Cumulative Impact and Equity Objectives in Energy Systems Modeling and Policy

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Engineering & Public Policy

Erin Noel Mayfield

B.S., Environmental Science, Rutgers University

M.S., Environmental Engineering, Johns Hopkins University

Carnegie Mellon University

Pittsburgh, Pennsylvania

May 2019

Acknowledgements

I am grateful for the funding sources that made this thesis possible. This work was funded through the Center for Air, Climate, and Energy Solutions and developed under Assistance Agreement No. RD83587301 awarded by the U.S. Environmental Protection Agency (EPA). This work has not been formally reviewed by the EPA. Additional funding was provided by the Scott Institute for Energy Innovation and the National Oceanic and Atmospheric Administration (Grant No. NA14OAR4310135). The views expressed in this document are solely those of the author and do not necessarily reflect those of the funding sources.

I owe my sincere gratitude to the many mentors, friends, and family who have contributed to my academic and personal growth over the years. First, thank you to my thesis advisors and co-chairs, Jared Cohon and Allen Robinson. I am truly grateful for your commitment to thoughtful, thorough, and meaningful research, which has been influential, and more so for your consistent kindness, and professional and personal encouragement.

I would like to thank my committee members – Nicholas Muller, Inês Azevedo, and Donna Vorhees – for their diversity of thought and always challenging my ideas for the sake of improvement. I feel fortunate to have served as a teaching assistant for Deanna Matthews, who has been a committed mentor. I also would like to thank the many other faculty members that have supported my research.

I appreciate the always helpful and warm staff, including Debbie Kuntz, Deb Scappatura, Diana Rotando, Elisabeth Udyawar, Ginny Delaney, Kate Tucker, Kim Martin, Mi Kim, Patti Steranchak, and Vicki Finney, with a special thanks to Adam Loucks for regularly greeting me with meows.

I would like to acknowledge the Doctoral Women of Engineering and Public Policy, including those that continue to lead and foster the organization – Erezi Ogbo, Vanya Britto, and Nicole Racine. Among my fellow EPP students, who have made this a great community to be part of, I would especially like to thank Aviva Loew, Barry Dewitt, Cristobal de la Maza, Chukky Udeani, Daniel Gingerich, Evan Sherwin, Jeremy Keen, Leslie Abrahams, Mike Ford, Mike Roth, Nichole Hanus, Rebecca Ciez, Shuchi Talati, Shayak Sengupta, Steph Seki, and Travis Carless.

Thank you to my dearest Pittsburgh friends. We have shared much together, but perhaps I can distill the essence of our experiences: figuring life out with Lynn Kaack, shocking story time and laughing with Octavio Mesner, endurance pursuits with Brian Sergi, life talks over bubble tea with Sinnott Murphy, marathon café "work" sessions with Tobi Adekanye, and truth-seeking conversations during geriatric lap swims with Julia Chelen.

Lastly, I owe the deepest thank you to my family for their unwavering support and love. I am grateful for my sister, Crystal – you always remind me of our roots – and my brother-in-law, Mike. I am also thankful for my nephews, Anthony, Vincent, and Dominic, to whom I can attribute my greatest moments of simple joy and that serve as the nearest inspiration for preserving the future. I would be remiss not to acknowledge Leroy Brown (and the long-distance meow phone calls) and all of my late cat and dog siblings over the years, including Moose, Petal, Flick, Gus, Pepper, Fritts, Cuddles, and Hobbit. I am blessed to have spent the many long days of childhood with my Grandma, the gentlest and most compassionate person that I have known. *You are an influential voice in my head. Thank you for everything.* And finally, I am grateful for my Mom and Dad and their unconditional love and patience. I can trace the foundations of my environmental philosophies to the many weekends that you shuttled us into the woods. It is through your example of using your brain and much more so your heart that I learned to enter and exist in the world. Thank you.

Abstract

Energy system development is driven by the complexity inherent in physical systems and the influence of a myriad of diverse, interacting stakeholders with heterogeneous preferences. Transforming energy systems entails balancing multiple and often conflicting societal objectives. This thesis presents new modeling approaches for energy systems planning and policy evaluation, with an emphasis on cumulative impacts, equity, and system heterogeneity. The application domain of this thesis is the U.S. natural gas system, although the analytical approaches and insight of this research are intended to extend to the broader domestic and global energy system.

