves E, Pan Y. 2006. Adverse health effects of outdoor air pollutants. Environ Int 32:815–830; doi:10.1016/j.envint.2006.03.012.

Dallmann TR, Harley RA. 2010. Evaluation of mobile source emission trends in the United States. J Geophys Res 115(D14):305; doi:10.1029/2010JD013862.

Darrow LA, Hess J, Rogers CA, Tolbert PE, Klein M, Sarnat SE. 2012. Ambient pollen concentrations and emergency department visits for asthma and wheeze. J Allergy Clin Immunol 130(3): 630–638; doi:10.1016/j.jaci.2012.06.020.

Darrow LA, Klein M, Sarnat JA, Mulholland JA, Strickland MJ, Sarnat SE, et al. 2009. The use of alternative pollutant metrics in time-series studies of ambient air pollution and respiratory emergency department visits. J Expo Sci Environ Epidemiol 21:10–9; doi:10.1038/jes.2009.49.

Darrow LA, Klein M, Sarnat JA, Mulholland JA, Strickland MJ, Sarnat SE, et al. 2011. The use of alternative pollutant metrics in time-series studies of ambient air pollution and respiratory emergency department visits. J Expo Sci Environ Epidemiol 21:10–19; doi:10.1038/jes.2009.49.

Deguillaume L, Beekmann M, Menut L. 2007. Bayesian Monte Carlo analysis applied to regional-scale inverse emission modeling for reactive trace gases. J Geophys Res Atmos 112; doi:10.1029/2006JD007518.

D'Onofrio D, Kim B-U, Kim Y, Kim K-H. 2016. Atlanta Roadside Emissions Exposure Study. Atlanta Regional Commission. *http://atlantaregional.org/wp-content/uploads/2017/03/arees-documentation.pdf*.

Dowle M, Srinivasan A. 2017. data.table: Extension of 'data.frame'. R package version 1.10.4-2. Available: *https://CRAN.R-project.org/package=data.table*.

Emery C, Liu Z, Russell A, Talat Odman M, Yarwood G, Kumar N. 2016. Recommendations on statistics and benchmarks to assess photochemical model performance. J Air Waste Manage Assoc 67(5):582–598, doi: 10.1080/10962247.2016.1265027.

Eskridge RE, Ku JY, Rao STT, Porter PS, Zurbenko IG. 1997. Separating different scales of motion in time series of meteorological variables. Bull Am Meteorol Soc 78:1473–1483; doi:10.1175/1520-0477(1997)078<1473:SDSOMI>2.0.CO;2.

European Space Agency. 2016. Tropospheric NO_2 from satellites, *www.temis.nl/airpollution/no2.html*. [accessed September 28, 2016].

Fang T, Verma V, Bates JT, Abrams J, Klein M, Strickland JM, et al. 2016. Oxidative potential of ambient water-soluble $PM_{2.5}$ in the southeastern United States: Contrasts in sources and health associations between ascorbic acid (AA) and dithiothreitol (DTT) assays. Atmos Chem Phys 16:3865–3879; doi:10.5194/acp-16-3865-2016.

Flanders WD, Klein M, Darrow LA, Strickland MJ, Sarnat SE, Sarnat JA, et al. 2011. A method to detect residual confounding in spatial and other observational studies. Epidemiology 22:823–826; doi:10.1097/EDE.0b013e3182305dac.

Flanders WD, Strickland MJ, Klein M. 2017. A new method for partial correction of residual confounding in time-series and other observational studies. Am J Epidemiol 185:941–949.

Foley KM, Dolwick P, Hogrefe C, Simon H, Timin B, Possiel N. 2015a. Dynamic evaluation of CMAQ part II: Evaluation of relative response factor metrics for ozone attainment demonstrations. Atmos Environ 103:188–195; doi:10.1016/j.atmosenv.2014.12.039.

Foley KM, Hogrefe C, Pouliot G, Possiel N, Roselle SJ, Simon H, et al. 2015b. Dynamic evaluation of CMAQ part I: Separating the effects of changing emissions and changing meteorology on ozone levels between 2002 and 2005 in the eastern U.S. Atmos Environ 103:247–255; doi:10.1016/j.atmosenv.2014.12.038.

Friedman M, Powell K, Hutwagner L, Graham LM, Teague WG. 2001. Impact of changes in transportation and commuting behaviors during the 1996 Summer Olympic Games in Atlanta on air quality and childhood asthma. JAMA 285:897–905.

Frost G. 2016. Understanding and improving emissions information through atmospheric observations and models. Community Model Anal Syst Conf. Available: *www.cmascenter.org/conference/2016/agenda.cfm*.

Gass K, Balachandran S, Chang HH, Russell AG, Strickland MJ. 2015. Ensemble-based source apportionment of fine particulate matter and emergency department visits for pediatric asthma. Am J Epidemiol 181:504–512; doi:10.1093/aje/kwu305.

Gauderman WJ, Urman R, Avol E, Berhane K, McConnell R, Rappaport E, et al. 2015. Association of improved air quality with lung development in children. N Engl J Med 372:905–913; doi:10.1056/NEJMoa1414123.

Gégo E, Gilliland A, Godowitch J, Rao S, Porter P, Hogrefe C. 2008. Modeling analyses of the effects of changes in nitrogen oxides emissions from the electric power sector on ozone levels in the eastern United States. J Air Waste Manage Assoc 58:580–588; doi:10.3155/1047-3289.58.4.580.

Gégo E, Porter PS, Gilliland A, Rao ST. 2007. Observation-based assessment of the impact of nitrogen oxides emissions reductions on ozone air quality over the eastern United States. J Appl Meteorol Climatol 46:994–1008; doi:10.1175/JAM2523.1.

Georgia Department of Transportation. 2017. 455 Reports — Mileage by Route Type and Functional Classification. Available: *http://www.dot.ga.gov/DS/Data#tab-2*.

Georgia Environmental Protection Division. 2009. Georgia's State Implementation Plan for the Atlanta 8-Hour Ozone Nonattainment Area. Georgia Department of Natural Resources Environmental Protection Division Air Protection Branch.

Georgia Environmental Protection Division. 2013. Rules for Air Quality Control. Georgia Department of Natural Resources Environmental Protection Division Air Protection Branch. Available: *http://rules.sos.ga.gov/nllxml/georgiacodesGetcv.aspx?urlRedirected=yes&data=admin&lookingfor=391-3-1*.

Georgia Power Company. 2007. Georgia Power Company's 2007 Integrated Resource Plan. Georgia Public Service Commission Docket #24505. Available: *www.psc.state.ga.us/factsv2/Docket.aspx?docketNumber=24505*.

Georgia Power Company. 2013. Georgia Power Company's 2013 Integrated Resource Plan. Georgia Public Service Commission Docket #36498. Available: *www.psc.state.ga.us/factsv2/Docket.aspx?docketNumber=36498*.

Gilliland AB, Hogrefe C, Pinder RW, Godowitch JM, Foley KL, Rao ST. 2008. Dynamic evaluation of regional air quality models: Assessing changes in O_3 stemming from changes in emissions and meteorology. Atmos Environ 42:5110–5123; doi:10.1016/j.atmosenv.2008.02.018.

Goldberg DL, Vinciguerra TP, Anderson DC, Hembeck L, Canty TP, Ehrman SH, et al. 2016. CAMx ozone source attribution in the eastern United States using guidance from observations during DISCOVER-AQ Maryland. Geophys Res Lett 2249–2258; doi:10.1002/2015GL067332.

Goldman GT, Mulholland JA, Russell AG, Strickland MJ, Klein M, Waller LA, et al. 2011. Impact of exposure measurement error in air pollution epidemiology: Effect of error type in time-series studies. Env Heal 10:61; doi:10.1186/1476-069x-10-61.

Goodman PG, Rich DQ, Zeka A, Clancy L, Dockery DW. 2009. Effect of air pollution controls on black smoke and sulfur dioxide concentrations across Ireland. J Air Waste Manage Assoc 59:207–213; doi:10.3155/1047-3289.59.1.207.

Götschi T, Heinrich J, Sunyer J, Künzli N. 2008. Long-term effects of ambient air pollution on lung function: A review. Epidemiology 19:690–701; doi:10.1097/EDE.0b013e318181650f.

Hanna SR, Lu Z, Frey HC, Wheeler N, Vukovich J, Arunachalam S, et al. 2001. Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. Atmos Environ 35:891–903; doi:10.1016/S1352-2310(00)00367-8.

Hansen DA, Edgerton ES, Hartsell BE, Jansen JJ, Kandasamy N, Hidy GM, et al. 2003. The Southeastern Aerosol Research and Characterization Study: Part 1 — Overview. J Air Waste Manage Assoc 53:1460–1471; doi:10.1080/10473289.2003.10466318.

Harrington W, Morgenstern R, Shih J-S, Bell ML. 2012. Did the Clean Air Act Amendments of 1990 really improve air quality? Air Qual Atmos Health 5:353–367; doi:10.1007/s11869-012-0176-5.

Health Effects Institute. 2010. Proceedings of an HEI Workshop on Further Research to Assess the Health Impacts of Actions Taken to Improve Air Quality. Communication 15. Boston, MA:Health Effects Institute.

Hedley AJ, Wong C-M, Thach TQ, Ma S, Lam T-H, Anderson HR. 2002. Cardiorespiratory and all-cause mortality after restrictions on sulphur content of fuel in Hong Kong: An intervention study. Lancet 360:1646–1652.

HEI Accountability Working Group. 2003. Assessing Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research. Communication 11. Boston, MA:Health Effects Institute.

Henneman LRF, Chang HH, Liao KJ, Lavoué D, Mulholland JA, Russell AG. 2017a. Accountability assessment of regulatory impacts on ozone and $PM_{2.5}$ concentrations using statistical and deterministic pollutant sensitivities. Air Qual Atmos Health 10:695–711. Available: *https://link.springer.com/article/10.1007/s11869-017-0463-2*.

Henneman LR, Holmes HA, Mulholland JA, Russell AG. 2015. Meteorological detrending of primary and secondary pollutant concentrations: Method application and evaluation using long-term (2000–2012) data in Atlanta. Atmos Environ 119:201–210; doi:10.1016/j.atmosenv.2015.08.007.

Henneman LRF, Liu C, Hu Y, Mulholland JA, Russell AG. 2017b. Air quality modeling for accountability research: Operational, dynamic, and diagnostic evaluation. Atmos Environ 166:551–565. Available: *https://doi.org/10.1016/j.atmosenv.2017.07.049*.

Henneman LRF, Liu C, Mulholland JA, Russell AG. 2016. Evaluating the effectiveness of air quality regulations: A review of accountability studies and frameworks. J Air Waste Manage Assoc 67:144–172; doi:10.1080/10962247.2016.1242518.

Henschel S, Atkinson R, Zeka A, Le Tertre A, Analitis A, Katsouyanni K, et al. 2012. Air pollution interventions and their impact on public health. Int J Public Health 57:757–768; doi:10.1007/s00038-012-0369-6.

Hidy GM, Blanchard CL. 2015. Precursor reductions and ground-level ozone in the continental United States. J Air Waste Manag Assoc 65:1261–1282; doi:10.1080/10962247.2015.1079564.

Hogrefe C, Rao ST. 2000. Interpreting the information in ozone observations and model predictions relevant to regulatory policies in the eastern United States. Bull Am Met Soc 81:2083–2106.

Hu Y, Balachandran S, Pachon JE, Baek J, Ivey C, Holmes H, et al. 2014. Fine particulate matter source apportionment using a hybrid chemical transport and receptor model approach. Atmos Chem Phys 14:5415–5431; doi:10.5194/acp-14-5415-2014.

Kampa M, Castanas E. 2008. Human health effects of air pollution. Environ Pollut 151:362–367; doi:10.1016/j.envpol.2007.06.012.

Kleinman MT, Phalen RF, Mautz WJ, Mannix RC, McClure TR, Crocker TT. 1989. Health effects of acid aerosols formed by atmospheric mixtures. Environ Health Perspect 79:137–145.

Konovalov IB, Beekmann M, Richter A, Burrows JP. 2006. Inverse modelling of the spatial distribution of NO_x emissions on a continental scale using satellite data. Atmos Chem Phys 6:1747–1770; doi:10.5194/acp-6-1747-2006.

Koo B, Kumar N, Knipping E, Nopmongcol U, Sakulyanontvittaya T, Odman MT, et al. 2015. Chemical transport model consistency in simulating regulatory outcomes and the relationship to model performance. Atmos Environ 116:159–171; doi:10.1016/j.atmosenv.2015.06.036.

Krall JR, Mulholland JA, Russell AG, Balachandran S, Winquist A, Tolbert PE, et al. 2016. Associations between source-specific fine particulate matter and emergency department visits for respiratory disease in four U.S. cities. Environ Health Perspect 125:97–103; doi:10.1289/EHP271.

Kuebler J, van den Bergh H, Russell AG. 2001. Long-term trends of primary and secondary pollutant concentrations in Switzerland and their response to emission controls and economic changes. Atmos Environ 35:1351–1363; doi:10.1016/S1352-2310(00)00401-5.

Laden F, Schwartz J, Speizer FE, Dockery DW. 2006. Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. Am J Respir Crit Care Med 173:667–672; doi:10.1164/rccm.200503-443OC.

Levy JI, Woo MK, Tambouret Y. 2016. Energy savings and emissions reductions associated with increased insulation for new homes in the United States. Build Environ 96:72–79; doi:10.1016/j.buildenv.2015.11.008.

Li Y, Wang W, Kan H, Xu X, Chen B. 2010. Air quality and outpatient visits for asthma in adults during the 2008 Summer Olympic Games in Beijing. Sci Total Environ 408:1226–1227; doi:10.1016/j.scitotenv.2009.11.035.

Liao K-J, Tagaris E, Napelenok SL, Manomaiphiboon K, Woo J-H, Amar P, et al. 2008. Current and future linked responses of ozone and $PM_{2.5}$ to emission controls. Environ Sci Technol 42:4670–4675.

Lin H, Zhang Y, Liu T, Xiao J, Xu Y, Xu X, et al. 2014. Mortality reduction following the air pollution control measures during the 2010 Asian games. Atmos Environ 91:24–31; doi:10.1016/j.atmosenv.2014.03.051.

Lin S, Jones R, Pantea C, Özkaynak H, Rao STT, Hwang S-A, et al. 2013a. Impact of NO_x emissions reduction policy on hospitalizations for respiratory disease in New York State. J Expo Sci Environ Epidemiol 23:73–80; doi:10.1038/jes.2012.69.

Lin Y, Huang K, Zhuang G, Fu JS, Xu C, Shen J, et al. 2013b. Air quality over the Yangtze River delta during the 2010 Shanghai expo. Aerosol Air Qual Res 13:1655–1666; doi:10.4209/aaqr.2012.11.0312.

Lipfert FW, Wyzga RE. 1993. On the spatial and temporal variability of aerosol acidity and sulfate concentration. J Air Waste Manage Assoc 43:489–491.

Marais EA, Jacob DJ, Jimenez JL, Campuzano-Jost P, Day DA, Hu W, et al. 2016. Aqueous-phase mechanism for secondary organic aerosol formation from isoprene: Application to the southeast United States and co-benefit of SO_2 emission controls. Atmos Chem Phys 16:1603–1618; doi:10.5194/acp-16-1603-2016.

Mauderly JL, Burnett RT, Castillejos M, Özkaynak H, Samet JM, Stieb DM, et al. 2010. Is the air pollution health research community prepared to support a multipollutant air quality management framework? Inhal Toxicol 22:1–19; doi:10.3109/08958371003793846.

Mauderly JL, Samet JM. 2009. Is there evidence for synergy among air pollutants in causing health effects? Environ Health Perspect 117:1–6; doi:10.1289/ehp.11654.

McDonald BC, Dallmann TR, Martin EW, Harley RA. 2012. Long-term trends in nitrogen oxide emissions from motor vehicles at national, state, and air basin scales. J Geophys Res Atmos 117: D00V18; doi:10.1029/2012JD018304.

McDonald BC, Gentner DR, Goldstein AH, Harley RA. 2013. Long-term trends in motor vehicle emissions in U.S. urban areas. Environ Sci Technol 47:10022–10031; doi:10.1021/es401034z.

Meskhidze N, Chameides WL, Nenes A, Chen G. 2003. Iron mobilization in mineral dust: Can anthropogenic SO_2 emissions affect ocean productivity? Geophys Res Lett 30:2085; doi:10.1029/2003GL018035.

Metzger K, Tolbert P, Klein M, Peel J, Flanders W, Todd K, et al. 2004. Ambient air pollution and cardiovascular emergency department visits. Epidemiology 15:46–56.

Morgenstern RD, Harrington W, Shih J-S, Bell ML. 2012. Accountability Analysis of Title IV Phase 2 of the 1990 Clean Air Act Amendments. Research Report 168. Boston, MA:Health Effects Institute.

Mozurkewich M. 1993. The dissociation constant of ammonium nitrate and its dependence on temperature, relative humidity and particle size. Atmos Environ Part A Gen Top 27:261–270; doi:10.1016/0960-1686(93)90356-4.

Muller N, Tong D, Mendelsohn RO. 2009. Regulating NO_x and SO_2 emissions in Atlanta. BE J Econ Anal Policy 9(2):1–32; doi: *https://doi.org/10.2202/1935-1682.1954*.

Napelenok SL, Foley KM, Kang D, Mathur R, Pierce T, Rao ST. 2011. Dynamic evaluation of regional air quality model's response to emission reductions in the presence of uncertain emission inventories. Atmos Environ 45:4091–4098; doi:10.1016/j.atmosenv.2011.03.030.

Napelenok SL, Pinder RW, Gilliland AB, Martin RV. 2008. A method for evaluating spatially-resolved NO_x emissions using Kalman filter inversion, direct sensitivities, and space-based NO_2 observations. Atmos Chem Phys Discuss 8:6469–6499.

National Research Council. 2004. Air Quality Management in the United States. Washington, DC:National Academies Press.

Nenes A. 1998. ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols. Aquat Geochemistry 4:123–152.

O'Lenick CR, Winquist A, Mulholland JA, Friberg MD, Chang HH, Kramer MR, et al. 2017. Assessment of neighbourhood-level socioeconomic status as a modifier of air pollution–asthma associations among children in Atlanta. J Epidemiol Comm Health 71:129–136.

Pachon JE, Balachandran S, Hu Y, Mulholland JA, Darrow LA, Sarnat JA, et al. 2012. Development of outcome-based, multipollutant mobile source indicators. J Air Waste Manage Assoc 62:431; doi:10.1080/10473289.2012.656218.

Parrish DD. 2006. Critical evaluation of U.S. on-road vehicle emission inventories. Atmos Environ 40:2288–2300; doi:10.1016/j.atmosenv.2005.11.033.

Peel JL, Klein M, W. Flanders WD, Mulholland JA, Tolbert PE. 2010. Impact of Improved Air Quality During the 1996 Summer Olympic Games in Atlanta on Multiple Cardiovascular and Respiratory Outcomes. Research Report 148. Boston, MA:Health Effects Institute.

Peel JL, Metzger KB, Klein M, Flanders WD, Mulholland JA, Tolbert PE. 2007. Ambient air pollution and cardiovascular emergency department visits in potentially sensitive groups. Am J Epidemiol 165:625–633; doi:10.1093/aje/kwk051.

Peel JL, Tolbert P, Klein M, Metzger K, Flanders W, Todd K, et al. 2005. Ambient air pollution and respiratory emergency department visits. Epidemiology 16:164–174.

Peters A, Breitner S, Cyrys J, Stölzel M, Pitz M, Wölke G, et al. 2009. The influence of improved air quality on mortality risks in Erfurt, Germany. HEI Research Report 137, Boston, MA:Health Effects Institute.

Peters J, Hedley AJ, Wong CM, Lam TH, Ong SG, Liu J, et al. 1996. Effects of ambient air pollution and environmental tobacco smoke on respiratory health of nonsmoking women in Hong Kong. Int J Epidemiol 28:859; doi:10.1093/ije/28.5.859.

Pleim J. 2016. A new version of the Community Multiscale Air Quality Model: CMAQv5.2. Community Model. Anal. Syst. Conf. Available: *www.cmascenter.org/conference/2016/agenda.cfm*.

Pope CA, Burnett RT, Thurston GD, Thun MJ, Calle EE, Krewski D, et al. 2004. Cardiovascular mortality and long-term exposure to particulate air pollution: Epidemiological evidence of general pathophysiological pathways of disease. Circulation 109:71–77; doi:10.1161/01.CIR.0000108927.80044.7F.

Pope CA, Ezzati M, Dockery DW. 2012. Validity of observational studies in accountability analyses: The case of air pollution and life expectancy. Air Qual Atmos Health 5:231–235; doi:10.1007/s11869-010-0130-3.

R Development Core Team. 2012. R: A Language and Environment for Statistical Computing Vienna Austria:R Foundation for Statistical Computing.

Raaschou-Nielsen O, Andersen ZJ, Beelen R, Samoli E, Stafoggia M, Weinmayr G, et al. 2013. Air pollution and lung cancer incidence in 17 European cohorts: Prospective analyses from the European Study of Cohorts for Air Pollution

Effects (ESCAPE). Lancet Oncol 14:813–822; doi:10.1016/S1470-2045(13)70279-1.

Rao S, Zurbenko I. 1994. Detecting and tracking changes in ozone air quality. J Air Waste Mgmt Assoc 44:1089–1092; doi: 10.1080/10473289.1994.10467303.

Rich DQ, Kipen HM, Huang W, Wang G, Wang Y, Zhu P, et al. 2012. Association between changes in air pollution levels during the Beijing Olympics and biomarkers of inflammation and thrombosis in healthy young adults. JAMA 307:2068–2078; doi:10.1001/jama.2012.3488.

Sahinler S, Topuz D. 2007. Bootstrap and jackknife resampling algorithms for estimation of regression parameters. J Appl Quant Methods 2:188–199.

Sandhu GS, Frey HC, Bartelt-Hunt S, Jones E. 2014. In-use measurement of the activity, fuel use, and emissions of front-loader refuse trucks. Atmos Environ 92:557–565; doi:10.1016/j.atmosenv.2014.04.036.

Sarnat JA, Russell A, Liang D, Moutinho J, Golan R, Weber RJ, et al. In press. Developing Multipollutant Exposure Indicators of Traffic Pollution: The Dorm Room Inhalation to Vehicle Emissions (DRIVE) Study. Research Report 196. Boston, MA:Health Effects Institute.

Sarnat SE, Klein M, Sarnat JA, Flanders WD, Waller LA, Mulholland JA, et al. 2010. An examination of exposure measurement error from air pollutant spatial variability in time-series studies. J Expo Sci Environ Epidemiol 20:135–146; doi:10.1038/jes.2009.10.

Sarnat SE, Sarnat JA, Mulholland J, Isakov V, Ozkaynak H, Chang HH, et al. 2013. Application of alternative spatiotemporal metrics of ambient air pollution exposure in a time-series epidemiological study in Atlanta. J Expo Sci Environ Epidemiol 23:593–605; doi:10.1038/jes.2013.41.

Saylor R, Myles L, Sibble D, Caldwell J, Xing J. 2015. Recent trends in gas-phase ammonia and $PM_{2.5}$ ammonium in the southeast United States. J Air Waste Manage Assoc 65:347–357; doi:10.1080/10962247.2014.992554.

Schwartz J, Marcus A. 1990. Mortality and air pollution in London: a time series analysis. Am J Epidemiol 131:185–194.

Schwartz J, Zanobetti A. 2000. Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. Epidemiology 11:666–672; doi:10.1097/00001648-200011000-00009.

Seaton A, MacNee W, Donaldson K, Godden D. 1995. Particulate air pollution and acute health effects. Lancet 345:176–178; doi:10.1016/S0140-6736(95)90173-6.

Seinfeld JH, Pandis SN. 2006. Atmospheric Chemistry and Physics: From Air Pollution to Climate Change. Hoboken, NJ:John Wiley & Sons.

Simon H, Baker KR, Phillips S. 2012. Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012. Atmos Environ 61:124–139; doi:10.1016/j.atmosenv.2012.07.012.

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, et al. 2008. A Description of the Advanced Research WRF Version 3. NCAR Tech. Note NCAR/TN-475+STR. doi:10.5065/D68S4MVH. Available: *http://opensky.ucar.edu/islandora/object/technotes:500* [accessed 7 December 2017].

Snyder M, Arunachalam S, Isakov V, Talgo K, Naess B, Valencia A, et al. 2014. Creating locally-resolved mobile-source emissions inputs for air quality modeling in support of an exposure study in Detroit, Michigan, USA. Int J Environ Res Public Health 11:12739–12766; doi:10.3390/ijerph111212739.

Snyder MG, Venkatram A, Heist DK, Perry SG, Petersen WB, Isakov V. 2013. RLINE: A line source dispersion model for near-surface releases. Atmos Environ 77:748–756; doi:10.1016/j.atmosenv.2013.05.074.

Souri AH, Choi Y, Jeon W, Li X, Pan S, Diao L, et al. 2016. Constraining NO_x emissions using satellite NO_2 measurements during 2013 DISCOVER-AQ Texas campaign. Atmos Environ 131:371–381; doi:10.1016/j.atmosenv.2016.02.020.

Stafoggia M, Cesaroni G, Peters A, Andersen ZJ, Badaloni C, Beelen R, et al. 2014. Long-term exposure to ambient air pollution and incidence of cerebrovascular events: Results from 11 European cohorts within the ESCAPE project. Environ Health Perspect 122:919–925; doi:10.1289/ehp.1307301.

Strickland MJ, Darrow LA, Klein M, Flanders WD, Sarnat JA, Waller LA, et al. 2010. Short-term associations between ambient air pollutants and pediatric asthma emergency department visits. Am J Respir Crit Care Med 182:307–316; doi:10.1164/rccm.200908-1201OC.

Strickland MJ, Darrow LA, Mulholland JA, Klein M, Flanders WD, Winquist A, et al. 2011. Implications of different approaches for characterizing ambient air pollutant concentrations within the urban airshed for time-series studies and health benefits analyses. Environ Health 10:36; doi:10.1186/1476-069X-10-36.

Strickland MJ, Gass KM, Goldman GT, Mulholland JA. 2013. Effects of ambient air pollution measurement error on health effect estimates in time-series studies: A simulation-based analysis. J Expo Sci Environ Epidemiol 25(2):160–166; doi:10.1038/jes.2013.16.

Strickland MJ, Klein M, Flanders WD, Chang HH, Mulholland JA, Tolbert PE, et al. 2014. Modification of the effect of ambient air pollution on pediatric asthma emergency visits: Susceptible subpopulations. Epidemiology 25:843–850; doi:10.1097/ede.0000000000000170.

Tolbert PE, Klein M, Metzger KB, Peel J, Flanders WD, Todd K, et al. 2000. Interim results of the Study of Particulates and Health in Atlanta (SOPHIA). J Expo Anal Environ Epidemiol 10:446–460; doi:10.1038/sj.jea.7500106.

Tolbert PE, Klein M, Peel JL, Sarnat SE, Sarnat JA. 2007. Multipollutant modeling issues in a study of ambient air quality and emergency department visits in Atlanta. J Expo Sci Environ Epidemiol 17 Suppl 2:S29–35; doi:10.1038/sj.jes.7500625.

Urbanek S, Horner J. 2015. Cairo: R graphics device using Cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDF, SVG, PostScript) and display (X11 and Win32) output. R package version 1.5-9. Available: *https://CRAN.R-project.org/package=Cairo*.

U.S. Energy Information Administration. 2014. Georgia Prime Supplier Sales Volumes of Petroleum Products. Available: *www.eia.gov/dnav/pet/pet_cons_prim_dcu_SGA_m.htm*.

U.S. Environmental Protection Agency. 1997. Regulatory Impact Analysis: Control of Emissions of Air Pollution from Highway Heavy-Duty Engines. EPA420-R-00-026.

U.S. Environmental Protection Agency. 1998. Regulatory Impact Analysis for the NBP. EPA-452/R-98-003. Available: *www3.epa.gov/ttn/ecas/docs/ria/transport_ria_final-nox-sip-call-vol2_1998-12.pdf*.

U.S. Environmental Protection Agency. 1999. Regulatory impact analysis — control of air pollution from new motor vehicles: Tier 2 motor vehicle emissions standards and gasoline sulfur control requirements. EPA420-R-99-023. Available: *https://nepis.epa.gov/Exe/ZyPDF.cgi/P100F1UV.PDF?Dockey=P100F1UV.PDF*.

U.S. Environmental Protection Agency. 2000. Technical Support Document for the Heavy-Duty Engine and Vehicle Standards and Highway Diesel Fuel Sulfur Control Requirements: Air Quality Modeling Analyses. EPA420-R-00-028. Available: *https://nepis.epa.gov/Exe/ZyPDF.cgi/P1001DPE.PDF?Dockey=P1001DPE.PDF*.

U.S. Environmental Protection Agency. 2002. Acid Rain Program 2001 Summary Report. Available: EPA-430-R-02-009. *www.epa.gov/sites/production/files/2015-08/documents/2001report_0.pdf*.

U.S. Environmental Protection Agency. 2003. NO_x Budget Trading Program 2003 Progress and Compliance Report. EPA-430-R-04-010. Available: *www.epa.gov/sites/production/files/2015-08/documents/noxreport03.pdf* (hosted here: *www.epa.gov/airmarkets/nox-budget-trading-program-historical-reports*)

U.S. Environmental Protection Agency. 2005a. Regulatory Impact Analysis for the Final Clean Air Interstate Rule. EPA-452/R-05-002. Available: *www3.epa.gov/ttn/ecas/docs/ria/transport_ria_final-clean-air-interstate-rule_2005-03.pdf* (hosted here: *www.epa.gov/economic-and-cost-analysis-air-pollution-regulations/regulatory-impact-analyses-air-pollution*)

U.S. Environmental Protection Agency. 2005b. Regulatory Impact Analysis of the Final Clean Air Mercury Rule. EPA-452/R-05-003. Available: *www3.epa.gov/ttn/ecas/docs/ria/utilities_ria_final-mercury-rule_2005-03.pdf*.

U.S. Environmental Protection Agency. 2005c. Technical Support Document for the Final Clean Air Interstate Rule Air Quality Modeling. Office of Air Quality Planning and Standards. Available: *https://archive.epa.gov/airmarkets/programs/cair/web/pdf/finaltech02.pdf*.

U.S. Environmental Protection Agency. 2008. Final Rule to Remove Georgia from the NO_x SIP Call Requirements. Available: *www.gpo.gov/fdsys/search/pagedetails.action?granuleId=E8-8673&packageId=FR-2008-04-22&acCode=FR&collectionCode=FR*.

U.S. Environmental Protection Agency. 2011. The Benefits and Costs of the Clean Air Act from 1990 to 2020 Final Report. U.S. Environmental Protection Agency Office of Air and Radiation. Available: *www.epa.gov/sites/production/files/2015-07/documents/fullreport_rev_a.pdf*.

U.S. Environmental Protection Agency. 2012a. Motor Vehicle Emission Simulator (MOVES): User Guide Version, MOVES2010b. EPA-420-B-12-001b. Available: *https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P100EP28.pdf*.

U.S. Environmental Protection Agency. 2012b. MOtor Vehicle Emissions Simulator (MOVES). Available: *www.epa.gov/otaq/models/moves* [accessed September 28, 2016].

U.S. Environmental Protection Agency. 2014a. 2011 Version 6 Air Emissions Modeling Platforms. Washington, DC:Office of Air and Radiation. Office of Air Quality Planning and Standards. Air Quality Assessment Division. Available: *www.epa.gov/air-emissions-modeling/2011-version-60-platform*.

U.S. Environmental Protection Agency. 2014b. CMAQ version 5.0.2. U.S. EPA Atmospheric Modeling and Analysis Division. *www.cmascenter.org/cmaq/*.

U.S. Environmental Protection Agency. 2015. Regulatory Impact Analysis for the Clean Power Plan Final Rule. 343.

U.S. Environmental Protection Agency. 2016a. Air Markets Program Data. Available: *https://ampd.epa.gov/ampd/*. [accessed August, 2015].

U.S. Environmental Protection Agency. 2016b. Health Effects of Ozone in the General Population. *www.epa.gov/ozone-pollution-and-your-patients-health/health-effects-ozone-general-population*.

U.S. Environmental Protection Agency. 2016c. National Emissions Inventory. Available: *www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data*.

U.S. Environmental Protection Agency. 2017. Air Quality System Data Mart [internet database] available at *https://aqs.epa.gov/aqsweb/documents/data_mart_welcome.html*.

Utell MJ. 1985. Effects of inhaled acid aerosols on lung mechanics: An analysis of human exposure studies. Environ Health Perspect 63:39–44.

van Erp AM, Cohen AJ. 2009. HEI's Research Program on the Impact of Actions to Improve Air Quality: Interim Evaluation and Future Directions. Communication 14. Boston, MA:Health Effects Institute.

van Erp AM, Kelly FJ, Demerjian KL, Pope CA, Cohen AJ. 2011. Progress in research to assess the effectiveness of air quality interventions towards improving public health. Air Qual Atmos Health 5:217–230; doi:10.1007/s11869-010-0127-y.

Venables WN, Ripley BD. 2002. Modern Applied Statistics with S. Fourth Edition. ISBN 0-387-95457-0. New York:Springer.

Vijayaraghavan K, DenBleyker A, Ma L, Lindhjem C, Yarwood G. 2014. Trends in on-road vehicle emissions and ambient air quality in Atlanta, Georgia, USA, from the late 1990s through 2009. J Air Waste Manage Assoc 64:808–816; doi:10.1080/10962247.2014.892039.

Wang S, Zhao M, Xing J, Wu Y, Zhou Y, Lei Y, et al. 2010. Quantifying the air pollutants emission reduction during the 2008 Olympic Games in Beijing. Environ Sci Technol 44:2490–2496; doi:10.1021/es9028167.

Weber RJ, Guo H, Russell AG, Nenes A. 2016. High aerosol acidity despite declining atmospheric sulfate concentrations over the past 15 years. Nat Geosci 9:1–5; doi:10.1038/NGEO2665.

Winquist A, Grundstein A, Chang HH, Hess J, Sarnat SE. 2016. Warm season temperatures and emergency department visits in Atlanta, Georgia. Environ Res 147:314–323; doi:10.1016/j.envres.2016.02.022.

Winquist A, Kirrane E, Klein M, Strickland M, Darrow LA, Sarnat SE, et al. 2014. Joint effects of ambient air pollutants on pediatric asthma emergency department visits in Atlanta, 1998–2004. Epidemiology 25:666–673; doi:10.1097/EDE.0000000000000146.

Winquist A, Klein M, Tolbert P, Flanders WD, Hess J, Sarnat SE. 2012a. Comparison of emergency department and hospital admissions data for air pollution time-series studies. Environ Health 11:70; doi:10.1186/1476-069X-11-70.

Winquist A, Klein M, Tolbert P, Sarnat SE. 2012b. Power estimation using simulations for air pollution time-series studies. Environ Health 11:68; doi:10.1186/1476-069X-11-68.

Woodruff TJ, Parker JD, Darrow LA, Slama R, Bell ML, Choi H, et al. 2009. Methodological issues in studies of air pollution and reproductive health. Environ Res 109:311–320; doi:10.1016/j.envres.2008.12.012.

Xie Y, Elleman R, Jobson T, Lamb B. 2011. Evaluation of O_3- NO_x-VOC sensitivities predicted with the CMAQ photochemical model using Pacific Northwest 2001 field observations. J Geophys Res 116(D20):303; doi:10.1029/2011JD015801.

Xing J, Zhang Y, Wang S, Liu X, Cheng S, Zhang Q, et al. 2011. Modeling study on the air quality impacts from emission reductions and atypical meteorological conditions during the 2008 Beijing Olympics. Atmos Environ 45:1786–1798; doi:10.1016/j.atmosenv.2011.01.025.

Xu L, Guo H, Boyd CM, Klein M, Bougiatioti A, Cerully KM, et al. 2015. Effects of anthropogenic emissions on aerosol formation from isoprene and monoterpenes in the southeastern United States. Proc Natl Acad Sci 112:37–42; doi:10.1073/pnas.1417609112.

Yazdani Boroujeni B, Frey HC. 2014. Road grade quantification based on global positioning system data obtained from real-world vehicle fuel use and emissions measurements. Atmos Environ 85:179–186; doi:10.1016/j.atmosenv.2013.12.025.

Ye D, Klein M, Chang HH, Sarnat JA, Mulholland JA, Edgerton ES, et al. 2017. Estimating acute cardiorespiratory effects of ambient volatile organic compounds. Epidemiology 28(2):197–206. doi: 10.1097/EDE.00000000000 00607. Available: *http://journals.lww.com/epidem/Fulltext/2017/03000/Estimating_Acute_Cardiorespiratory_Effects_of.7.aspx*.

Zeger SL, Thomas D, Dominici F, Samet JM, Schwartz J, Dockery D, et al. 2000. Exposure measurement error in time-series studies of air pollution: Concepts and consequences. Environ Health Perspect 108:419–426; doi:10.1289/ehp.00108419.

Zhai X, Russell A, Sampath P, Mulholland J, Kim B-U, Kim Y, et al. 2016. Calibrating R-LINE Model Results with Observational Data to Develop Annual Mobile Source Air Pollutant Fields at Fine Spatial Resolution: Application in Atlanta. Atmos Environ 147:446–457. Available: *https://doi.org/10.1016/j.atmosenv.2016.10.015*.

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; doi:10.1093/aje/kwu263.

Zigler CM, Dominici F, Wang Y. 2012. Estimating causal effects of air quality regulations using principal stratification for spatially correlated multivariate intermediate outcomes. Biostatistics 13:289–302; doi:10.1093/biostatistics/kxr052.

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.

Zurbenko IG. 1991. Spectral analysis of nonstationary time series. Int Stat Rev 39:163–173.

HEI QUALITY ASSURANCE STATEMENT

The conduct of this study was subjected to independent audits by RTI International staff members Dr. Linda Brown and Dr. Prakash Doraiswamy. These staff members are experienced in quality assurance oversight for air quality monitoring, emission inventories and modeling, data analysis, and related epidemiological studies. Other participants on the RTI QA oversight team included Dr. Jeremy Aldworth, a statistician who reviewed the epidemiological statistical model codes.