Chapter 2 adopts a traditional economic efficiency optimization approach, coupled with methane emissions and abatement cost simulations reflecting system heterogeneity, to evaluate and design system-wide and superemitter policies related to methane abatement in the U.S. transmission and storage system. We find that most emissions, given the existing suite of technologies, have the potential to be abated. We also demonstrate that there are high societal benefits from abatement policies, and minimal (if any) private costs under standard and tax instruments. Superemitter policies, which target the highest emitting facilities, may reduce the private cost burden and achieve high emission reductions, especially if emissions across facilities are highly skewed. However, detection across all facilities is necessary regardless of the policy option, and there are nontrivial societal benefits resulting from abatement of relatively low-emitting sources.

Chapters 3 aims to develop and demonstrate a data-driven approach for characterizing systems-level cumulative impacts of current energy systems. Specifically, we comprehensively assess the spatially-and temporally-resolved air, climate, and employment impacts from extraction to end use and over the life of natural gas plays in the Appalachian basin from 2004 to 2016. Our approach highlights the attribution of impacts across the supply chain, the tradeoffs between near- and long-term impacts, and the evolution and accumulation of impacts over time with changing regulation, natural gas activity, and technological and operational efficiencies and practices. We show that short-lived air quality and employment impacts track with the boom-and-bust cycle, while climate impacts persist for generations well beyond the period of natural gas activity. We also find that employment effects are spatially concentrated in rural areas with thin labor markets where development is occurring, and more than half of cumulative premature mortality is within source emissions states. We show that most premature mortality is associated with end uses, while upstream and midstream segments also account for a substantial portion of impacts. With respect to climate change impacts, the magnitude of methane emissions across the supply chain produces temperature impacts nearly equivalent to that of carbon dioxide over a 30-year time horizon, but over longer integration periods, the warming impact from carbon dioxide dominates. We estimate a tax on production of $2 per thousand

cubic foot (+172%/-76%) would compensate for cumulative climate and air quality externalities across the supply chain.

In Chapter 4, we develop a multiobjective optimization model incorporating cumulative impact objectives to facilitate future energy system planning. We develop natural gas system pathways by optimizing impacts with respect to sequential natural gas decisions regarding the timing and location of infrastructure and activity from extraction to end use. Environmental and employment objectives are conflicting if we follow a natural gas pathway consistent with the status quo; however, a collection siting, emissions abatement, and renewable integration policies may collectively resolve and reverse these conflicts.

In Chapter 5, we develop and demonstrate an approach for evaluating the equity state of an energy system. We apply variants of standard methods and present new methods and metrics to quantify spatial, temporal, and distributional equity, leveraging impact estimates of the shale gas boom in the Appalachian basin from Chapter 3. We find that there are high temporal and spatial inequities with respect to cumulative air and employment impacts, and that spatial inequities are constant over time reflecting largely fixed infrastructure and consumption patterns. We also present indicators of temporal climate inequities, estimating that long-term global temperature impacts are 100 times that of near-term impacts. With respect to distributional equity of air quality impacts, we do not observe a disparity in mortality rates across subpopulations on the basis of income and poverty; however, there is a trend of increasing income corresponding to decreasing damages, which demonstrates the higher health burden of lower income communities. With respect to distributional equity of labor markets, we find statistically significant declines in the income disparity and poverty rates in producing counties. Pairwise comparisons of impacts reveal that changes in air and climate impacts are sensitive to changes in employment impacts.

In Chapter 6, we develop future natural gas system pathways that optimize for the multiple dimensions of equity. We expand upon the multiobjective optimization model developed in Chapter 4, deriving objectives that instill different normative concepts of spatial, temporal, and distributional equity that apply to air, climate, and employment impacts. We find that there are inherent conflicts between different equity dimensions, as well as, between equity and cumulative impact objectives in a fossil-fuel dominated energy system. However, low-carbon technologies have the potential to reduce inequities.

Table of Contents

List of Figures

List of Tables

1 Introduction

Requisite to transforming an energy system is addressing the multi-dimensional needs of today and the future. This necessitates both a systems-level thinking to broadly scaffold the structure and elucidate the dynamics of societal objectives, as well as, the nuanced decomposition of objectives to enable the evaluation and design of policies targeting specific system components. This thesis focuses on multiple objectives that influence decision-making processes and are often the subject of public discourse and concern, including climate change, employment, and air pollution impacts and equity. Specifically, this thesis develops and demonstrates new approaches for energy systems modeling and policy evaluation, with an emphasis on systems heterogeneity, cumulative impacts, and equity. In the following sections, we provide a brief overviews of energy system transitions, energy systems analysis and optimization, U.S. environmental policy research and evaluation, and the application domain, the U.S. natural gas system. We additionally describe the overarching objectives of this research and outline the thesis chapters.

1.1 Motivation

1.1.1 U.S. energy system transitions

Climate change has resulted in impacts to natural and human systems, and avoidance of serious impacts requires rapid and deep reductions in greenhouse gas (GHG) emissions and associated policies.[1–3] The Paris Agreement, adopted in 2015, aims to reduce GHG emissions to a level consistent with limiting the average global temperature to well below 2°C above pre-industrial levels.[2,4] Studies have shown that a deeply decarbonized U.S. energy system can provide equivalent energy services as the status quo; however, deep decarbonization entails unprecedented and transformational changes in the energy system, such as drastically increasing efficiency of end uses, decreasing the carbon intensity of electricity, and switching end uses from direct combustion of fossil fuels to electricity.[5–7]

Beyond technoeconomic challenges of decarbonization, accelerated transitions also depend upon widespread social acceptance and balancing potentially countervailing societal needs and objectives.[8,9]

Given the coupling of air and climate impacts related to fossil fuel production and use, transitioning to low-carbon energy also contributes to reducing near-term air pollution impacts.[10] Labor markets are also inherently linked to the transition, with a redistribution of jobs away from carbon-intensive energy sources, such as coal, to green infrastructure development.[6,11] There are also distributional effects and equity implications of climate change and mitigation.[12,13]

1.1.2 Energy systems optimization

Technoeconomic optimization techniques are often used to evaluate energy systems and technologies, including their costs and environmental effects. There are a wide array of energy systems models that vary in complexity and sectoral, spatial, and temporal scope, but few models currently focus on energy system transitions and deep decarbonization.[5] We identify common elements of several existing energy optimization models that limit their application to future energy system contexts. Many architectures use illustrative systems and networks that are not reflective of critical real world dynamics.[8] Understanding the costs of energy transitions is fundamental, and thus, models almost exclusively employ monetized objectives, adopt an economic efficiency or utilitarian framing, and do not consider impacts.[14–16] Dominating are models that focus exclusively on electric power systems, ignoring the broader, interconnected energy sector and social system.[17] In addition, energy system optimization models often do not consider long time horizons and multiperiod objectives.

1.1.3 U.S. environmental policy

Conventional approaches, such as benefit-cost analysis (BCA), that are widely used in U.S. environmental policy research, evaluation, and design have many limitations that partially parallel those described for energy systems optimization research, especially in relation to the treatment of long time horizons and focus on efficiency criteria and monetization.[18–20] BCA has been described from a practical perspective as a consistent, systematic accounting framework for comparing benefits and costs in social decisions, and from a theoretical perspective as an optimizing tool that maximizes social welfare.[21] While costs can be reasonably measured, benefits may be difficult to quantify and monetize, and even if benefits can be monetized, the implied exchange rate between benefits and costs may not be reflective of preferences.[22] Underlying most federal environmental policy making is the Kaldor-Hicks (KH) efficiency criterion which maintains that a change from the current social state should be undertaken if the gainers compensate losers such that everyone would be better off; the KH criterion satisfies the Pareto criterion in which a change in state makes at least one individual better off without making any other worse off.[23] Inevitably, regulatory policies involve winners and losers, even when aggregate benefits exceed costs,[21] and the KH criterion does not require that compensation actually occurs, effectively ignoring the distributional aspects of a change.[24] Factors such as equity within and across generations are often ignored in U.S. environmental policy, despite being fundamental in many decision contexts[25], including energy system transitions. Another dominating construct in environmental policy evaluation, that has implications for long-term decisions and temporal equity, is the use of discounting, which renders benefits and costs that occur in different time periods comparable.[26–28]

With respect to equity in U.S. environmental policy making, there are no federal environmental justice regulations; however, Executive Order (E.O.) 12898, *Federal Actions to Address Environmental Justice in Minority Populations and Low-Income Populations*, calls for agencies to integrate environmental justice into its mission "by identifying and addressing, as appropriate, disproportionately high and adverse human health or environmental effects of its programs, policies, and activities on minority populations and low-income populations."[29,30] While the U.S. Environmental Protection Agency provides technical guidance on conducting environmental justice analyses, equity is treated as a subsidiary or supplemental, rather than central, criterion in regulatory decision making.