The QA oversight program consisted of an initial on-site audit of the research study at Georgia Institute of Technology and at Emory University for conformance to the study protocol and standard operating procedures, and a final remote audit of the final report and the data processing steps. The onsite audit was performed by Drs. Brown and Doraiswamy. The final remote audit was performed by Drs. Brown, Aldworth, and Doraiswamy. The dates of the audits and reviews are listed below.

April 22–23, 2015 (Audit Phase 1, Georgia Institute of Technology and Emory University)

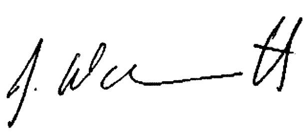
The auditors conducted an on-site audit at the Georgia Institute of Technology, School of Civil and Environmental Engineering, and at the Rollins School of Public Health at Emory University, Department of Environmental Health, Atlanta, GA. The audit reviewed the following study components: progress reports; personnel and staff; adequacy of equipment and facilities; internal quality assurance procedures; air quality data processing and documentation; health data processing and quality checks; and backup procedures. Program codes were inspected to verify proper documentation. Analytic data sets and codebooks were examined. The audit included an observation of the demonstration of the script executions, file tree structure on the server, and model diagnostics. The audit also evaluated future analysis plans for the health data. No errors were noted, but recommendations were made for documenting model development, assumptions, QA/QC procedures and codes, and developing appropriate backup procedures for air quality data analysis performed on student desktops.

September–December 2017 (Final Remote Audit)

The final remote audit consisted of two parts: (a) review of final report for the project; and (b) audit of data processing steps. The audit of the final report focused on ensuring that it is well documented and easy to understand, and highlighted key findings and limitations of the study. This review also provided guidance on specific aspects of the data processing sequence that could be reviewed remotely. The audit of the data included reviewing the scripts for the data reduction, processing and analysis, model development, and visualization. This specific portion of the audit was restricted to the central portion of the research project that focused on detrending, generation of counterfactuals, and epidemiological analyses. Scripts and input data for the air quality detrending and model development component were sent to RTI for review. The scripts were executed at RTI, and the results were compared to some of the figures and tables in the report. For the epidemiological component, scripts were sent to RTI for review. However, due to restrictions on health data, a video conference was organized for the auditors to observe code execution and comparison of outputs obtained with SAS and R. No major quality-related issues were identified that would impact the findings. Certain typographical errors and mislabeling of figures in the report were discovered. Minor editorial corrections and recommendations for fixing the errors were made.

Written reports of each activity were provided to HEI. These quality assurance oversight audits demonstrated that the study was conducted by a well-coordinated, experienced team according to the study protocol and standard operating procedures. Interviews with study personnel revealed a consistently high concern for data quality. The final report, except as noted in the comments, appears to be an accurate representation of the study.

Linda Brown, M.P.H., Dr.P.H., Epidemiologist, Quality Assurance Auditor



Jeremy Aldworth, Ph.D., Statistician, Quality Assurance Auditor

Prakash Doraiswamy, Ph.D., Air Quality Specialist, Quality Assurance Auditor

Date: December 22, 2017

MATERIALS AVAILABLE ON THE HEI WEBSITE

Appendices A–E and Additional Materials 1 contain supplemental material not included in the printed report. They are available on the HEI website, *www.healtheffects.org/publications.*

Appendix A: Supplementary Figures and Tables

Appendix B: Estimation of Uncertainty in Empirical Counterfactuals

Appendix C: Background Demographic Information and Hospital Information

Appendix D: Evidence for a Potential Overestimation in Mobile Source NO_x Emissions

Appendix E: Linked Response of Aerosol Acidity and Ammonia to SO_2 and NO_x Emissions Reductions in the U.S.: A Focus on the Southeast

Additional Materials 1: White Paper: Attributing 20 Years of Electricity Generating Unit Emissions Reductions in Atlanta, GA to Specific Policy Actions

ABOUT THE AUTHORS

Armistead (Ted) G. Russell, the principal investigator on this project, is a professor of civil and environmental engineering at the Georgia Institute of Technology, where his research is aimed at better understanding the dynamics of air pollutants at urban and regional scales and assessing their impacts on health and the environment. He earned his M.S. and Ph.D. in mechanical engineering at the California Institute of Technology, and his bachelor's degree from Washington State University. Russell was a member of the U.S. EPA's Clean Air Science Advisory Committee and of the Board on Environmental Studies and Toxicology at the National Research Council, where he continues to serve on associated NRC committees. He previously served on the Health Effects Institute's Review Committee.

Paige Tolbert, an environmental epidemiologist, is lead on the Emory subaward. She is O. Wayne Rollins Professor and Chair of Environmental Health at the Rollins School of Public Health of Emory University, jointly appointed in the Department of Epidemiology. Tolbert has a bachelor's degree in biochemistry from Harvard College and an M.S. P.H. in environmental science and engineering and Ph.D. in epidemiology, both from the University of North Carolina–Chapel Hill. Much of her recent work has focused on health effects of ambient air pollution, and she and Russell codirect the Emory/Georgia Institute of Technology Clean Air Research Center, funded by the U.S. EPA. She recently completed service on the U.S. EPA's chartered Science Advisory Board.

Lucas R.F. Henneman contributed to this work as a PhD. student at the Georgia Institute of Technology, and is currently a postdoctoral fellow at the Harvard T. H. Chan School of Public Health. He earned his bachelor of science degree in environmental engineering from Johns Hopkins University in 2012 and his M.S. and Ph.D. in environmental engineering from the Georgia Institute of Tech in 2014 and 2017, respectively. In 2012, he was awarded a National Science Foundation Graduate Research Fellowship and the Georgia Tech President's Fellowship. The study described in this report serves as the basis for his Ph. D. thesis. Henneman assessed regulatory frameworks, modeled emissions, performed the empirical air quality analyses, and coordinated data transfers from Georgia Tech to Emory University.

Joseph Abrams contributed to this work as a Ph.D. student at Emory University and is currently an epidemiologist for the Centers for Disease Control and Prevention. He earned his bachelor's degree in biology from the University of Virginia in 2002 and his M.P.H. and Ph.D. in epidemiology from the Rollins School of Public Health at Emory. In 2014, he was awarded a traineeship on a National Institute for Occupational Safety and Health Training Grant. The study described in this report serves as the basis for aspects of his Ph.D. thesis. Abrams implemented the health impact models, generating estimated health outcome results and figures under a variety of model parameterizations and pollution-control scenarios.

Cong Liu was a postdoctoral research associate working with Ted Russell at the Georgia Institute of Technology's School of Civil and Environmental Engineering. He received a Ph.D. in civil engineering from the Department of Building Science of Tsinghua University, China, in 2014. He is an associate professor at Southeast University, China. His main research interests include regional air quality modeling, indoor–outdoor interaction, and gas-particle partitioning. For this study, Liu primarily ran CMAQ simulations and analyzed the mechanistic response of air quality to emissions and meteorology.

Mitchel Klein is an assistant research professor in the Department of Environmental Health at Emory University's Rollins School of Public Health, with a joint appointment in the Department of Epidemiology. He received his doctoral degree in epidemiology from Emory. His research focuses on methodological issues in epidemiology, particularly in air pollution epidemiology. For this study Klein was involved in the design and implementation of the health analyses.

James Mulholland is a professor of civil and environmental engineering at the Georgia Institute of Technology. He earned his Ph.D. in chemical engineering from the Massachusetts Institute of Technology. His research addresses the sources of air pollution and characterizing spatial and temporal distributions of ambient air pollutants. In this study, Mulholland has helped guide the work to estimate impacts of regulatory actions on air quality.

Stefanie Ebelt Sarnat is an associate professor of environmental health at the Rollins School of Public Health of Emory University. She received a Ph.D. in environmental health from the Harvard T.H. Chan School of Public Health in 2005. Her research focuses primarily on assessing health impacts of air pollution, meteorological conditions, and weather extremes using population- and panel-based approaches. In this study, Sarnat participated in planning the data analysis for the health impact assessment and its interpretation.

Yongtao Hu is a senior research scientist at the School of Civil and Environmental Engineering, Georgia Institute of Technology. He earned his Ph.D. in environmental sciences at Peking University, China. His research focuses on meteorology and air quality modeling, air quality forecasting and dynamic management, sensitivity analysis, and source apportionment using air quality models and hybrid approaches. In this project, he contributed to emissions and air quality modeling.

Howard H. Chang is an assistant professor in the Department of Biostatistics and Bioinformatics at Emory University. He received a Ph.D. in biostatistics from Johns Hopkins University. His current research interests include spatial epidemiology, environmental statistics, and Bayesian methods. For this study, Chang assisted in statistical analyses and results interpretation.

Talat Odman is a principal research engineer in the School of Civil and Environmental Engineering at the Georgia Institute of Technology. He received a Ph.D. in mechanical engineering from Carnegie Mellon University. His current research interests include ozone and secondary aerosol formation, air quality impacts of fires, air quality forecasting, and dynamic air quality management. Odman participated in air quality modeling portions of this work.

Matthew J. Strickland, Ph.D., M.P.H., is associate professor of epidemiology in the School of Community Health Sciences, University of Nevada, Reno. He received a Ph.D. in epidemiology from Emory University in 2007. He is an environmental epidemiologist interested in better understanding the health effects of ambient air pollution mixtures. For this study, Strickland helped with the design and interpretation of the epidemiological analyses.

Huizhong Shen is a postdoctoral fellow in the Georgia Institute of Technology School of Civil and Environmental Engineering. He received a Ph.D. in environmental geography from Peking University. His current research focuses on assessing the impacts of land use policy on the nitrogen cycle and on air quality in the United States. For this study, Shen led the analysis of evidence for a potential bias in the mobile-source NO_x emissions in the southeastern United States using multiple approaches.

Abiola Lawal is a graduate student at the Georgia Institute of Technology and was involved in empirical analysis of pH trends in the Southeast. She has a bachelor's degree in chemical engineering from the University of Minnesota and a master's degree in chemical engineering from Auburn University.

OTHER PUBLICATIONS RESULTING FROM THIS RESEARCH

Henneman LRF, Chang HH, Liao KJ, Lavoué D, Mulholland JA, Russell AG. 2017. Accountability assessment of regulatory impacts on ozone and $PM_{2.5}$ concentrations using statistical and deterministic pollutant sensitivities. Air Qual Atmos Health10:695–711. Available: *https://link.springer.com/article/10.1007/s11869-017-0463-2.*

Henneman LRF, Liu C, Hu Y, Mulholland JA, Russell AG. 2017. Air quality modeling for accountability research: Operational, dynamic, and diagnostic evaluation. Atmos Environ 166:551–565. Available: *www.sciencedirect.com/science/article/pii/S1352231017304983.*

Henneman LRF, Shen H, Liu C, Hu Y, Mulholland JA, Russell AG. 2017. Responses in ozone and its production efficiency attributable to recent and future emissions changes in the eastern United States. Environ Sci Technol 51(23): 13797–13805. doi: 10.1021/acs.est.7b04109

Henneman LRF, Liu C, Chang H, Lavoué D, Mulholland JA, Russell AG. 2016. Estimating the impact of air pollution controls on ambient concentrations. In: Air Pollution Modeling and its Application XXIV. pp 141–146. Switzerland:Springer International *www.springer.com/us/book/9783319244761.*

Henneman LRF, Liu C, Mulholland JA, Russell AG. 2016. Evaluating the effectiveness of air quality regulations: a review of accountability studies and frameworks. J Air Waste Manage Assoc 67:144–172; *www.tandfonline.com/doi/abs/10.1080/10962247.2016.1242518.*

Henneman LRF, Holmes HA, Mulholland JA, Russell AG. 2015. Meteorological detrending of primary and secondary pollutant concentrations: method application and evaluation using long-term (2000–2012) data in Atlanta. Atmos Environ 119:201–210. doi:10.1016/j.atmosenv.2015.08.007. *http://linkinghub.elsevier.com/retrieve/pii/S1352231015302521.*

COMMENTARY

Review Committee

Research Report 195, *Impacts of Regulations on Air Quality and Emergency Department Visits in the Atlanta Metropolitan Area, 1999–2013*, A.G. Russell et al.

INTRODUCTION

Accountability research evaluates whether regulatory and other actions aimed at improving air quality have resulted in actual decreases in air pollutant concentrations and improvements in public health. Frequently, researchers follow the impacts along a chain of accountability that includes the regulation and its implementation, and effects on emissions, air quality, exposure, and public health (Boogaard et al. 2017; Health Effects Institute 2010; van Erp et al. 2008). In establishing whether a particular regulatory action has been successful in providing the intended improvements, it is very useful to establish mechanistic links between the steps in the chain: Has the action improved air quality? If so, has it reduced people's exposure to air pollutants? And finally, has it in turn led to improved public health?

Because of the various time scales and geographical areas covered by regulations, and because the right data are often not available, along with many methodological issues, accountability research is challenging. The past decades have seen an increasing number of accountability studies. For practical reasons, many early studies focused on the impacts of local actions taking place over relatively short time frames, such as measures to reduce traffic in a city (Dockery et al. 2013; Lee et al. 2007; Peel et al. 2010; Zhang et al. 2011). Several studies to date have focused on the impact of long-term policies, such as the impact of the U.S. Clean Air Act Amendments (Gilliland et al. 2017; Morgenstern et al. 2012; Zigler et al. 2016). One of the challenges in evaluating longer-term regulatory actions is that studies may be confounded by other simultaneous changes (for example, improvements in access to healthcare) that may also affect air quality or health. An accountability research approach that addresses this challenge is to compare the observed changes in air quality and health after implementation of the regulation with projected *what-if scenarios* — also called *counterfactual scenarios* — where researchers estimate what the air quality and health outcomes would have been without the intervention.

The Health Effects Institute has a long history in accountability research. HEI has contributed both to the development of the conceptual framework for accountability research and to the funding for a number of studies designed to assess the health outcomes of actions to improve air quality (HEI Accountability Working Group 2003). This history is described more fully in the Preface to this report. Of the first set of nine studies that HEI funded, a majority evaluated actions that were at a local scale or implemented relatively rapidly. For example, researchers studied the effects of banning the sale of coal for heating in Dublin and other Irish cities (Dockery et al. 2013) or of reducing emissions from traffic or local and regional sources during a unique event, such as the Olympic Games (Peel et al. 2010; Rich et al. 2012; Zhang et al. 2013). HEI also funded some studies that evaluated longer-term, national changes, such as air quality improvements after the reunification of Germany (Peters et al. 2009) and reductions in emissions from power plants under the Clean Air Act (Morgenstern et al. 2012). Summaries of these studies have been published that discuss the advances made, and challenges encountered, in conducting such research (Boogaard et al. 2017; Health Effects Institute 2010; HEI Accountability Working Group 2003; Henneman et al. 2016; Rich 2017; van Erp et al. 2008).

After assessing the results from this first wave of nine studies, HEI issued Request for Applications (RFA*) 11-1, "Assessing the Health Outcomes of Air Quality Actions," in 2011. The goals of this RFA were to fund research to (1) evaluate regulatory and other actions at the national or regional level implemented over multiple years; (2) evaluate complex sets of actions targeted at improving air quality in large urban areas, including those in the vicinity

Dr. Armistead (Ted) G. Russell's 3-year study, "Impacts of Emission Changes on Air Quality and Acute Health Effects in the Southeast, 1993-2012," began in January 2013. Total expenditures were $671,913. The draft Investigators' Report from Russell and colleagues was received for review in October 2016. A revised report, received in May 2017, was accepted for publication in June 2017. During the review process, the HEI Review Committee and the investigators had the opportunity to exchange comments and to clarify issues in both the Investigators' Report and the Review Committee's Commentary.

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

of major ports; or (3) develop methods to support such health outcomes research.

In response to RFA 11-1, Dr. Russell and colleagues proposed to quantify how emissions reductions programs and meteorological variations affect air quality and public health, using a comprehensive air quality model and new statistical models. The overall goals were to (1) measure and model changes in emissions and air quality in the southeastern United States over a 15-year period (1999–2013), focusing on six regulatory programs to control air pollution emissions from power-generating utilities and from light- and heavy-duty vehicles, taking into account meteorological trends; and (2) assess the impacts of the regulatory programs on acute emergency department (ED) visits in the Atlanta area.

Dr. Russell and colleagues proposed a step-wise approach, following the stages outlined in the HEI chain of accountability: First, they would estimate changes in emissions that could be attributed to each of the different air quality regulations. Second, they would estimate the changes in air quality associated with each regulation (or combinations thereof). Third, they would estimate several counterfactual scenarios — that is, emissions and air quality over the same period but in scenarios in which the regulations had not been implemented. Subsequently, in collaboration with Dr. Paige Tolbert and colleagues at Emory University, they would estimate the potential health benefits associated with the changes in concentrations due to these six regulatory programs under the various emissions scenarios (actual and projected).

The HEI Research Committee recommended the proposal by Russell and colleagues for funding because they thought it was a strong study design, in particular regarding the proposed analyses of emissions and air quality data using a variety of state-of-the-art modeling approaches and the scenario approach. They also liked the proposed detrending methods to remove the influence of short- and long-term variation in meteorology on estimates of air quality associated with the regulations, an issue in prior analyses in the Atlanta area (Friedman et al. 2001; Peel et al. 2010). They also noted that the investigators had extensive experience with emissions and air quality modeling in the region and had access to a large database of health outcomes covering a twenty-year period that had been well studied (see section on Previous Studies in Atlanta below). The project started in January 2013.

This Commentary provides the HEI Review Committee's 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 Investigators' Report into scientific and regulatory context.

REGULATORY AND SCIENTIFIC BACKGROUND

REGULATIONS AND OTHER FACTORS AFFECTING AIR POLLUTANT EMISSIONS

Air pollution in the United States is regulated by the Clean Air Act, which sets allowable concentrations, known as National Ambient Air Quality Standards (NAAQS) for six major pollutants known as the criteria pollutants (particulate matter [PM], ozone [O_3], carbon monoxide [CO], oxides of sulfur [SO_x], oxides of nitrogen [NO_x], and lead). Emissions of air pollutants are controlled under federal rules and state regulations adopted in response to federal requirements so that these NAAQS can be attained. Georgia's state-specific regulations to reduce air pollutant emissions are described in its federally mandated State Implementation Plans (SIPs). At this time, Atlanta is out of compliance with the NAAQS for O_3; it is in compliance with the NAAQS for PM $\leq$2.5 µm in aerodynamic diameter ($PM_{2.5}$) and has submitted a maintenance plan through year 2024. Federal and state regulatory programs target emissions from large stationary sources like power plants or from mobile sources like cars and trucks. The regulations seek to reduce emissions of air pollutants by mandating changes in fuel composition, requiring installation of control technologies, or specifying total or facility-specific emission rates.

Commentary Table 1 briefly describes key rules and regulations affecting air quality in Atlanta by limiting emissions from electricity-generating units (EGUs) and mobile sources, and the implementation phases of each rule. However, regulatory programs are not the only consideration for operators of power-generating utilities or manufacturers designing a new vehicle. Factors such as improved performance and costs of operation or construction may also be taken into account. For example, fuel changes may be made to reduce operational costs of EGUs, and fuel costs can influence consumers to purchase more efficient motor vehicles. As a result, actual emissions are difficult to link to specific regulations.