While economic efficiency is a dominate framing in existing technoeconomic policy research and is part of U.S. writ, the approaches developed and demonstrated in this thesis are largely premised on the observation that conventional methods and assumptions of policy evaluation and design have largely not resulted in nor meaningfully iterated towards processes or outcomes that address inequities and the long-term effects of climate change.

1.1.4 U.S. natural gas system

The application domain of this thesis is the U.S. natural gas system, although the analytical approaches and insight of this research are intended to extend to the broader domestic and global energy system. Here, we provide some background on the technological, economic, and policy evolution of the natural gas system.

Rapid increases in U.S. natural gas production, resulting from advancements in horizontal drilling and hydraulic fracturing, have had reverberations on world energy markets and the domestic energy outlook. The U.S. has been the largest natural gas consumer and producer over the past decade, comprising 20% of the world market, and the domestic shale gas market has contributed to price volatility and a shift in global flows of natural gas.[31] Increases in domestic supply and reserves have contributed to the displacement of coal[32], but have also potentially delayed investment, development and deployment of renewables.[33] In much of this thesis, we focus on the Appalachian basin, the largest natural gas basin in the U.S. with respect to both reserves and production over the past decade.[34]

Recent studies indicate that the climate benefits of natural gas relative to other fossil fuels are countered, perhaps significantly, by emissions of methane, the primary constituent of natural gas and a potent greenhouse gas.[35,36] Despite advances in emission controls, methane emissions from the U.S. natural gas system are already substantial and may increase with increasing production and without regulatory intervention. Targeting emission reductions from new and modified oil and natural gas facilities, the Environmental Protection Agency (EPA) finalized New Source Performance Standards (NSPS) under the Clean Air Act in 2016 during the Obama Administration; the Trump Administration, which has advocated for the revival of the coal industry, has since proposed a rollback of these rules.[37–41]

1.2 Research objectives

The intent of this thesis is to develop and apply multiobjective, systems-level, and data-driven approaches for energy system modeling and policy evaluation that seek to address limitations of existing evaluative tools, with an emphasis on cumulative impacts, equity, and systems heterogeneity.

The first research objective is to present an approach for characterizing and quantifying spatially- and temporally-resolved cumulative environmental and socioeconomic impacts of energy systems across supply chains and development cycles. Many existing policy research and evaluation tools, such as benefit-cost analysis and life cycle assessment, and corresponding applications are not as comprehensive in scope.

The second research objective is to develop and demonstrate an approach for characterizing and quantifying the multi-dimensional equity state of existing energy systems. The systems-level spatial, temporal, and distributional effects and equity of energy systems – i.e., how benefits and costs are and should be distributed spatially, temporally, or among sub-populations on the basis of demographics – are still largely unexplored.[29,42]

Building upon the first and second objectives, the third research objective is to construct a multiobjective energy system optimization model that incorporates cumulative impact and equity objectives. The purpose of this model is to facilitate future energy system planning through the generation future development pathways, with respect to the timing, location, and magnitude of energy system activity and low-carbon interventions. The goal is to develop a modeling architecture that is distinct from many existing energy systems optimization models because of the unique cumulative impact and equity objectives, spatial resolution, long time horizons, air and climate impact (rather than emission and monetary) measures leveraging reduced complexity models, and parameterization and structure reflecting a real (rather than illustrative) energy system.

The final research objective, which intersects with the previously described objectives, is to draw upon the quantitative measures and insight derived from the systems-level quantitative modeling to design new policy and planning mechanisms to address system heterogeneity, cumulative impacts, and equity, which are largely unaccounted for in U.S. environmental policy and energy systems planning.

1.3 Outline

The following outlines the chapters of this thesis and maps them to the previously described research objectives.

In Chapter 2, we evaluate and design optimal system-wide and superemitter methane abatement strategies and policies for the U.S. natural gas system, with a focus on the transmission and storage segment We formulate economic efficiency optimization models from private and societal perspectives, coupled with methane emissions and abatement cost simulations reflecting system heterogeneity, to assess policies across four broad dimensions: private and social costs and benefits, detection and abatement technologies, emissions reduction potential, and policy instruments.

Whereas Chapter 2 develops an evaluative framework that in part adopts a traditional economic efficiency framing and targets specific policies for a segment of the natural gas supply chain, we take a more comprehensive approach and consider other societal objectives in Chapters 3 and 4 to facilitate long-term decision making and comparisons of natural gas to low-carbon energy technologies. In Chapter 3, we develop an approach for evaluating and characterizing cumulative impacts from extraction to end use and over the life of natural gas plays. Specifically, we demonstrate this approach through a retrospective case study of the shale gas boom in the Appalachian basin from 2004 to 2016, and we model data-driven, spatially- and temporally-resolved air quality, climate change, and employment impact estimates.

While Chapter 3 provides a retrospective assessment of impacts, the goal of Chapter 4 is to evaluate how natural gas systems can develop in the future if cumulative air, climate, and employment impact objectives are incorporated into decision making. We formulate a multiobjective optimization model, in which we develop future natural gas system pathways by optimizing impacts with respect to sequential natural gas decisions regarding the timing and location of infrastructure and activity from extraction to end use.

We also develop and demonstrate approaches for evaluating the equity of energy systems in Chapters 5 and 6. In Chapter 5, we take a systematic, but exploratory, approach to quantify and characterize the multi-dimensional equity state of energy systems, using established economic methods in addition to developing supporting systems-level equity metrics. We focus on establishing the state of spatial, temporal, and distributional equity as it relates to air, climate, and employment impacts during the shale gas boom in Appalachia.

In Chapter 6, we develop future natural gas system pathways that optimize for the multiple dimensions of equity. We expand upon the multiobjective energy system optimization model developed in Chapter 4, deriving objectives that instill different normative concepts of equity that apply to air, climate, and employment impacts.

In Chapter 7, we summarize key findings and policy implications, as well as, outline a proposal for future work.

2 System-wide and Superemitter Policy Options for the Abatement of Methane Emissions from the U.S. Natural Gas System

This paper assesses tradeoffs between system-wide and superemitter policy options for reducing methane emissions from compressor stations in the U.S. transmission and storage system. Leveraging recently collected national emissions and activity datasets, we developed a new processed-based emissions model implemented in a Monte Carlo simulation framework to estimate emissions for each component and facility in the system. We find that approximately 83% of emissions, given the existing suite of technologies, have the potential to be abated, with only a few emission categories comprising a majority of emissions. We then formulate optimization models to determine optimal abatement strategies. Most emissions across the system (approximately 80%) are efficient to abate, resulting in net benefits ranging from \$160M to \$1.2B annually across the system. The private cost burden is minimal under standard and tax instruments, and if firms market the abated natural gas, private net benefits may be generated. Superemitter policies, namely those that target the highest emitting facilities, may reduce the private cost burden and achieve high emission reductions, especially if emissions across facilities are highly skewed. However, detection across all facilities is necessary regardless of the policy option and there are nontrivial net benefits resulting from abatement of relatively low-emitting sources.

This paper has been published as: Mayfield, E. N., Robinson, A. L. & Cohon, J. L. System-wide and Superemitter Policy Options for the Abatement of Methane Emissions from the U.S. Natural Gas System. *Environ. Sci. Technol.* (2017). doi:10.1021/acs.est.6b05052

2.1 Introduction

Recent studies indicate that the climate benefits of natural gas relative to other fossil fuels are countered, perhaps significantly, by emissions of methane, the primary constituent of natural gas and a potent greenhouse gas.[35,43] Methane emissions from the natural gas sector alone account for approximately 24% of U.S. anthropogenic methane emissions.[44] Despite advances in emission controls, methane emissions from the U.S. natural gas system are already substantial and may increase with increasing production and without regulatory intervention.

In an effort to reduce emissions across the natural gas supply chain, the White House Climate Action Plan, which includes a strategy for reducing methane emissions, was released in March 2014; this was followed by an announcement that President Obama will use his executive authority to issue regulations with the intent of reducing emissions from the oil and natural gas sector by 40-45 percent from 2012 levels by 2025.[45,46] Targeting emission reductions from new and modified oil and natural gas facilities, the Environmental Protection Agency (EPA) finalized New Source Performance Standards (NSPS) under the Clean Air Act in June 2016.[37–41] The EPA also proposed a new voluntary emissions reduction program, the Natural Gas STAR Methane Challenge, in July 2015.[47]