ASSESSING THE IMPACT OF REGULATIONS

The United States Environmental Protection Agency (U.S. EPA) is required to assess the costs and benefits of the Clean Air Act as a whole, as well as of individual regulations. To date, the U.S. EPA has released one retrospective (1997) and two prospective (1999b, 2011) studies of the overall benefits — including health — of the Clean Air Act relative to its costs. The U.S. EPA has also released Regulatory Impact Assessments (RIAs) estimating the expected costs and benefits of a number of individual rules and their alternatives proposed under the Clean Air Act (e.g., U.S. EPA 1992, 1998, 1999a, 2000, 2005). In general, these RIAs compare expected future scenarios with and without regulation (or different versions of the regulation) to assess whether the proposed rules are likely to be cost-effective and meet their stated goals. For the Acid Rain Program, the RIA focused nearly entirely on the costs of implementation (U.S. EPA 1992). Later RIAs added detailed sections describing the anticipated effects of the rules on emissions, air quality, and health (e.g., U.S. EPA 1998, 1999a, 2000, 2005). The U.S. EPA uses estimates of avoided mortality, hospital admissions, or other outcomes, and economic assumptions about the value of those avoided outcomes, to characterize the monetary benefits of improved health from the regulation or intervention. Thus, predictions of the benefits of all of the federal rules and programs listed in Commentary Table 1 are available.

To what extent have the rules achieved their expected goals in reducing emissions, air pollution concentrations, and adverse health impacts? These are questions that accountability research attempts to answer and that Russell and colleagues proposed to address for the various regulations affecting the Atlanta area. They used a scenario approach similar to that used for the RIAs but applied it retrospectively and based their models on actual, observed changes after the regulations were implemented. The scenario approach has been used in a few previous accountability studies (e.g., Morgenstern et al. 2012; Tonne et al. 2008). Most accountability studies of national-scale regulations have used approaches other than the scenario approach, such as time-series analyses corrected as much as possible for other long-term trends (e.g., Dockery et al. 2013; Peters et al. 2009); cohort comparisons over a period that included reductions in air pollutant concentrations (Gilliland et al. 2017); or causal modeling (Zigler et al. 2016). The counterfactual approach used in the current study differs from approaches typically used to evaluate the impact of regulations because the current study did not rely on the use of a control area (i.e., where regulations had not been implemented) or a control period (i.e., before regulations were implemented) as the basis for evaluation. In addition, the current study assesses both the impact of changes in emissions on air quality and the impact of changes in air quality on health.

PREVIOUS STUDIES IN ATLANTA

The study by Russell and colleagues built on previous work by their group in the southeastern United States, focusing on emissions and air quality modeling, including the influence of meteorology. Levels of all of the pollutants considered in the current study have been measured at the U.S. EPA Atlanta Supersite Project site near downtown Atlanta (hereafter referred to as the "Jefferson Street monitoring site") since at least 1998. The Jefferson Street monitoring site is also part of the Southeastern Aerosol Research Characterization Study and Aerosol Research Inhalation Epidemiology Study networks. This site was operated by Atmospheric Research and Analysis, Inc., with support from the Electric Power Research Institute and southeastern utilities (Solomon et al. 2003). Data quality at the sites is verified with independent audits by outside consultants.

In addition to being of generally high quality, the Jefferson Street monitoring site has been used in extensive prior research, including comparisons to other nearby stations (Hansen et al. 2012; Solomon et al. 2003) and has also been used in previous work attempting to link decreasing air pollutant levels with emissions reductions (Blanchard and Hidy 2005; Blanchard et al. 2010). Statistical air quality models in the current study built on regional and theoretical studies that separated the long-term and short-term impacts of meteorology on air quality (Eskridge et al. 1997; Flaum et al. 1996; Kuebler et al. 2001; Rao et al. 1997). The investigators previously used the Community Multiscale Air Quality modeling system (CMAQ), a model that is used to externally evaluate the statistical models in the current study in order to test the sensitivity of O_3 and $PM_{2.5}$ levels in the southeastern United States to NO_x and volatile organic compound (VOC) emissions (Cohan et al. 2005; Liao et al. 2008; Napelenok et al. 2006).

The health analyses conducted for the current study built on prior epidemiological analyses performed by the investigators as part of the Study of Particles and Health in Atlanta (SOPHIA) (Metzger et al. 2004; Peel et al. 2005; Tolbert et al. 2000, 2007). The goal of SOPHIA was to study the relationship between daily air pollutant concentrations and daily ED visits for cardiovascular and respiratory diseases. In previous publications that cover the period 1993–2000, the investigators reported positive associations of cardiovascular ED visits with ambient levels of nitrogen dioxide (NO_2), CO, $PM_{2.5}$, and the organic carbon (OC) and elemental carbon (EC) content of $PM_{2.5}$ (Metzger et al. 2004). They also reported positive associations of respiratory ED visits with the levels of O_3, NO_2, CO, and sulfur dioxide (SO_2) (Darrow et al. 2011; Peel et al. 2005). In a follow-up study that evaluated a longer period and used multipollutant models, they found that cardiovascular disease ED visits during 1993–2004 remained associated with the same pollutants as observed previously, using single-pollutant models, but that CO was the strongest predictor in all multipollutant models tested (Tolbert et al. 2007). Similarly, multipollutant modeling showed that respiratory disease was associated with O_3 and PM_{10}, with O_3 as the stronger predictor (Tolbert et al. 2007).

Commentary Table 1. Key Regulations to Control Emissions from Power Plants and Motor Vehicles in the Atlanta Region[a]

Name[b]	Brief Description	Implementation Phases
EGU Programs		
Acid Rain Program		
Title IV of the 1990 Amendments to the Clean Air Act Acid Rain Program	Cap-and-trade program to reduce SO_2 and NO_x emissions and reduce the acidity of rain and natural waters	• Phase I (largest power plants): 1995 for SO_2 and 1996 for NO_x • Phase II (all other plants): 2000
Emissions of Nitrogen Oxides from Major Sources ($GRAQC_{yy}$)	Required major sources of nitrogen oxides in the Atlanta and Macon nonattainment areas for the 1-hr O_3 standard to install "reasonably available control technology" for control of NO_2	• Older sources in Atlanta exceeding 50 tons-per-year NO_2 • Older sources in Macon emitting at least 25 tons-per-year NO_2 by May 1, 2007 • Older sources in Atlanta emitting at least 25 tons-per-year NO_2 in the 13-county area by May 1, 2007 • New sources in Atlanta by April 1, 2004 • New sources in Barrow county (Atlanta) by March 1, 2009
Nitrogen Budget Program		
NO_x Budget Trading Program and Associated State Implementation Plan Call	Optional cap-and-trade program to reduce NO_x emissions contributing to nonattainment of ozone standards in downwind states	• All 20 states in the SIP Call met their requirements by 2003
NO_x Emissions from Electric Utility Steam Generating Units ($GRAQC_{jjj}$)	NO_x emissions limits on EGU sources in the 13-county Atlanta nonattainment area during the summer (1 May–30 September)	• Compliance by some units required by summer 1999 • More plants added each summer between 2000 and 2007
Multipollutant Control of Interstate Air Pollutant Transport		
Clean Air Interstate Rule	Set up SO_2 and NO_x trading programs and required NO_x ozone-season controls to reduce pollutant (e.g., $PM_{2.5}$ and ozone) transport across state borders	• Implemented and kept in place by court decision in 2008
Multipollutant Control for Electricity Utility Steam Generating Units ($GRAQC_{sss}$) (also known as the "Georgia Multipollutant Control Rule")	Year-round controls: SCR to reduce NO_x and FGD to reduce SO_2 from coal-fired power plants	• Dates of compliance for specific named plants ranged from December 31, 2008, to January 1, 2018

Table continues next page

[a] EGU = electricity-generating unit; FGD = flue gas desulfurization, an SO_2 emissions control device; GRAQC = Georgia Rules for Air Quality Control, with subscripts referring to sections of the Rules; NMHC = nonmethane hydrocarbon; NO_2 = nitrogen dioxide; NO_x = oxides of nitrogen; PM = particulate matter; $PM_{2.5}$ = PM ≤2.5 µm in aerodynamic diameter; RVP = Reid vapor pressure; SCR = selective catalytic reduction; SO_2 = sulfur dioxide; VOC = volatile organic hydrocarbon.

[b] Regulations are grouped in five subsets that were evaluated by the investigators. Modified from IR Table 1 and IR Figure 2.

Commentary Table 1 (*Continued*). Key Regulations to Control Emissions from Power Plants and Motor Vehicles in the Atlanta Region[a]

Name[b]	Brief Description	Implementation Phases
Mobile Source Programs		
Inspection and Maintenance		
Inspection and Maintenance (Georgia)	Enhanced emissions and safety testing on gasoline-powered cars and light trucks registered in 13 counties surrounding Atlanta	• 1996
Light Duty Vehicles and Gasoline		
Tier 2 Light and Medium Duty Vehicle Program	Updated NO_x and nonmethane hydrocarbon engine emissions standards for all passenger cars, light trucks, and medium-duty passenger vehicles	• Phased in between 2004 and 2009
Gasoline Marketing Rule ($GRAQC_{bbb}$)	Limited RVP to reduce evaporative VOC emissions and reduced sulfur content in gasoline sold in and around Atlanta to reduce sulfur dioxide and sulfate emissions and increase SCR efficiency	• Limited RVP to 7.0 psi from June 1–September 15, starting in 1999 • Reduced seasonal average gasoline sulfur content limits (150 ppm in 1999, 90 ppm in 2003, and 30 ppm in 2004) • Overridden by Tier 2 limit in 2006
Diesel		
Heavy Duty Diesel Rule	Limited sulfur content in fuels to increase SCR efficiency and enforced reduced NO_x and NMHC emissions standards on new or rebuilt highway heavy-duty engines to reduce O_3 levels by reducing O_3 precursor emissions, also required diesel particulate filters	• Fuel sulfur limits of 15 ppm for most diesel fuel sold by major refiners for use in highway vehicles by June 2006 • Diesel fuel sold in Georgia had less than 500 ppm sulfur by 2008 and 15 ppm sulfur by 2012 • Intermediate NO_x and NMHC emissions standards enforced beginning on model year 2004 vehicles • Stronger PM, NO_x and NMHC standards phased in 2007 to 2010

[a] EGU = electricity-generating unit; FGD = flue gas desulfurization, an SO_2 emissions control device; GRAQC = Georgia Rules for Air Quality Control, with subscripts referring to sections of the Rules; NMHC = nonmethane hydrocarbon; NO_2 = nitrogen dioxide; NO_x = oxides of nitrogen; PM = particulate matter; $PM_{2.5}$ = PM ≤2.5 µm in aerodynamic diameter; RVP = Reid vapor pressure; SCR = selective catalytic reduction; SO_2 = sulfur dioxide; VOC = volatile organic hydrocarbon.

[b] Regulations are grouped in five subsets that were evaluated by the investigators. Modified from IR Table 1 and IR Figure 2.

Other work by researchers at the Georgia Institute of Technology and at Emory University has assessed the sensitivity of the air pollution exposure assignments and epidemiological results to various key assumptions incorporated into the analyses for the current study. For example, several studies examined the effect of using a single monitoring station on exposure assignments and measurement error when people in fact move around over the course of the day (Darrow et al. 2011; Goldman et al. 2011; Sarnat et al. 2010, 2013); others examined the effect of spatial confounding on health outcomes (Flanders et al. 2011, 2017; Strickland et al. 2015).

The Atlanta region also featured two previous accountability studies that showed the importance of controlling for meteorology in assessments of the impacts of interventions on air quality. An initial, often cited study by Friedman and colleagues (2001) evaluated the impact of measures to reduce traffic during the 1996 Summer Olympic Games in Atlanta. They reported reductions in the number of asthma and nonasthma pediatric acute care events in and around Atlanta that were associated with a modest traffic reduction and a small decrease in O_3 concentrations. A follow-up study by Peel and colleagues (2010) looked specifically at the influence of regional meteorology on air quality and included analyses of the same summer period in the previous and following years for comparison. They showed that the observed decreases in O_3 concentrations were regional in nature and were likely due to favorable weather in the region. Although they confirmed some health improvements in association with the lower O_3 concentrations during the Olympic Games period, it was not possible to definitively attribute the changes in O_3 to the traffic interventions (Peel et al. 2010). This work emphasized the importance of accounting for changes in both the period of the intervention and the population affected by it.

SUMMARY OF THE STUDY

OBJECTIVES

The objectives of the current study were to:

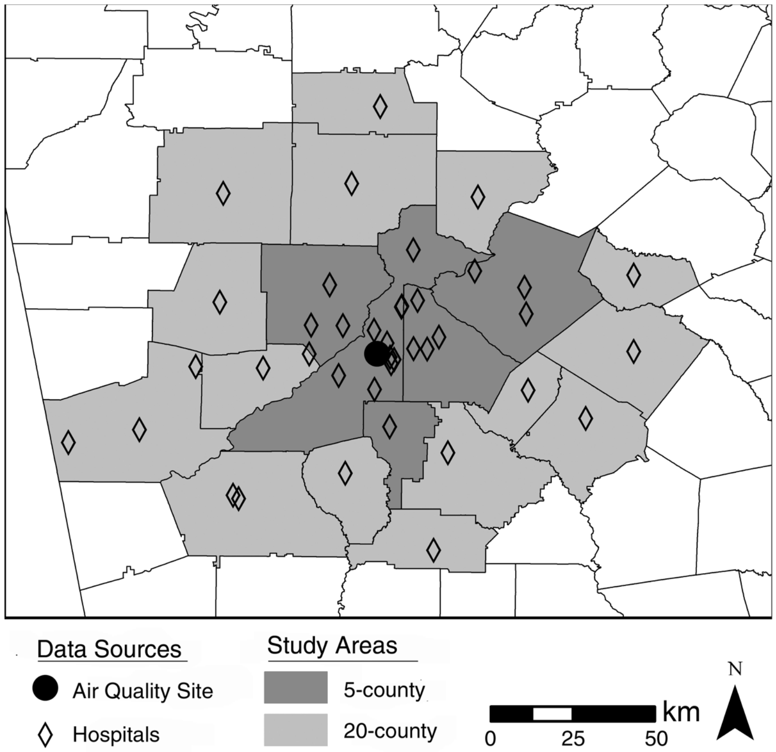
1. Quantitatively assess the impacts of controls on NO_x and SO_x emissions from EGUs on air quality in the southeastern United States.
2. Quantitatively assess the impacts of controls on NO_x, VOC, and CO emissions from light-duty vehicles and on NO_x and PM emissions from heavy-duty vehicles on air quality in the southeastern United States.
3. Assess the impact of meteorological trends on air quality in the southeastern United States.
4. Assess the impacts of the regulatory programs on acute ED visits, with a focus on cardiorespiratory health outcomes.
5. Evaluate the effects of methodological choices used in estimating the impact of pollution-control policies.
6. Conduct uncertainty analyses capturing potential error in parameter estimation at multiple analytical stages to construct comprehensive confidence interval estimates.

The regions of interest in the southeastern United States were the 20-county Atlanta nonattainment area for PM and O_3 with a focus on the 5-county metropolitan Atlanta area (see Commentary Figure 1). The investigators proposed two approaches to estimating impacts on air quality: (1) a chemical transport model linking emissions, meteorology, and air quality over the southeastern United States during two years near the beginning (2001) and end (2011) of the study period and (2) an empirical model that was used to statistically relate air quality to emissions and meteorology for the full study period and was limited to the metropolitan Atlanta area. The health analysis was done using the empirical approach, and the chemical transport results were used to test the sensitivity of the air quality results to meteorology. Initially, the investigators proposed to analyze hospital admissions (in addition to ED visits), as well as impacts on susceptible and vulnerable populations. During the course of the study, the investigators decided not to pursue those additional analyses at the recommendation of the HEI Research Committee. Although potentially informative, they were considered to be too ambitious to complete in the available time.

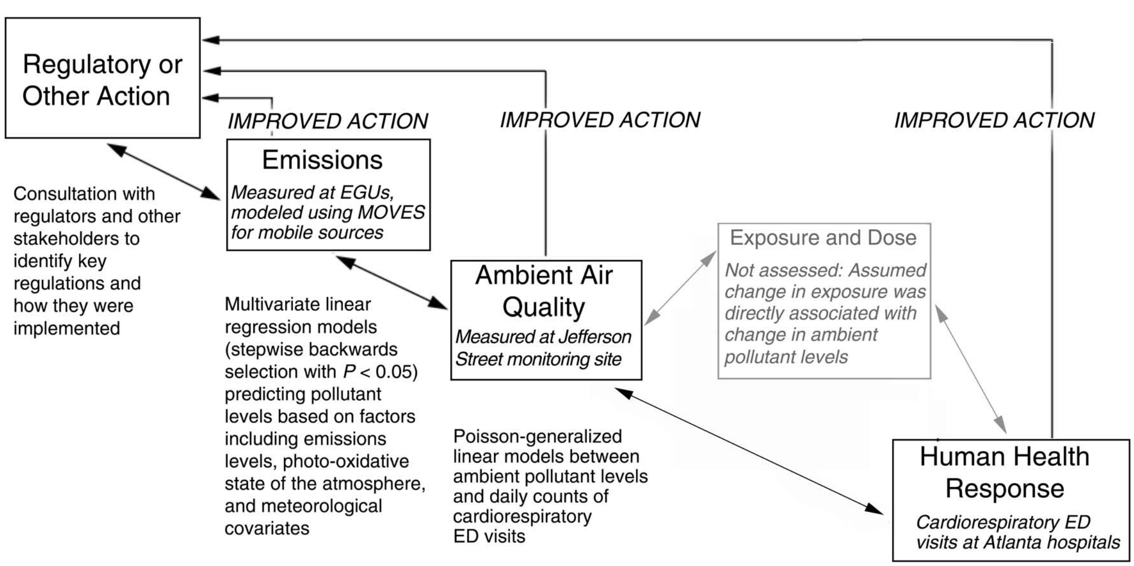
STUDY DESIGN AND METHODS

Study Approach

The study was conducted within the framework of the chain of accountability (Commentary Figure 2). The investigators sequentially addressed four steps in the chain of accountability, where the results of each step served as inputs for the next step: (1) identification of regulatory actions; (2) assessment of their impacts on emissions; (3) assessment of the changes in emissions on air quality; and (4) ultimately assessment of their impacts on health. The fundamental approach taken in this study differed from that of many previous accountability studies in that it did not rely on use of a control area (i.e., where regulations had not been implemented) or control period (i.e., before regulations were implemented) as the basis for evaluation. The investigators anticipated that it would be difficult either to define an appropriate control region because the study area was too large (regulatory programs were implemented both regionally and nationally) or to define an appropriate control period because the study period was too long (programs were implemented over a nearly 15-year period). Instead, they therefore compared observed changes in emissions, air quality, and health following the implementation of regulations to what *would have happened* under different *counterfactual* scenarios, in which the various regulatory programs affecting electricity-generating units (EGUs) and mobile source emissions were assumed not to



Commentary Figure 1. Map of the 5-county Atlanta (dark grey) and 20-county (light and dark grey) study areas in Atlanta. The data sources were the Jefferson Street air quality monitoring site (filled circle, center of map) and 42 Atlanta hospitals (open diamonds). The study included all emergency departments visited by residents of the study area at nonfederal acute care hospitals. The number of participating hospitals varied over the study period as hospitals opened, closed, or merged.



Commentary Figure 2. Summary of the study approach within the chain of accountability. Major methods connecting the steps on the chain of accountability are listed next to the connecting arrows, and major data sources are briefly described in italics inside the boxes. MOVES = U.S. EPA MOtor Vehicle Emissions Simulator.

have been implemented. The difference between the counterfactual estimates and the actual air quality and health outcomes with all regulations in place was interpreted as the effect of the regulatory program.

Selection of Regulations

The investigators included major regulations that they thought were likely to affect air pollutant emissions or air quality in the southeastern United States, including Atlanta, Georgia (see Commentary Table 1). They evaluated three national program sets to reduce emissions from power plants (EGUs): the Acid Rain Program, the NO_x Budget Trading Program and the associated State Implementation Plan Call, and the Clean Air Interstate Rule. In addition, they evaluated three programs associated with the Clean Air Act Amendments (CAAA) and Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Requirement: the Georgia Gasoline Marketing Rule, Inspection and Maintenance, and the 2007 Heavy Duty Highway Rule. Potential differences among the actual implementation of the regulation, the reported changes, and factors other than the regulation were not explored. However, the investigators did consult with experts in the power-generating industry and the Georgia Environmental Protection Division throughout the process to match their scenarios as closely as possible to reality and to gather feedback on the interpretation of their results.

Analysis of Emissions and Air Quality

Emissions For each of the major regulatory programs, Russell and colleagues first used emissions inventories to compare emissions before each regulation was implemented to those at the end of the study period (in 2013). Then they estimated *actual* (i.e., with all regulations in place) and *counterfactual* (i.e., without regulation) time series of daily emissions for the period 1999–2013. They used measurements where available, and models where measurements were not available, to compare actual and counterfactual time-series scenarios of emissions from 1999 through 2013. Commentary Table 2 provides an overview of the main and sensitivity analyses conducted.

For EGUs, they then compared modeled counterfactual time series of NO_x and SO_2 emissions to actual emissions measured at the EGUs. Daily counterfactual EGU emissions were constructed from models that assumed that each actual megawatt-hour of energy generated resulted in the same emissions as those measured prior to the promulgation of the regulations.

They also compared time series of modeled mobile source emissions of NO_x, VOCs, and $PM_{2.5}$ between scenarios with and without regulations. For the mobile sources, both actual and counterfactual emissions were estimated using the MOtor Vehicle Emissions Simulator (MOVES, version 2010b) (U.S. EPA 2012) because no measurements of mobile source emissions were available. To construct the counterfactuals for mobile source emissions, the investigators assumed that emissions from each vehicle depended on the age of the vehicle and fuel used, but that changes in control technology had not occurred.

Air Quality The investigators' assessment of effects of regulations on air quality was based on air quality measurements at the Jefferson Street monitoring site (see earlier section, "Previous Studies in Atlanta") and the counterfactual emissions estimates. Before doing any air quality modeling, the investigators compared measured levels of gaseous (O_3, NO_2, SO_2, and CO) and particulate ($PM_{2.5}$, sulfate, nitrate, ammonium, OC, EC, and acidity of $PM_{2.5}$) pollutants in the year 1999 to those in the year 2013.

Next, counterfactual air pollutant levels were estimated using statistical models. The investigators started with daily time series of measured air pollutant levels from the Jefferson Street monitoring site. They removed meteorological trends and built statistical models of the relationships of air pollutant levels to emissions and day-to-day variation in meteorology. To construct the counterfactual air pollutant levels, the investigators replaced actual emissions (measured for EGUs and modeled for mobile sources) in the statistical models with the counterfactual (modeled) emission estimates. The investigators compared actual measured air pollutant levels (i.e., those used to build the statistical models of air quality and health relationships) to the counterfactual air pollutant level time series to estimate the effect of regulations on air quality and used these two time series as inputs in the health analyses.

Sensitivity Analyses Earlier accountability studies in the Atlanta area had shown that the relationships between policy changes, air quality, and health could be temporally confounded by regional meteorology (see Peel et al. [2010]). Therefore, to assess the impact of meteorological trends relative to changes in emissions on air quality, the researchers used two approaches: an empirical approach removing the meteorological influence from their regression models and a simulation approach comparing chemical transport model results under different assumptions about emissions and meteorology. For example, they conducted sensitivity analyses using a physics-based regional air quality model (CMAQ) to assess whether changes in emissions or meteorology were more likely to account for differences in the air

Commentary Table 2. Main Results of Temporal Trend and Sensitivity Analyses for Emissions and Air Quality

Topic	Analysis	Main Result
Temporal Trends		
Emissions		
	Compared emissions inventory of EGU and mobile source emissions in the Southeast and Atlanta area between the base year when each regulation was promulgated and the year 2013.	EGU and mobile source emissions decreased substantially between the late 1990s and 2013, with magnitude of change depending on the area and pollutant (see IR Table 2).
	Compared time series of EGU emissions from 1999 to 2013 for scenarios with (actual measurements) and without (modeled counterfactual) programs affecting EGU emissions.	EGU emissions of NO_x, and SO_2 varied seasonally but gradually decreased from year to year over the study period; the emissions in 2013 were about an order of magnitude less than those in 1999 (see IR Figure 4).
	Compared time series of modeled mobile source emissions from 1999 to 2013 for scenarios with and without programs affecting mobile source fuels and controls.	Mobile emissions of NO_x, VOC, and $PM_{2.5}$ gradually decreased over the period and the emissions in 2013 were less than half those in 1999 (see IR Figure 5).
Air Quality		
	Compared measured gaseous and particulate matter concentrations at the Jefferson Street monitoring site in Atlanta in 1999 and 2013.	Large decreases in air pollutant levels were observed for most pollutants measured (see IR Table 5).
	Compared actual (measured) and counterfactual (modeled) time series of air quality at the Jefferson Street monitoring site in Atlanta from 1999–2013.	Contributions of EGUs and mobile sources to air pollution at the Jefferson Street monitoring site decreased from 1999–2013, with larger differences for EGUs than mobile sources and substantial seasonal variability.
	Used multivariate regression models and CMAQ to determine whether acidity of $PM_{2.5}$, a key factor in particulate matter formation and dynamics with potential health implications, changed over time.	Acidity of $PM_{2.5}$ remained essentially the same, although the amount of acidic particulate matter decreased (see IR Figure 12, IR Appendix E).
Sensitivity Analyses of Emissions and Air Quality		
Bias in Estimates of Mobile Source Emissions		
	Investigate the potential impacts of bias in mobile source NO_x emissions estimates from the MOVES model using trend analysis, air quality modeling, satellite data, and a ratio-of-ratios analysis.	MOVES may have overestimated NO_x emissions; needs more study (see IR Appendix D).
Meteorological Impact on Air Quality Trends		
	Compared measured and meteorologically detrended air quality at the Jefferson Street monitoring site.	Downward trend in air pollutant levels remained after removing meteorological influence (see IR Figure 6) (Henneman et al. 2015).
	Evaluated the relative influence of changes in emissions and meteorology on differences between air quality in 2001 and 2011 in the southeastern United States using CMAQ.	Compared with emissions, long-term meteorological trends played a negligible role in air quality changes between 2001 and 2011 (see IR Figure 8: change in O_3, and IR Figure 9: change in $PM_{2.5}$) (Henneman et al. 2017b).

quality results near the start (year 2001) and near the end (year 2011) of the study period. The CMAQ model output also allowed them to explore whether acidity, an important property of PM in the air, had changed over time. The investigators also assessed whether bias in the emissions models could have influenced the main results.

Analyses of Health Outcomes

Emergency Department Visits The investigators sought to estimate the numbers of ED visits that might have been prevented due to the implementation of the various regulatory policies affecting EGUs and mobile sources in the 5-county Atlanta metropolitan area from 1999–2013. Their basic approach was to develop multipollutant models to describe the relationships between observed daily air pollutant concentrations and ED visits during the study period, and to use those models to estimate ED visits under assumption of the counterfactual air quality time series.

Health outcomes of interest were ED visits for respiratory disease, cardiovascular disease, asthma, and congestive heart failure. This selection was based on associations observed in previous studies using the same Atlanta ED data. Data on ED visits were obtained from 42 hospitals in the 20-county nonattainment area of Atlanta over the 15-year study period from 1999–2013. Annual ED visits increased by 74.2% over this period, from 710,414 in 1999 to 1,237,541 in 2013. Data sets for the study period included about 1.6 million ED visits for respiratory disease and 0.4 million ED visits for cardiovascular disease. During the same period, the population of Atlanta increased by about 24%. On average, more than 97% of the included hospitals were reporting data on any given day; the data set included which hospitals were not reporting data on each day. Variables for each patient included date of admission, primary International Classification of Diseases 9th Revision (ICD-9) diagnostic code, date of birth, sex, race, and 5-digit residential ZIP code.

To estimate associations of daily exposure to multiple pollutants with ED visits, the investigators used time-stratified Poisson generalized linear regression models (Bhaskaran et al. 2013) that account for overdispersion (meaning the presence of more variability in the measurements than would be expected based on the Poisson model). Lags (i.e., time between exposure and the outcome) were chosen a priori based on previous research (Sarnat et al. 2013). The investigators used 3-day moving averages (lag 0–2) of concentrations for respiratory disease and asthma ED visits and same-day (lag 0) concentrations for cardiovascular disease and congestive heart failure ED visits. These lags were generally consistent with those in an earlier study that showed that both lag time and amount of smoothing could affect the results, but that lag times of 2 or fewer days were more strongly associated with hospitalizations than were longer lag times (Katsouyanni et al. 2009).

The investigators explored four alternative pollutant models:

1. $PM_{2.5}$ only
2. 5 criteria pollutants ($PM_{2.5}$, O_3, CO, SO_2, and NO_2)
3. 7-pollutant model ($PM_{2.5}$, O_3, CO, SO_2, NO_2, OC, and NO_3^-)
4. 9-pollutant model (7 pollutants plus EC [correlation of 0.80 with OC] and SO_4^{2-} [correlation of 0.79 with $PM_{2.5}$])

The 7-pollutant model was chosen for the main analysis and was conducted in the 5-county data set. The model included linear and cubic terms for each pollutant and interactions of each pollutant with each other pollutant in the model.

Based on previous analyses of ED visits in Atlanta, the investigators specified a priori several covariates and model parameters to control for various temporal trends that might affect the results. They used a time-stratified design with variables for year, month, and day to account for changes that were not specifically included in the model (e.g., in population and demographic patterns). Hospital indicators were added to the model to distinguish periods when data were unavailable for each hospital. In addition, to assess whether changes over time other than those considered in the study might explain changes in ED visits, the authors also reported results for the number of ED visits for finger wounds as a negative control that should not be affected by changes in air quality. Thus, if changes were observed in ED visits for cardiovascular or respiratory diseases, but not for finger wounds, it would increase the confidence that the observed changes in health were actually related to changes in air quality.

Next, the investigators estimated the difference between the actual number of ED visits and the number of ED visits that would have been expected without regulatory actions (the counterfactual number) by calculating *daily risk ratios*: that is, the ratio of the number of actual ED visits to the number of ED visits estimated by applying the 7-pollutant health model to the counterfactual pollutant levels. From these ratios, and the data on the actual numbers of daily ED visits, the investigators estimated the *number of daily ED visits prevented* that could be attributed to the regulatory actions. Effects of regulations on ED visits were modeled for both individual regulations and for groups of regulations listed in Commentary Table 1.

Sensitivity Analyses The investigators also tested the sensitivity of their models to a number of analytic assumptions in the estimation of the number of ED visits (Commentary Table 3). The investigators reported results for the relationships between air quality and ED visits that were estimated using data from two periods: 1999–2005 and 1999–2013. The earlier period ended at the end of 2005 because that was when the reporting system for Atlanta-area hospitals changed. The investigators hypothesized that, if the models fit to the earlier period better represented the

Commentary Table 3. Main Results of Sensitivity Analyses in the Epidemiological Modeling and Application of the Health Model

Analysis	Purpose	Main Result
Sensitivity to Epidemiological Model Assumptions		
Form of the Model Linking Air Quality to ED Visits		
Compare the impacts of the shape of the relationships between pollutant levels and ED visits (linear vs. cubic, with and without interactions among pollutants).	Determine whether assumptions about how air quality affects health influence the main results.	Results were similar for all models (see IR Figure 18).
Number of Pollutants		
Compare the predicted changes in ED visits using models with 1 ($PM_{2.5}$), 5 ($PM_{2.5}$, O_3, CO, SO_2, NO_2), 7 (5-pollutant model plus OC and nitrate), and 9 (7-pollutant model plus EC and sulfate) pollutants.	Assess whether including more pollutants allows for more thorough analysis of the impacts of the air pollutant mixture as a whole.	Predicted benefits increased with increasing number of pollutants up to 7, and decreased for the 9-pollutant model (see IR Figure 19).
Data Used to Fit the Health Model		
Compare actual and counterfactual ED visits using health models of the relationships between air quality and ED visits using data from 1999–2005 (early in the study period) with those using data from 1999–2013 (full study period).	Determine whether the results were sensitive to the subset of data used to fit the health model.	Differences between actual and counterfactual scenarios were lower for associations estimated using data from the entire 1999–2013 period than from the period 1999–2005.
Size of Study Area		
Compare the difference between counterfactual scenarios and actual conditions in 5-county and 20-county Atlanta areas.	Test for potential impact of exposure uncertainty due to using a single central air quality monitoring site for all analyses.	Modeled impacts in the 20-county area were slightly smaller than those in the 5-county area (see IR Figure 17).
Sensitivity in the Health Impact Assessment Model Application		
Impacts of Individual Regulations		
Counterfactual scenarios were developed for each set of regulations with related goals, the mobile source regulations only, the EGU regulations only, and the combined effect of all regulations considered.	Determine which regulations likely contributed the most to any differences between actual conditions and scenarios without that regulation.	Policies affecting emissions from EGUs appeared to have a greater effect than policies affecting mobile sources, although that may have been related to different methods for the two source types (see IR Figure 20).
Result Report Period		
Results were presented for the full 1999–2013 period and for the final years of the study (2012–2013).	Compare the estimated benefits during implementation and after all regulations were implemented.	Annual reported benefits increased over time and were largest at the end of the study period (see IR Figure 15: complete time series; and IR Figures 16–20: 2012–2013 only).

relationships between exposures and health impact prior to full implementation of the regulations, the models from the earlier period would be more appropriate to estimate the counterfactual impacts (i.e., without regulation). In contrast, they hypothesized that if the relationship between air quality and health was related to factors other than the regulatory actions, then using the health model fit with data from the full period would be more appropriate. In the absence of evidence regarding which assumptions were correct, the investigators reported results for both periods.

In addition, although their main analysis was originally for a 20-county area, the investigators changed their main study area to a 5-county area because the HEI Research Committee overseeing the study was concerned that measurement error due to the use of a single air quality measurement site would be expected to be larger for the 20-county area than for a 5-county area. The investigators retained the original 20-county area for sensitivity analysis.

Characterization and Propagation of Uncertainty Through the Chain of Accountability

Russell and colleagues estimated uncertainty in each of the major steps in the chain of accountability in an effort to understand the contributions of different sources of uncertainty to overall uncertainty in the counterfactual concentrations. They used Monte Carlo simulation techniques to propagate uncertainty from emissions estimates, through to air quality, and then health outcomes. First, they estimated uncertainties in measured EGU emissions and modeled mobile source emissions, air quality measurements, and meteorological measurements using a combination of published values in the literature and their own analytic evaluations. Second, they sampled distributions of uncertainty in emissions and other inputs to the air quality models to create a set of 5,000 alternative estimates of counterfactual time-series concentrations. Third, they estimated uncertainty in the numbers of ED visits using the 5,000 sets of counterfactual air pollutant time-series concentrations as alternative inputs to the health models.

The 95% uncertainty intervals for differences between actual and counterfactual air pollutant concentrations and ED visits were represented by the range between the 2.5th and 97.5th percentiles of the simulated distributions. They conducted this analysis for the six sets of EGU and mobile source regulatory programs listed in Commentary Table 1, for EGU regulations and mobile source regulations separately, as well as for the combined EGUMOB scenario that included all major regulatory actions affecting EGUs and mobile sources over the period of the study. Detailed methods for uncertainty analysis are available in the Investigators' Report Appendix B.

RESULTS

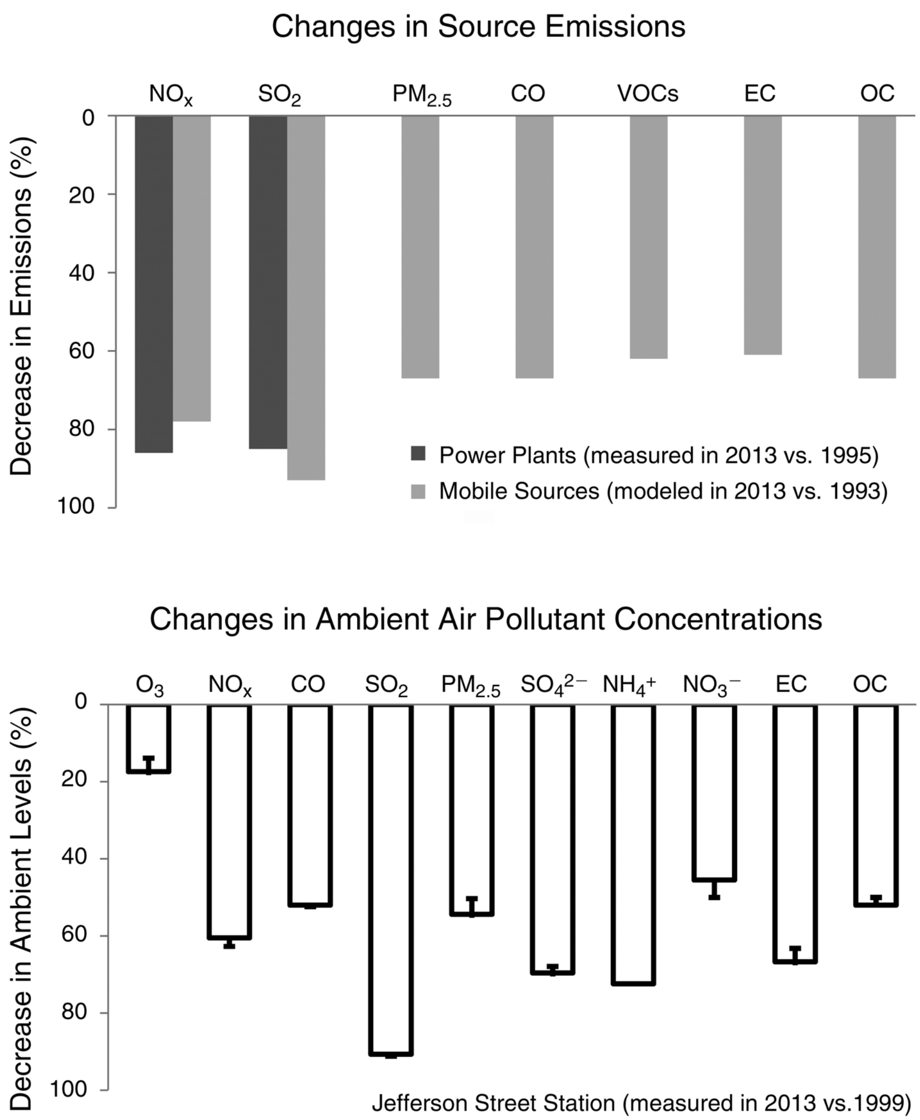
The investigators reported analyses for three of the linkages in the chain of accountability (Commentary Figure 2). First, the investigators reported their confidence in the selection and attribution of observed and estimated changes in emissions to groups of regulations based on their consultation with regulators and other stakeholders. Second, they reported that their multivariate air pollutant models estimated large improvements in air quality that were associated with the observed and estimated decreases in emissions. Third, they estimated differences between observed and counterfactual (i.e., without regulation) air quality scenarios for all regulations combined and reported that those differences translated into substantial and statistically significant differences in ED visits for asthma. Results for respiratory disease, cardiovascular disease, congestive heart failure, and finger wounds were consistent with no difference between observed and counterfactual air quality scenarios (although all point estimates were positive, the confidence intervals included the null) (see Investigators' Report Figures 15 and 16). Thus, the investigators concluded that emissions reductions associated with regulations as a whole led to lower air pollutant concentrations (both measured and modeled), and those in turn were associated with a reduced number of ED visits for asthma.