Emissions inventories and estimates are often based on relatively small datasets that may not be representative of the current natural gas system.[48] There have been substantial recent efforts to collect additional methane emissions data and refine activity counts to better understand the current and changing fleet of U.S. natural gas facilities. These studies have revealed that, for many source categories and across the natural gas supply chain, methane emissions distributions are highly skewed – a few *superemitters* are responsible for a majority of emissions.[48–53]

The heavy-tailed distributions of observed component- and facility-level emissions suggest that abating emissions from the highest emitting components or facilities has the potential to be a relatively efficient strategy for reducing emissions. While potentially efficient from an abatement perspective, lacking *a priori* knowledge of the location of high-emitting components or facilities, detection of emissions may be challenging and entail nontrivial costs. Previous studies that evaluated abatement options and costs generally did not account for heterogeneity in emissions and uncertainty in other key input parameters.[40,54,55]

This paper assesses *system-wide* and *superemitter* policy options for reducing methane emissions from the U.S. natural gas system. For the purposes of this analysis, system-wide and superemitter policies are broadly differentiated by their regulatory scope; the former includes detection and abatement of methane emissions at all facilities across the U.S. natural gas system, whereas the latter entails detection at all facilities and abatement at only a targeted subset of facilities with the highest absolute emissions or emissions rate. Further distinctions between these policy options are developed and discussed in the succeeding sections and summarized in Table 1.

Table 1. Description of system-wide and superemitter policy options.

	System-wide Policy		**Superemitter Policy**	
	Unconstrained	**Facility-Level Percent Emissions Reduction Target**	**Absolute Annual Emissions Threshold**	**Proportional Loss Rate Threshold**
Scope of detection	All facilities	All facilities	All facilities	All facilities
Scope of potential abatement	All facilities	All facilities	Facilities w/ emissions above threshold (50^{th}/90^{th} percentile)	Facilities with loss rate above threshold (50^{th}/90^{th} percentile)
Regulated unit	Component	Facility	Facility and component	Facility and component
Policy instrument	Component-level standard OR Tax	Facility-level percent emissions reduction target (10/50/75%)	Component-level standard for subset of facilities	Component-level standard for subset of facilities
Emissions reduction criterion	Maximize net benefits OR Maximum achievable reduction	Minimize private costs OR Maximum achievable reduction	Maximize net benefits OR Maximum achievable reduction	Maximize net benefits OR Maximum achievable reduction

2.2 Methods

We developed a modeling framework to analyze system-wide and superemitter policy tradeoffs across four broad dimensions: i) private and social costs and benefits, ii) detection and abatement technologies, iii) emissions reduction potential, and iv) policy instruments (e.g., standards versus taxes). SI Figure A-1 shows the six interacting sub-modules used in the assessment of policy options. The formulation, assumptions, and data for these sub-modules are introduced in the following sections and elaborated upon in SI Appendices B-G.

This study focuses on existing compressor stations in the transmission and storage (T&S) sector, as described in SI Appendix A. We first develop a baseline methane emissions model for each component and facility in the T&S sector. Then we simulate potential abatement measures, incorporating simulated emissions, as well as uncertainties in natural gas price forecasts, firm structure, abatement costs, and abatement efficacy. We also evaluate costs and sensitivity of existing and emerging detection methods and assess marginal benefits associated with emission reductions, employing the social cost of methane. Combining results from the preceding sub-modules, we formulate an optimization model, in which net social benefits are maximized or private costs are minimized, in order to evaluate abatement strategies for the T&S system. Finally, we summarize the sources of variability and uncertainty that we incorporate or omit in the model (SI Appendix G).

2.2.1 Baseline Emissions Simulation

We used a Monte Carlo simulation to model the baseline annual methane emissions for each component and facility in the U.S. T&S sector. SI Figure C-1 is a depiction of the types of components (e.g., rod-packing, isolation valves) or emission sources at transmission compressor stations and storage facilities. Rather than using emission estimates reported in the EPA Greenhouse Gas Inventory (GHGI), we employ emissions data from a study recently published by Zimmerle et al. (2015) because this dataset is relatively large and includes component-level estimates which are necessary for modeling heterogeneity in emissions.[49] We simulate annual baseline emissions for each emission category and component in the population, explicitly accounting for variability in emission factors and annual operating hours. Emission categories are delineated by component type and operating mode (e.g., reciprocating compressor rod-packing in an operating pressurized mode). Baseline emissions for each facility in the system are simulated by assigning emissions for each component to a facility. Each facility is described by a facility profile or inventory (i.e., component counts). These profiles are known, at least in part, for some facilities, and we simulate unknown profiles, accounting for correlations between counts and types of components. We additionally simulate a proportional loss rate for each facility, which is the quotient of annual emissions and throughput. Overall, we simulated 100 realizations of emissions from each component and facility in the entire U.S. T&S system. SI Appendix C provides greater detail regarding the mathematical formulation of the baseline emissions simulation.