Effects of Regulations on Emissions and Air Quality

The current study reported that emissions from EGUs and mobile sources decreased from the beginning to the end of the study period (Commentary Figure 3). For example, by 2013, measured NO_x emissions from EGUs in Atlanta decreased by 86% (since 1997) and in the southeastern United States by 82% (since 1995). Over the same periods, measured EGU emissions of SO_2 decreased by 85% in Atlanta and 83% in the southeastern United States. Mobile source emissions could not be directly measured, but modeled mobile source emissions controls were estimated to result in 61% to 93% lower emissions, depending on the pollutant (NO_x, SO_2, $PM_{2.5}$, CO, EC, OC, or VOCs). Using air quality models, the investigators tried to parse out to what extent each of the six regulations considered in this study contributed to these decreases in emissions.

The investigators reported that air pollutant levels measured at the Jefferson Street air quality monitoring station were lower in 2013 than in 1999 for all pollutants considered (NO_x, SO_2, $PM_{2.5}$, CO, EC, OC, O_3, sulfate, ammonium, and nitrate; Commentary Figure 3). If observed concentrations at the Jefferson Street site were lower than counterfactuals, the investigators concluded that the decreases in air pollution were attributable to the regulations.

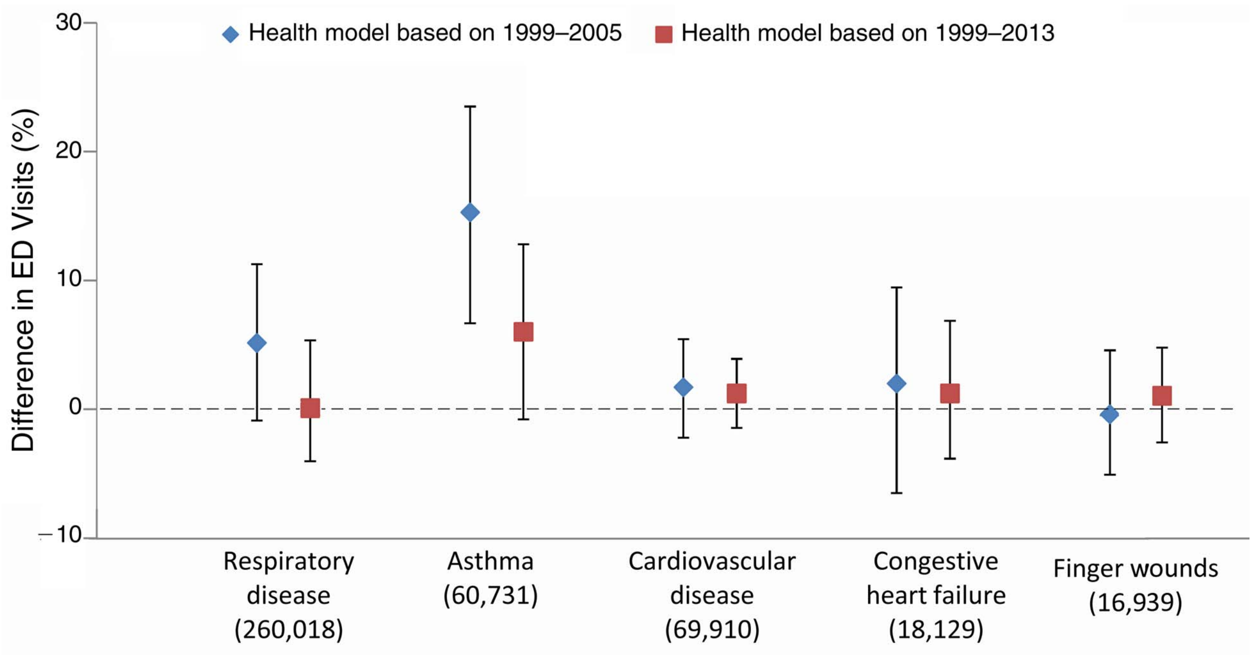
According to the investigators, the *combined effects of EGU and mobile source controls* (EGUMOB) resulted in



Commentary Figure 3. Decreases in source emissions and air pollutant concentrations in Atlanta during the study period. Top: Change from the base years to 2013 in the Atlanta nonattainment area. **Bottom:** Changes in air pollutant levels at the Jefferson Street air quality monitoring station between 1999 and 2013. Error bars show the difference in percentage change after removing the influence of meteorology. NH_4^+ = ammonium, NO_3^- = nitrate, SO_4^{2-} = sulfate.

lower annual average concentrations in 2013 relative to the counterfactual concentrations for NO_2, CO, SO_2, $PM_{2.5}$, sulfate, ammonium, nitrate, EC, and OC (see Table 6 in Investigators' Report). The investigators did not report substantial changes in annual average O_3 concentrations or acidity of $PM_{2.5}$ related to EGU or mobile source controls. They reported decreases in air pollutant levels related to changes in emissions from both EGUs (both within and outside of the Atlanta nonattainment area) and mobile sources (see Figure 7 in the Investigators' Report). They also reported decreases in *EGU emissions* of subsets of each of these pollutants related to the Acid Rain Program, NO_x Budget Trading Program, and Clean Air Interstate Rule independently. (See Figures 4 and 10 in the Investigators' Report for time series of the effects on emissions and air quality, respectively, of the different regulations affecting EGU emissions.)

The models also showed decreased concentrations of $PM_{2.5}$, NO_2, ammonium, CO, SO_2, OC, and EC due to programs to reduce *mobile emissions*, with the amount of the improvement increasing over time. Over the study period, annual average O_3 levels decreased, with lower peak O_3 levels in summer and higher minimum O_3 levels in winter, whereas nitrate levels decreased in the summer but were not affected in the winter. Thus, lower concentrations of



Commentary Figure 4. Estimated percentage difference between actual and counterfactual ED visits in 2012–2013 for two different models for all pollution control policies combined. Results are presented for the 5-county Atlanta metropolitan area, by outcome and by period used to estimate the relationship between air quality and health. Positive percentage differences represent actual ED visits that were lower than those in the counterfactual (without regulation) scenarios. Whiskers represent the 95% confidence intervals. Actual numbers of ED visits in 2012–2013 are listed in parentheses for each outcome. The health model included 7 pollutants with all cubic polynomial and interaction terms. Finger wounds were analyzed as a negative control for residual bias.

all pollutants analyzed were observed, but some were not lower during the winter season. Both the gasoline fuel programs and the Heavy Duty Diesel Rule contributed to the total reductions, although the Inspection and Maintenance programs appeared to have minimal effect on the pollutants considered. The mobile source programs reduced emissions of $PM_{2.5}$, SO_2, and NO_2 during the period evaluated, but did not affect sulfate emissions, which are a small part of the total sulfur emissions. (See Figures 5 and 11 in the Investigators' Report for time series of the effects on emissions and air quality, respectively, of the different regulations affecting mobile source emissions.)

When the investigators compared multiple methods for estimating mobile NO_x emissions, they concluded that NO_x emissions throughout the period may have been overestimated by MOVES but that the trends in emissions were accurately predicted. Similarly, while meteorology had an important influence on air quality on a daily scale, they concluded that results from the meteorological detrending and empirical ambient concentration–emissions models suggested that multiyear trends in ambient air pollution concentrations at the Jefferson Street monitoring site were driven by changes in emissions.

Effects of Air Quality Improvements on Cardiorespiratory ED Visits

The investigators estimated that the air quality improvements collectively resulting from all regulations were associated with substantial reductions in hospital ED visits (Commentary Figure 4). Using relationships between air quality and ED visits from 1999–2005, they estimated that in the 5-county Atlanta area there were a total of 55,794 fewer cardiovascular and respiratory ED visits over the 1999–2013 study period than would have occurred without regulation. Of those ED visits, they estimated that 17,977 (about one-third) would have occurred in the final two years of the study period (2012–2013) after the regulations being evaluated were closer to their full implementation. Estimates of differences between actual and counterfactual ED visits were generally larger for health models based on the relationships between air quality and health from 1999–2005 than for health models based on data from 1999–2013. One possible explanation for the differences was that the air pollution mixture may have changed sufficiently over the study period, leading to changes in the relationship between air quality and health. The actual numbers of ED visits that occurred in the later years of the study period might already reflect impacts of changes in the air pollution mixture related to the regulatory programs, and thus could explain the reduction of observed differences between actual ED visits and the counterfactual estimates when the health models based on the full period from 1999–2013 were used.

The investigators reported most of their results as percentage differences between the actual and counterfactual (i.e., without regulation) scenarios for several reasons. First, they reasoned that the main results should be presented in such a way that they would not depend too strongly on the definition of the study area and period. This would allow them to more easily compare results using two different study area boundaries (5-county or 20-county) and over a long period during which population and total ED visits increased. If only the numbers of ED visits were compared, changes in the difference in ED visits with and without regulation in the later years compared to the early years could be attributed to control policies being more fully implemented, the increasing population of Atlanta, or both.

Most of the difference in total ED visits between scenarios with and without regulation resulted from differences in asthma ED visits. For the last two years of the study period the models predicted about 6.8% (health model using relationships between air quality and health from 1999–2013) or 16.5% (health model using relationships between air quality and health from 1999–2005) fewer actual asthma ED visits than would have occurred without regulation (Commentary Figure 4). The differences between actual and counterfactual ED visits for cardiovascular and congestive heart failure were similar to those for the negative control (i.e., finger wounds) regardless of the data used to create the health model. The investigators reported larger differences between actual and counterfactual ED visits for policies to reduce EGU emissions than for policies to reduce emissions from mobile sources, although the emissions estimates were based on measurements for EGUs and on models for mobile sources.

The investigators stated that the sensitivity analyses showed that these overall results were unlikely to have been strongly influenced by major assumptions related to model parameterization, study area size, or exact levels of emissions. However, the results were sensitive to the period used to build the health models. They reported that the 95% uncertainty intervals for estimates of ED visits were smaller than the difference between the counterfactuals and the observations and that contributions to uncertainty were larger for statistical model parameters than for changes in EGU or mobile emissions.

HEI REVIEW COMMITTEE'S EVALUATION

INTRODUCTION

In its independent review of the report, the HEI Review Committee noted that the study was an ambitious application of HEI's accountability framework as it encompassed a broad suite of regulatory programs designed to reduce multipollutant emissions from power plants and mobile sources in Georgia and nearby states over the period from 1999 to 2013. The Committee thought that the investigators had tackled an important public health question, examining whether the regulations had individually or collectively reduced emissions, improved air quality, and ultimately reduced ED visits for respiratory and cardiovascular outcomes in the Atlanta area.

DISCUSSION OF DATA AND METHODS

Emissions, Air Quality, and Health Data

The investigators used large, high-quality, and well-documented data sets of air pollutant emissions in the southeastern United States and of air quality and ED visit s in Atlanta. The Committee thought that the size, quality, and extensive prior evaluation of the data sets in earlier studies were reassuring and were important starting points for the current study. They used air quality measurements from the Jefferson Street monitoring site, which is well known for both the quality of the data and the long period it has been in operation.

The investigators indicated that they chose to use air quality measurements from only the Jefferson Street site because it was the only site where measurements of all considered pollutants were colocated for the entire study period and that bringing in additional monitors would have complicated the analyses and potentially introduced additional error. They had originally proposed to analyze a 20-county area, but ultimately decided to analyze the 5-county area in their main analyses and conduct a 20-county area analysis as a sensitivity analysis. In discussing their results, the investigators suggest that their estimates of slightly larger relative impacts of regulations on ED visits in the 5-county area than in the 20-county area are evidence that using a single air pollution monitoring site has biased their results toward underestimating the true effects in the larger area. While the exposure measurement error is a plausible hypothesis for the differences observed, the Review Committee thought a more systematic consideration of other factors that might account for the differences between the 5- and 20-county results would also provide insight into these differences.

The hospital data used in this study have also been extensively vetted and tested for the sensitivity of results in epidemiological studies of air quality and health to a large number of methodological assumptions. For example, the team at Emory University has used these data in previous studies on the sensitivity of results in epidemiological studies of air quality and health to a large number of methodological assumptions, including the appropriateness of assigning exposures in time-series studies based on air quality at the Jefferson Street monitoring site (Darrow et al. 2011; Sarnat et al. 2010, 2013; Strickland et al. 2011, 2015). In developing estimates of the health impacts averted by reductions in emissions and air pollution, the Committee agreed that use of relationships between ED Visits and air quality specific to the Atlanta area was preferable to reliance on relationships observed in other populations that might have different underlying health status or other characteristics that might change their susceptibility to air pollution exposure.

Evaluation of the Chain of Accountability

A strength of the study was that the investigators systematically and logically addressed four steps in the chain of accountability: from the identification of key EGU and mobile source regulations to estimating their influence on emissions to the effect of emissions on air quality and finally to the effect of air quality on ED visits. Another strength of the study was that the investigators used a logical approach to develop robust counterfactual methods to estimate the differences between conditions with and without regulations at each link in the analysis.

The investigators faced several major challenges common to accountability studies of regulations that take place over a long period. The regulations went into effect at different points of time over the study period, and some of the regulations may not have been fully implemented by the end of the study period. In addition to meteorological conditions — which were controlled for in the models — a number of factors including economic conditions can influence temporal changes in air quality and complicate the task of isolating the effect of regulatory actions. Similarly, changes in hospital ED visits for specific health outcomes can be driven by multiple factors including changes in population characteristics, in healthcare access, and in treatment practices, further complicating the task of isolating the effects of changes in air pollution and regulations.

Characterization of Regulatory Programs and Their Impact on Emissions The Committee thought that the investigators had used a logical and systematic approach to identifying the key regulatory programs affecting EGUs and mobile sources, to estimating emissions reductions, and to apportioning the reductions to legislative or regulatory programs. The investigators worked with regulated industries, as well as with the appropriate agencies in charge of the regulations, to clarify details regarding the actual implementation of the regulatory programs and other factors affecting decisions made regarding operation and facility improvements. The investigators tested the sensitivity of their emissions estimates to a number of assumptions and methodological choices. For example, they explored four different methods of evaluating the potential for systematic bias in estimates of NO_X emissions from mobile sources. The Review Committee concluded that these efforts all added to the credibility of the work.

Several large-scale trends could have affected the results and may not have been fully addressed, so that explanations other than regulation could explain changes in emissions. For instance, the assessment is made more complex by the significant switch from coal to natural gas over the last decade, which in part may be a response to regulations, but likely also was a response to the falling natural gas prices. Another analysis that would have informed the comparison of the results based on the different health models was a detailed evaluation of possible changes in the composition of emissions or air pollutants such as PM due to changes in technology during the study period. The investigators did evaluate acidity (see Commentary Table 2), but other potential changes in PM composition were not assessed. It would have been useful to know whether such changes occurred over the time frame of the study, in particular in relation to the changes in vehicle technology.

Characterization of the Impact of Emissions on Air Quality The Committee liked the investigators' use and intercomparison of two different methods — multivariate regression and the CMAQ chemical transport model — to estimate the relationship between emissions and pollutant concentrations and to remove the effects of meteorological trends on the pollutant concentrations over time. Although the CMAQ modeling results were not used in the counterfactual and health analyses, the investigators were able to compare the two independent approaches to finding relationships between emissions and air quality. This comparison provided evidence that meteorological trends were effectively removed from the air quality time series and thus not likely to have influenced the results of the health analysis. The Review Committee agreed with this conclusion.

By developing new detrending methods to remove the influence of meteorology and other temporal factors (e.g., day-of-week or season) on emissions and air quality, the investigators decreased the potential for temporal confounding. The Review Committee thought this was an important contribution, both because a previous evaluation of traffic-reduction policies during the 1996 Olympic Games in Atlanta had shown that regional meteorology can be an important confounder (Peel et al. 2010) and because the current study was conducted over a long period (15 years), when other temporal factors can come into play.

Other accountability studies have used time-varying counterfactual models similar to those used in the current study to evaluate the effects of regulations (Morgenstern et al. 2012; Tonne et al. 2008). However, either those studies did not go beyond detailed air pollution modeling of the impacts on air quality to also address the impacts on health (Morgenstern et al. 2012) or they evaluated air quality impacts on health without detailed models of changes in air quality over time (Tonne et al. 2008). Morgenstern and colleagues (2012) applied counterfactuals to EGU emissions for SO_2 using a similar method of building statistical models of $PM_{2.5}$ based on meteorology and emissions. The current study was more comprehensive than these earlier studies because it addressed in detail and on a daily basis all the linkages among regulations, emissions, air quality, and health. The current study also included more detailed analysis of sensitivity to model assumptions than is typically done.

Multipollutant Health Effects Modeling The Review Committee found the investigators' approach to the multipollutant modeling to be appropriate and comprehensive. The Committee thought that the investigators selected a reasonable set of health outcomes for evaluating ED visits, although omission of mortality outcomes neglects a major potential benefit of the regulations. In sensitivity analyses, the investigators explored several multipollutant model formulations to assess the joint effect of reductions in pollutant concentrations, including varying numbers of pollutants and cubic or interaction terms in the models. In the current study, the investigators wanted to characterize more thoroughly the effects of changes in the overall air quality mixture than they had done in their previous studies. They used their multipollutant approach as evidence that "if assumptions about pollutants being independently and linearly associated with health outcomes are false, these more detailed models could result in more accurate and complete assessments of the impact of pollution control policies." The Committee was not completely convinced by this reasoning because they thought overparameterization was likely, but thought that building models with different numbers of pollutants for comparison was useful.

Both the model design and choice of time-varying covariates were selected to control for temporal trends that might otherwise explain or mask the trends in ED visits attributed to air pollution trends. The Committee thought that the addition of a parallel analysis of ED visits for finger wounds as a negative control — an outcome not expected to be affected by air pollution — provided useful insights.

A strength of the analyses was that, for many decisions, modeling choices were made a priori based on previous analyses in these data sets. These choices included the specific 7 pollutants to consider in the main health model, the lags (the number of days between the exposure and the ED visit) for each health outcome, an indicator variable to adjust for the time-varying contribution of data from different hospitals, multiple terms for meteorological variables, time terms (day-of-week, month, and year), and time-interaction terms to control for long-term trends.

In other aspects of the analyses, however, the Committee had some concerns that the modeling choices were in part influenced by results. The investigators argued that the 7-pollutant models most fully captured the joint effect of air pollution on ED visits. The Review Committee generally agreed that this was a plausible hypothesis. However, the Committee thought that it remained unclear whether the 7-pollutant models were truly a better representation of the relationship between air pollution and health relative to alternative multipollutant models, and to what extent the models may have been overparameterized or reflected changes in composition of the pollution mixture.

Another key sensitivity analysis related to the periods during which data were selected to construct the health models, that is, the entire study period (1999–2013) versus the first 7 years of the study period (1999–2005). Results for ED visits in 2012–2013 using health models based on those two periods were different; larger estimates of effects of air pollutant changes on ED visits occurred when data from the earlier period were used to construct the model. It was unclear why such differences were found, or whether they were attributable to modeling artifacts rather than to true differences in health responses across the respective periods.

Uncertainty Analyses

The Committee thought the investigators' efforts to estimate and propagate uncertainty starting with emissions and through to the health effects estimates were a useful complement to the many sensitivity analyses conducted in the development and evaluation of emissions models, air quality models, and health models. Despite the importance of understanding uncertainty in analyses like these, such comprehensive analyses are rarely done. Yet, done well, uncertainty analyses can be informative, both for identifying which elements of the analysis are least well understood and might benefit from further research and for understanding how much weight to put on the final analyses as input to decisions. Without propagation of uncertainty through all steps of the analysis, uncertainty is represented only by uncertainty in the health effects model parameters and is very likely to be underestimated.

In this study, the authors used standard Monte Carlo sampling techniques to conduct their analysis. They essentially treat as known the underlying *form of the models* used to (1) predict counterfactual air pollutant concentrations from emissions and (2) predict the relative risk of ED visits, and they treat the set of *input parameters* to those models as uncertain. The distributions chosen to characterize uncertainty in the parameters were generally standard parametric distributions (e.g., normal and uniform).

The investigators' discussion would have benefited from more assessment of the propagation of uncertainty in the investigators' analyses, in particular in relation to some of the other key assumptions tested — for example, the choice of period with which to fit the time-series models, which the investigators' sensitivity analysis showed are more influential than the uncertainty analyses in some cases. This reality is a reminder that model uncertainty, which is often difficult to characterize, can be more influential than the parameter uncertainties taken into account in the Monte Carlo analysis conducted in the current study. Although the uncertainty estimates were not necessarily comprehensive, and included assumptions that could be considered arbitrary, the Committee thought that the investigators' choices were reasonable, and inclusion of different periods of analyses was a strength of this study as it highlighted the influence of different assumptions and approaches. Moreover, this work set an important example that may be used as a starting point for future studies.

DISCUSSION OF FINDINGS

Main Findings along the Chain of Accountability

The Committee agreed with the investigators' basic conclusions that emissions of NO_x, SO_2, CO, VOCs, $PM_{2.5}$, EC, and OC from both power plants and mobile sources had decreased over the study period, and in later years of the study period were lower than what would have been expected had regulatory actions not been taken. However, there was no discussion about the degree to which the transition in power plant fuel usage from predominantly coal (four times more coal than natural gas in 2002) to nearly equal usage of coal and natural gas near the end of the study period (Cabral 2017) was a result of response to regulatory requirements for cleaner combustion, or to significant marketplace changes, or both.

In evaluating the step from changes in emissions to changes in air quality, the Committee thought the investigators' method to account for meteorological trends and to isolate the effect of emissions reductions on air quality was technically sound, especially given the limitations of data available for the retrospective study. However, an important source of uncertainty was the use of air quality measurements from a single fixed-site monitoring location in the regression analysis of relationships between emissions and air pollutant levels and ultimately in the estimation of the relative risks of ED visits. One fixed site is often not representative of broader regional areas, in particular in estimating exposure to emissions from mobile sources for pollutants with high spatial variability (e.g., NO_x) (HEI Panel on the Health Effects of Traffic-Related Air Pollution 2010; Sarnat et al. 2010). This was a potential concern raised by both the HEI Research Committee during study oversight and by the Review Committee during review of the final report. In previous work, the investigators evaluated the effect of using different monitors and study areas on exposure estimates in the 20-county Atlanta region. They showed that air quality improvements were greater at urban center sites in Atlanta than at more rural sites; however, the effect of this difference on associations between health outcomes and air quality within ~30 km of the site, or within the 5-county area, was negligible (Sarnat et al. 2010). Based on these previous studies, the investigators thought that the effects of the mobile source regulations were likely underestimated, because the central monitoring site was not able to adequately capture near-source traffic-related air pollution exposure. In addition, the effects of using a single monitor in health impact studies that rely on associations between air quality and health, like the current study, depend on whether the pollutants are spatially uniform or locally elevated (Strickland et al. 2011). The Review Committee agreed.

The Committee agreed with the investigators that, for all regulations combined, there were fewer ED visits for asthma in later years than would have been expected without regulatory actions. The Committee also noted limitations in some of the results. In particular, the Committee thought that the extent of differences in actual and counterfactual ED visits was quite modest for cardiovascular disease and congestive heart failure despite a pronounced decrease in air pollution concentrations and that these results were similar to those for finger wounds (the negative control). Nevertheless, the results for asthma and respiratory ED visits suggest that regulations have improved health, even if the exact magnitude of the effect is unclear.

Contribution of Individual Regulations to Changes in Air Quality and Health

One of the attractive features of this study was that it offered the opportunity to examine the effect of individual, as well as collective, control policies for power plants and for mobile source emissions on air quality and health. The Committee appreciated the investigators' presentation of results for all programs combined (EGUMOB), for EGU and mobile source programs separately, and for individual programs within those categories.

However, the Committee ultimately concluded that for multiple reasons the results for the combined suite of control policies were more robust than those for individual programs. For one, the timing of implementation and effects on emissions of regulations were not known with equal levels of certainty for all regulations considered. Further, uncertainties in the health outcome analyses, resulting in part from overlapping regulatory programs implemented over a long period, made it difficult to verify the attribution of differences between actual and counterfactual scenarios to specific programs. The Committee suggests that the relative impact of the individual programs should be interpreted with caution. Similar challenges have been encountered in other accountability studies that looked at the impacts of one specific regulation (Morgenstern et al. 2012) or broad changes in air quality and health but not individual regulations (Gilliland et al. 2017; Peters et al. 2009).

Despite the uncertainties in estimating impacts of distinct regulatory programs, the study suggests that relative to their respective counterfactuals, controls on EGUs have had more impact than controls on motor vehicles. Estimation uncertainties may partly account for this finding. A single air quality monitor is expected to better represent impacts of power plant emissions on air quality (via secondary pollutant formation and more regionally uniform contributions) than it would represent mobile source impacts that are more spatially varied. In addition, the level of uncertainty in modeled mobile source emissions was larger than the level of uncertainty in measured EGU emissions, and this may have decreased the ability to estimate emissions reductions from mobile sources, especially since there were proportionally lower estimated reductions in emissions from mobile sources than in those from EGUs. Another factor in the relatively lower estimated impact of the mobile source programs may be that several programs came into effect later in the study period; for example, the Heavy Duty Diesel Rule was phased in between 2004 and 2012. Therefore, the full effect of the mobile source regulations on emissions, air quality, and health may not have yet been observed as their effects would continue well beyond 2013 as fleet turn-over progresses. In addition, the implementation and enforcement of those rules have been uneven (Yang et al. 2017).

Temporal Confounding and Sensitivity of the Results

As noted above, the estimates of the difference between actual and counterfactual ED visits were different when the health effects model was based on data from 1999–2005 rather than from the full study period of 1999–2013. The Committee concurred with the investigators' approach of presenting the results based on both periods and attempting to characterize the uncertainties. However, the Committee thought this difference was unexplained and wondered whether the evident sensitivity of the relationship of ED visits with air quality to the period of the health effects model reflected unmeasured temporal confounding. Still, it is possible that the investigators' conclusions reflect a stronger weighting of the results from the earlier period, and perhaps the results should be weighted more equally since it is not clear which health model was more appropriate.

The investigators made some plausible arguments for why the coefficients might have changed from one period to another, such as changes in composition of the pollutant mixture and nonlinearities in the relationships between air quality and health over the range of concentrations. It would have been useful to test some of these suggestions with existing data on the composition of the air pollution mixture. For example, $PM_{2.5}$ composition measurements are available from the Jefferson Street monitoring station and other sources. Also, additional sensitivity analyses could have been done to directly evaluate the issue of nonlinear relationships.

Some other changes over time may provide alternative explanations for the differences between the periods analyzed, including changes in population demographics and healthcare, and these could have been explored. For example, the population in the 5-county Atlanta metropolitan area grew by 24.2% between 1999 and 2013 (United States Census Bureau 2013), but the population increased more slowly than the annual numbers of ED visits at Atlanta hospitals. In addition, there were changes in healthcare that could have decreased the number of ED visits, such as the passing of the Affordable Care Act and better ED visit prevention by, for example, increased availability of medications to treat asthma or cardiovascular disease. In light of these concurrent changes, the Committee found that there were more unanswered questions regarding the link between health effects and air quality than regarding the link between emissions and air quality. The Committee also noted it was unclear what the implications of changes in healthcare access and treatment practice may have been.

Also, the Committee noted that the investigators used hospital indicators adjusted for the size of the hospital for whether any given hospital was reporting on a particular day (Russell AG, personal communication, 14 March 2018). To address the possible effect of missing data, the investigators provided summary statistics on the hospitals, showing that the number of hospitals reporting did not change substantially throughout the study period (see Investigators' Report Appendix Table C.4, available on the HEI website). The Committee was reassured by these summary statistics, although they thought there may have been more subtle issues with the completeness of the hospital data that perhaps could only be found by studying subsets of the hospitals in greater detail. The addition of finger wounds as a negative control was reassuring in this regard; however, comparison to finger wounds is not likely to account for changes over time in healthcare or treatments that might be more directly related to the more serious outcomes of interest for this study. A more nuanced discussion of these potential sources of confounding would have been valuable. Therefore, although the investigators were able to control for temporal confounding due to meteorology, they may not have fully captured other sources of confounding of the relationship between air pollution and ED visits in their data.

An ongoing question regarding accountability studies is whether they can provide information on whether there is a causal relationship between the regulatory actions and any observed improvements in air quality and health, regardless of whether the investigators used causal methods (Dominici and Zigler 2017). Because they did not use causal methods, the authors have carefully discussed their results in terms of *estimates* of prevented ED visits. Even so, in the Committee's view, the term "prevented" suggests a more definitive conclusion than may be merited, considering the challenges in accounting for changes in healthcare access and treatment practices.

Comparison to Regulatory Impact Assessments

One of the purposes of accountability studies is to evaluate whether regulatory programs actually produced the benefits anticipated in the Regulatory Impact Assessments (RIAs) that justified them in the first place (Henneman et al. 2016; Rich 2017). As anticipated in the RIAs for the regulatory programs considered in the current study, the investigators found reduced emissions, improved air quality, and reduced health impacts that could be attributed to regulations to control air pollutant emissions from EGUs and mobile sources. Unfortunately, direct comparison with the RIAs was not feasible because the various studies had different spatial and temporal scales, methods, and assumptions. For example, the populations, study area, and number of pollutants considered in the current study were different from those in the RIAs, and there was little overlap in health outcomes. For these and similar reasons, the comparison of retrospective accountability study results to RIA predictions is challenging and often not done.

Some commenters have suggested that challenges related to the comparison of prospective and retrospective studies could be minimized in the future if important long-term regulatory programs were to have built-in accountability components (Hubbell 2012; Hubbell and Greenbaum 2014; Rich 2017). Prospective study design for accountability would allow for comparable metrics (e.g., air pollution measurements and health outcome records) with a detailed plan to collect the appropriate data to efficiently evaluate the impacts of the regulation on the environment and health.

CONCLUSIONS

The Committee thought that this report by Russell and colleagues is a valuable addition to the accountability literature. This is one of few accountability studies to examine the effects of regulations on emissions all the way through to health outcomes, using scenarios based on observed data. Though labor-intensive, this approach is valuable and worth considering for future accountability studies in the United States or elsewhere in the world.

Russell and colleagues addressed important questions raised in RFA 11-1 about (1) our ability to discern the effects of regulatory and other actions at the national or regional level implemented over multiple years and (2) approaches that can be used to evaluate complex sets of actions targeted at improving air quality in large urban areas. Specifically, the investigators were able to show that major regulations targeting power plants and mobile sources were effective in reducing pollutant emissions, improving air quality, and ultimately reducing ED visits (in particular those for asthma) in the Atlanta area. This was a complex and multifaceted study, and the Committee thought the methods were appropriate, the analyses were done correctly, and the report was well written.

The study accomplished its objectives, most notably to link changes in air quality to regulatory programs. This was a formidable task and required a series of well-reasoned decisions. A major strength of this study is that it followed the accountability framework laid out in a previous HEI report (HEI Accountability Working Group 2003). Other strengths were the well-characterized data sets on air pollutant levels and ED visits, parallel approaches (empirical modeling and chemical transport modeling) to construct the counterfactual time series of pollutant concentrations, and extensive sensitivity and uncertainty analysis that were done in the development and application of the air quality and health models.

The Review Committee thought this study used counterfactual scenarios in a novel way in accountability research by applying them to health analyses, and agreed with the major findings of the study that actual emissions, air pollutant levels, and ED visits were lower than for counterfactuals that reflected estimated projections of conditions under which the regulations had not been implemented. On the other hand, the Committee also thought the differences in estimates for ED visits using data from two different periods suggested there was uncertainty that was not fully accounted for, and did not fully agree with some interpretations. For example, the Committee thought the estimated differences in ED visits for cardiovascular disease and congestive heart failure between with- and without-regulation scenarios were very small and were not different from those for finger wounds, whereas the investigators had assigned the results for cardiovascular disease and congestive heart failure more significance.

Although Russell and colleagues presented results for individual regulatory programs, the Committee concluded that the results for the combined suite of control policies were more robust than those for individual programs. Actual implementation of the regulations may have been different from the reported changes, and alternative explanations other than regulation are plausible. For example, there could have been changes to reduce operational costs of EGUs or improve motor vehicle efficiency that could have reduced emissions from those sources.

The investigators' finding that regulations targeting power plants had more impact on improving air quality and reducing ED visits than regulations targeting mobile sources needs further study. The Committee raised several issues related to direct comparisons of the impact of the different regulations, including that there were different levels of uncertainty in the emissions levels associated with the different sources, that mobile source programs were implemented later than the EGU programs, and that rules were unevenly enforced. The Committee concluded that the relative impact of the individual programs should be interpreted with caution.

Russell and colleagues reported that the overall results were robust to the geographical scale of assessment and the number of pollutants in the health models, and less robust to the period of assessment. The Review Committee had more confidence in the link between emissions and air quality than in the links between regulations and emissions or between air quality and ED visits, because the investigators not only statistically linked air quality with changing emissions, but also used alternative modeling strategies to rule out other likely causes of changes in air quality (e.g., meteorology) as dominating factors.

In the future, other researchers could apply these methods to the long-term impacts of regulations on health outcomes in other locations, although it would be recommended to more thoroughly account for changes in medical practice and healthcare access, where possible. In particular, disentangling the effects of specific regulations among a suite of regulations remains challenging, and efforts should continue. This body of work is a strong contribution to HEI's accountability research portfolio because it sequentially and carefully addresses multiple links in the chain of accountability. The results suggesting that reductions in emissions and improved air quality were linked to health benefits are important in terms of continued evaluation of the public health benefits of air pollution regulation in the context of implementation and compliance issues that may hamper achievement of the intended benefits.

ACKNOWLEDGMENTS

The 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 Hanna Boogaard and Annemoon van Erp for their oversight of the study, to Allison Patton and Annemoon van Erp for their assistance in preparing its Commentary, to Carol Moyer for science editing of this Report and its Commentary, and to Fred Howe, Hilary Selby Polk, and Ruth Shaw for their roles in preparing this Research Report for publication.

REFERENCES

Bhaskaran K, Gasparrini A, Hajat S, Smeeth L, Armstrong B. 2013. Time series regression studies in environmental epidemiology. Int J Epidemiol 42:1187–1195; doi:10.1093/ije/dyt092.

Blanchard CL, Hidy GM. 2005. Effects of SO_2 and NO_x emission reductions on $PM_{2.5}$ mass concentrations in the southeastern United States. J Air Waste Manag Assoc 55:265–272; doi:10.1080/10473289.2005.10464624.

Blanchard CL, Hidy GM, Tanenbaum S. 2010. NMOC, ozone, and organic aerosol in the southeastern United States, 1999–2007: 2. Ozone trends and sensitivity to NMOC emissions in Atlanta, Georgia. Atmos Environ 44:4840–4849; doi:10.1016/j.atmosenv.2010.07.030.

Boogaard H, van Erp AM, Walker KD, Shaikh R. 2017. Accountability studies on air pollution and health: The HEI experience. Curr Environ Health Rep; doi:10.1007/s40572-017-0161-0.

Cabral L. 2017. Natural gas power generation share grew in Southern states for a decade as coal declined. Today in Energy, 5 December 2017. U.S. Energy Information Administration. Available: *www.eia.gov/todayinenergy/detail.php?id=33992* [accessed 12 December 2017].

Cohan DS, Hakami A, Hu Y, Russell AG. 2005. Nonlinear response of ozone to emissions: source apportionment and sensitivity analysis. Environ Sci Technol 39:6739–6748; doi: 10.1021/es048664m.

Darrow LA, Klein M, Sarnat JA, Mulholland JA, Strickland MJ, Sarnat SE, et al. 2011. The use of alternative pollutant metrics in time-series studies of ambient air pollution and respiratory emergency department visits. J Expo Sci Environ Epidemiol 21:10–19; doi:10.1038/jes.2009.49.

Dockery DW, Rich DQ, Goodman PG, Clancy L, Ohman-Strickland P, George P, et al. 2013. Effect of Air Pollution Control on Mortality and Hospital Admissions in Ireland. Research Report 176. Boston MA:Health Effects Institute.

Dominici F, Zigler C. 2017. Best practices for gauging evidence of causality in air pollution epidemiology. Am J Epidemiol 186:1303–1309; doi: 10.1093/aje/kwx307.

Eskridge RE, Ku JY, Rao ST, Porter PS, Zurbenko IG. 1997. Separating different scales of motion in time series of meteorological variables. Bulletin of the American Meteorological Society 78:1473–1483; doi:10.1175/1520-0477(1997)078<1473:sdsomi>2.0.co;2.

Flanders WD, Klein M, Darrow LA, Strickland MJ, Sarnat SE, Sarnat JA, et al. 2011. A method to detect residual confounding in spatial and other observational studies. Epidemiology 22:823–826; doi:10.1097/EDE.0b013e3182305dac.

Flanders WD, Strickland MJ, Klein M. 2017. A new method for partial correction of residual confounding in time-series and other observational studies. Am J Epidemiol 185:941–949; doi:10.1093/aje/kwx013.

Flaum JB, Rao ST, Zurbenko IG. 1996. Moderating the influence of meteorological conditions on ambient ozone concentrations. J Air Waste Manag Assoc 46:35–46; doi:10.1080/10473289.1996.10467439.

Friedman MS, Powell KE, Hutwagner L, Graham LM, Teague WG. 2001. Impact of changes in transportation and commuting behaviors during the 1996 summer Olympic Games in Atlanta on air quality and childhood asthma. JAMA 285:897–905; doi:10.1001/jama.285.7.897.

Gilliland F, Avol E, McConnell R, Berhane K, Gauderman WJ, Lurmann FW, et al. 2017. The Effects of Policy-Driven Air Quality Improvements on Children's Respiratory Health. Research Report 190. Boston, MA:Health Effects Institute.

Goldman GT, Mulholland JA, Russell AG, Strickland MJ, Klein M, Waller LA, et al. 2011. Impact of exposure measurement error in air pollution epidemiology: effect of error type in time-series studies. Environ Health 10:61; doi:10.1186/1476-069X-10-61.

Hansen DA, Edgerton ES, Hartsell BE, Jansen JJ, Kandasamy N, Hidy GM, et al. 2012. The Southeastern Aerosol Research and Characterization Study: Part 1—overview. J Air Waste Manag Assoc 53:1460–1471; doi:10.1080/10473289.2003.10466318.

Health Effects Institute. 2010. Proceedings of an HEI Workshop on Further Research to Assess the Health Impacts of Actions Taken to Improve Air Quality. Communication 15. Boston, MA:Health Effects Institute.

HEI Accountability Working Group. 2003. Assessing Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research. Communication 11. Boston MA:Health Effects Institute.

HEI Panel on the Health Effects of Traffic-Related Air Pollution. 2010. Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. Special Report 17. Boston, MA:Health Effects Institute.

Henneman LR, Liu C, Mulholland JA, Russell AG. 2016. Evaluating the effectiveness of air quality regulations: A review of accountability studies and frameworks. J Air Waste Manag Assoc 67:144–172; doi:10.1080/10962247.2016.1242518.

Henneman LRF, Liu C, Hu Y, Mulholland JA, Russell AG. 2017. Air quality modeling for accountability research: operational, dynamic, and diagnostic evaluation. Atmos Environ 166:551–565; doi:10.1016/j.atmosenv.2017.07.049.

Hubbell B. 2012. Assessing the results of air quality management programs. *EM: Air and Waste Management Association's Magazine for Environmental Managers* pp. 8–15.

Hubbell B, Greenbaum D. 2014. Counterpoint: Moving from potential-outcomes thinking to doing — changing research planning to enable successful health outcomes research. Am J Epidemiol 180:1141–1144; doi:10.1093/aje/kwu266.

Katsouyanni K, Samet J, Anderson HR, Atkinson R, Le Tertre A, Medina S, et al. 2009. Air Pollution and Health: A European and North American Approach (APHENA). Research Report 142. Boston, MA:Health Effects Institute.

Kuebler J, van den Bergh H, Russell AG. 2001. Long-term trends of primary and secondary pollutant concentrations in Switzerland and their response to emission controls and economic changes. Atmos Environ 35:1351–1363; doi: 10.1016/S1352-2310(00)00401-5.

Lee, J, Son, J, Cho Y. 2007. Benefits of mitigated ambient air quality due to transportation control on childhood asthma hospitalization during the 2002 Summer Asian

Games in Busan, Korea. J Air Waste Manag Assoc 57:968–973; doi:10.3155/1047-3289.57.8.968.

Liao K-J, Tagaris E, Napelenok SL, Manomaiphiboon K, Woo J-H, Amar P, et al. 2008. Current and future linked responses of ozone and $PM_{2.5}$ to emission controls. Environ Sci Technol 42:4670–4675; doi:10.1021/es7028685.

Metzger KB, Tolbert PE, Klein M, Peel JL, Flanders WD, Todd K, et al. 2004. Ambient air pollution and cardiovascular emergency department visits. Epidemiology 15:46–56; doi:10.1097/01.EDE.0000101748.28283.97.

Morgenstern RD, Harrington W, Shih JS, Bell ML. 2012. Accountability Analysis of Title IV Phase 2 of the 1990 Clean Air Act Amendments. Research Report 168. Boston MA:Health Effects Institute.

Napelenok SL, Cohan DS, Hu Y, Russell AG. 2006. Decoupled direct 3D sensitivity analysis for particulate matter (DDM-3D/PM). Atmos Environ 40:6112–6121; doi:10.1016/j.atmosenv.2006.05.039.

Peel JL, Tolbert PE, Klein M, Metzger KB, Flanders WD, Todd K, et al. 2005. Ambient air pollution and respiratory emergency department visits. Epidemiology 16:164–174.

Peel JL, Klein M, Flanders WD, Mulholland JA, Tolbert PE. 2010. Impact of Improved Air Quality During the 1996 Summer Olympic Games in Atlanta on Multiple Cardiovascular and Respiratory Outcomes. Research Report 148. Boston, MA:Health Effects Institute.

Peters A, Breitner S, Cyrys J, Stölzel M, Pitz M, Wölke G, et al. 2009. The Influence of Improved Air Quality on Mortality Risks in Erfurt, Germany. Research Report 137. Boston, MA:Health Effects Institute.

Rao ST, Zurbenko IG, Neagu R, Porter PS, Ku JY, Henry RF. 1997. Space and time scales in ambient ozone data. Bull Am Meteorol Soc 78:2153–2166; doi:10.1175/1520-0477(1997)078<2153:satsia>2.0.co;2.

Rich DQ, Kipen HM, Huang W, Wang G, Wang Y, Zhu P, et al. 2012. Association between changes in air pollution levels during the Beijing Olympics and biomarkers of inflammation and thrombosis in healthy young adults. JAMA 307:2068–2078; doi:10.1001/jama.2012.3488.

Rich DQ. 2017. Accountability studies of air pollution and health effects: lessons learned and recommendations for future natural experiment opportunities. Environ Int 100:62–78; doi:10.1016/j.envint.2016.12.019.

Sarnat SE, Klein M, Sarnat JA, Flanders WD, Waller LA, Mulholland JA, et al. 2010. An examination of exposure measurement error from air pollutant spatial variability in time-series studies. J Expo Sci Environ Epidemiol 20:135–146; doi:10.1038/jes.2009.10.

Sarnat SE, Sarnat JA, Mulholland J, Isakov V, Özkaynak H, Chang HH, et al. 2013. Application of alternative spatiotemporal metrics of ambient air pollution exposure in a time-series epidemiological study in Atlanta. J Expo Sci Environ Epidemiol 23:593–605; doi:10.1038/jes.2013.41.

Solomon PA, Chameides W, Weber R, Middlebrook A, Kiang CS, Russell AG, et al. 2003. Overview of the 1999 Atlanta Supersite Project. J Geophys Res: Atmospheres 108; doi:10.1029/2001JD001458.

Strickland MJ, Darrow LA, Mulholland JA, Klein M, Flanders WD, Winquist A, et al. 2011. Implications of different approaches for characterizing ambient air pollutant concentrations within the urban airshed for time-series studies and health benefits analyses. Environ Health 10:36; doi:10.1186/1476-069X-10-36.

Strickland MJ, Gass KM, Goldman GT, Mulholland JA. 2015. Effects of ambient air pollution measurement error on health effect estimates in time-series studies: a simulation-based analysis. J Expo Sci Environ Epidemiol 25:160–166; doi:10.1038/jes.2013.16.

Tolbert PE, Klein M, Metzger KB, Peel J, Flanders WD, Todd K, et al. 2000. Interim results of the study of particulates and health in Atlanta (SOPHIA). J Expo Anal Environ Epidemiol 10:446–460; doi:10.1038/sj.jea.7500106.

Tolbert PE, Klein M, Peel JL, Sarnat SE, Sarnat JA. 2007. Multipollutant modeling issues in a study of ambient air quality and emergency department visits in Atlanta. J Expo Sci Environ Epidemiol 17:S29–S35; doi:10.1038/sj.jes.7500625.

Tonne C, Beevers S, Armstrong B, Kelly F, Wilkinson P. 2008. Air pollution and mortality benefits of the London Congestion Charge: spatial and socioeconomic inequalities. Occup Environ Med 65:620–627; doi:10.1136/oem.2007.036533.

United States Census Bureau. 2013. Decennial Census of Population and Housing. Available: *www.census.gov/programs-surveys/decennial-census/decade.html* [accessed 1 May 2017].

U.S. EPA (United States Environmental Protection Agency). 1992. Regulatory Impact Analysis of the Final Acid Rain Implementation Regulations. Office of Atmospheric and Indoor Air Programs, Acid Rain Division. Available: *https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P100RD3G.txt* [accessed 27 February 2018].

U.S. EPA (United States Environmental Protection Agency). 1997. The Benefits and Costs of the Clean Air

Act, 1970 to 1990. Available: *www.epa.gov/clean-air-act-overview/benefits-and-costs-clean-air-act-1970-1990-retrospective-study* [accessed 12 January 2018].

U.S. EPA (United States Environmental Protection Agency). 1998. Regulatory Impact Analysis for the NO_x SIP Call, FIP, and Section 126 Petitions, Volume 2: Health and Welfare Benefits. EPA-452/R-98-003. Available: *https://nepis.epa.gov/Exe/ZyPDF.cgi/P100DSAQ.PDF?Dockey=P100DSAQ.PDF* [accessed 7 August 2017].

U.S. EPA (United States Environmental Protection Agency). 1999a. Regulatory Impact Analysis — Control of Air Pollution from New Motor Vehicles: Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Control Requirements. Available: *www.epa.gov/regulations-emissions-vehicles-and-engines/final-rule-control-air-pollution-new-motor-vehicles-tier* [accessed 27 February 2018].

U.S. EPA (United States Environmental Protection Agency). 1999b. The Benefits and Costs of the Clean Air Act: 1990 to 2010. EPA-410-R-99-001. Available: *www.epa.gov/clean-air-act-overview/benefits-and-costs-clean-air-act-1990-2010-first-prospective-study* [accessed 12 January 2018].

U.S. EPA (United States Environmental Protection Agency). 2000. Regulatory Impact Analysis: Heavy-Duty Engine and Vehicle Standards and Highway Diesel Fuel Sulfur Control Requirements. Available: *www.epa.gov/regulations-emissions-vehicles-and-engines/final-rule-control-air-pollution-new-motor-vehicles-heavy* [accessed 27 February 2018].

U.S. EPA (United States Environmental Protection Agency). 2005. Regulatory Impact Analysis for the Final Clean Air Interstate Rule. EPA-452/R-05-002. Available: *https://archive.epa.gov/airmarkets/programs/cair/web/pdf/finaltech08.pdf* [accessed 7 August 2017].

U.S. EPA (United States Environmental Protection Agency). 2011. The Benefits and Costs of the Clean Air Act from 1990 to 2020. Final Report. Revision A. Office of Air and Radiation. Available: *www.epa.gov/sites/production/files/2015-07/documents/fullreport_rev_a.pdf* [accessed 27 February 2018].

U.S. EPA (United States Environmental Protection Agency). 2012. MOtor Vehicle Emissions Simulator (MOVES). Available: *www.epa.gov/otaq/models/moves* [accessed September 28, 2016].

van Erp AMM, O'Keefe R, Cohen AJ, Warren J. 2008. Evaluating the effectiveness of air quality interventions. J Toxicol Environ Health, Part A 71:583–587; doi:10.1080/15287390801997708.

Yang Z, Muncrief R, Bandivadekar A. 2017. Global Baseline Assessment of Compliance and Enforcement Programs for Vehicle Emissions and Energy Efficiency. The International Council on Clean Transportation (ICCT). Available: *www.theicct.org/sites/default/files/publications/PV-C%26E-global-baseline-assessment_ICCT-report_14112017_vF.pdf* [accessed 27 February 2018].

Zhang J, Ouyang Z, Miao H, Wang X. 2011. Ambient air quality trends and driving factor analysis in Beijing, 1983–2007. J Environ Sci 23:2019–2028; doi:10.1016/s1001-0742(10)60667-5.

Zhang J, Zhu T, Kipen H, Wang G, Huang W, Rich D, et al. 2013. Cardiorespiratory Biomarker Responses in Healthy Young Adults to Drastic Air Quality Changes Surrounding the 2008 Beijing Olympics. Research Report 174. Boston MA:Health Effects Institute.

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.

ABBREVIATIONS AND OTHER TERMS

ANAA	Atlanta nonattainment area
ARP	Acid Rain Program
C	Clean Air Act Amendments
CAIR	Clean Air Interstate Rule
CEM	continuous emissions monitoring
CHF	congestive heart failure
CMAQ	Community Multiscale Air Quality
CO	carbon monoxide
CTM	chemical transport model
CVD	cardiovascular disease
DSP	Heavy-Duty Diesel Rule
EC	elemental carbon
ED	emergency department
EGU	electricity generating unit
EY	emission year
EPD	Environmental Protection Division
GRAQC	Georgia Rules for Air Quality Control
$GRAQC_{bbb}$	Gasoline Marketing Rule
$GRAQC_{jjj}$	NO_x Emissions from Electric Utility Steam Generating Units
$GRAQC_{sss}$	Multipollutant Control for Electricity Utility Steam Generating Units
$GRAQC_{yy}$	Emissions of Nitrogen Oxides from Major Sources
GSP	Tier 2 Gasoline Program
IM	inspection and maintenance
IND	industrial emissions
ICD-9	International Classification of Diseases 9th Revision
MDA8h	maximum daily 8-hour
MOB	on-road mobile
MOVES	U.S. EPA MOtor Vehicle Emissions Simulator
NAA	nonattainment area
NAAQS	national ambient air quality standards
NBP	NO_x Budget Trading Program
NH_4^+	ammonium
NMB	normalized mean bias
NME	normalized mean error
NMHCs	nonmethane hydrocarbons
NO_2	nitrogen dioxide
NO_3^-	nitrate
NO_x	oxides of nitrogen
O_3	ozone
OC	organic carbon
PM	particulate matter
$PM_{2.5}$	particulate matter ≤2.5 µm in aerodynamic diameter
PM_{10}	particulate matter ≤10 µm in aerodynamic diameter
PS*	emissions-independent atmospheric photochemical state
R-LINE	research LINE-source dispersion model for near-surface releases
RD	respiratory disease
REG	regional, referring to EGU emissions in Alabama, Georgia, Mississippi, North Carolina, South Carolina, and Tennessee minus those in the ANAA
RFA	Request for Applications
RH	relative humidity
RIA	Regulatory Impact Assessment
SEARCH	SouthEastern Aerosol Research and Characterization
SD	standard deviation
SIP	State Implementation Plan
SO_2	sulfur dioxide
SO_4^{2-}	sulfate
SOA	secondary organic aerosol
SOPHIA	Study of Particles and Health in Atlanta
SO_x	oxides of sulfur
STM	short-term meteorology
U.S. EPA	U.S. Environmental Protection Agency
VMT	vehicle miles traveled
VOCs	volatile organic compounds

RELATED HEI PUBLICATIONS: HEALTH OUTCOMES STUDIES AND STUDIES ASSESSING LONG-TERM HEALTH EFFECTS FROM EXPOSURE TO AIR POLLUTION

Number	Title	Principal Investigator	Date
Research Reports			
190	The Effects of Policy-Driven Air Quality Improvements on Children's Respiratory Health	F. Gilliland	2017
187	Causal Inference Methods for Estimating Long-Term Health Effects of Air Quality Regulations	C.M. Zigler	2016
178	National Particle Component Toxicity (NPACT) Initiative Report on Cardiovascular Effects	S. Vedal	2013
177	National Particle Component Toxicity (NPACT) Initiative: Integrated Epidemiologic and Toxicologic Studies of the Health Effects of Particulate Matter Components	M. Lippmann	2013
176	Effect of Air Pollution Control on Mortality and Hospital Admissions in Ireland	D.W. Dockery	2013
174	Cardiorespiratory Biomarker Responses in Healthy Young Adults to Drastic Air Quality Changes Surrounding the 2008 Beijing Olympics	J. Zhang	2013
170	Impact of the 1990 Hong Kong Legislation for Restriction on Sulfur Content in Fuel	C.-M. Wong	2012
168	Accountability Analysis of Title IV Phase 2 of the 1990 Clean Air Act Amendments	R.D. Morgenstern	2012
163	The London Low Emission Zone Baseline Study	F. Kelly	2011
162	Assessing the Impact of a Wood Stove Replacement Program on Air Quality and Children's Health	C.W. Noonan	2011
155	The Impact of the Congestion Charging Scheme on Air Quality in London	F. Kelly	2011
148	Impact of Improved Air Quality During the 1996 Summer Olympic Games in Atlanta on Multiple Cardiovascular and Respiratory Outcomes	J.L. Peel	2010
140	Extended Follow-Up and Spatial Analysis of the American Cancer Society Study Linking Particulate Air Pollution and Mortality	D. Krewski	2009
139	Effects of Long-Term Exposure to Traffic-Related Air Pollution on Respiratory and Cardiovascular Mortality in the Netherlands: The NLCS-AIR Study	B. Brunekreef	2009
137	The Influence of Improved Air Quality on Mortality Risks in Erfurt, Germany	A. Peters	2009
HEI Communications			
15	Proceedings of an HEI Workshop on Further Research to Assess the Health Impacts of Actions Taken to Improve Air Quality		2010
14	HEI's Research Program on the Impact of Actions to Improve Air Quality: 2009 Interim Evaluation and Future Directions		2009
11	Assessing the Health Impact of Air Quality Regulations: Concepts and 2003 Methods for Accountability Research		2003

Copies of these reports can be obtained from HEI; pdf's are available for free downloading at www.healtheffects.org/publications.

HEI BOARD, COMMITTEES, and STAFF

HEI BOARD, COMMITTEES, and STAFF

Officers and Staff

Daniel S. Greenbaum *President*
Robert M. O'Keefe *Vice President*
Rashid Shaikh *Director of Science*
Jacqueline C. Rutledge *Director of Finance and Administration*
April Rieger *Corporate Secretary*

Hanna Boogaard *Consulting Senior Scientist*
Aaron J. Cohen *Consulting Principal Scientist*
Maria G. Costantini *Principal Scientist*
Philip J. DeMarco *Compliance Manager*
Hope Green *Editorial Project Manager*
Kathryn Liziewski *Research Assistant*
Allison P. Patton *Staff Scientist*
Hilary Selby Polk *Managing Editor*
Lindy Raso *Executive Assistant*
Anna S. Rosofsky *Staff Scientist*
Robert A. Shavers *Operations Manager*
Annemoon M.M. van Erp *Managing Scientist*
Donna J. Vorhees *Director of Energy Research*
Katherine Walker *Principal Scientist*

HEALTH
EFFECTS
INSTITUTE

75 Federal Street, Suite 1400
Boston, MA 02110, USA
+1-617-488-2300
www.healtheffects.org

RESEARCH
REPORT

Number 195
April 2018