2.2.2 Emission factors and operating hours

The input datasets are based on recent emission measurements by Subramanian et al. (2015) and the system-wide emissions model by Zimmerle et al. (2015). We develop annual emission factor distributions for 27 emission categories, based on emission measurements from approximately 22,000 components at 45 compressor stations across U.S facilities.[48] We fit parametric distributions that attempt to capture the data skewness and uncertainty; as summarized in SI Appendix G, recent studies review parametric and nonparametric approaches for modeling uncertainties.[36,56] Operating hour distributions are developed based on a database, compiled by Zimmerle et al. (2015), of hours reported for 24,000 components at 514 compressor stations. Using paired operating hours and emission factor measurements, we find that there is little correlation between operating hours and emission factors. Input distributions for each emission category are summarized in SI Appendix B.1.

2.2.3 Facility profiles and activity counts

Zimmerle et al. (2015) compiled facility profile data at 922 facilities in 2012, including data from facilities that annually report to the Greenhouse Gas Reporting Program (GHGRP) and the Federal Energy Regulatory Commission (FERC), as well as data collected by Subramanian et al. (2014). Although the precise number of T&S facilities is unknown, we assume that there are 1,758 compressor stations based on the central estimate in Zimmerle et al. (2015). Incorporating uncertainty in the number of facilities and activity counts is omitted because it would unnecessarily complicate the model and is unlikely to impact the overarching policy recommendations in a material way. The baseline emissions model provides a snapshot of existing components and facilities in 2012 and does not reflect expanding and changing natural gas infrastructure.

2.2.4 Proportional loss rates

Proportional loss rate is a measure of the proportion of facility-level throughput that is emitted. We fit probability distributions to paired horsepower and efficiency data for reciprocating and centrifugal compressors at 922 facilities compiled by Zimmerle et al. (2015). Combining these simulated results with the facility profile assignment model, emission estimates, and operating hours for compressors, we develop estimates of throughput and proportional loss rates for each facility.

2.2.5 Abatement Cost Simulation

Monte Carlo simulation was also used to model abatement costs for each component in the T&S sector. The mathematical formulation is provided in SI Appendix D. Abatement costs for each emission category and component are a function of the simulated baseline natural gas emissions, abatement efficacy for each category, annualized cost of abatement for each category and component, and natural gas prices (see SI Equation D-3). The following sections introduce input assumptions and data, and Appendix B.2 elaborates on these inputs.

2.2.6 Abatement efficacy

We model abatement measures (e.g., repair or replacement of a valve) for each emission category. We assume point estimates for efficacy derived from multiple sources, including supporting information for recent EPA rule-makings and abatement studies conducted by Carbon Limits, the EPA Natural Gas Star Program, and ICF International.[38–40,57–60]

The efficacy of an abatement measure is defined as the percentage reduction in emissions relative to baseline. While similar definitions of efficacy are implicitly adopted in the NSPS Regulatory Impact Analysis and an ICF economic analysis, there are limitations to operationalizing efficacy as a percentage reduction.[40,58] Percentage reduction does not account for potential correlation between baseline emissions and efficacy. In addition, efficacy conceivably decreases over time after an abatement action is conducted. To account for structural limitations of modeling efficacy and to demonstrate the sensitivity of efficacy on the conclusions drawn, we perform a parametric analysis, conservatively reducing the point estimates of efficacy by half.

2.2.7 Abatement costs

We estimate annualized abatement costs as the present value of the capital and labor costs (beyond the status quo) of an abatement measure over an assumed abatement interval (e.g., replacement interval), converted to equal annual payments. Annualized costs are a function of the capital recovery factor (CRF) and total abatement costs for each component type (see SI Equations D-1 and D-2). The CRF is a function of the discount factor and the abatement interval. We employ three discount rates, 3%, 5%, and 7%, which are consistent with those used in the 2015 NSPS Regulatory Impact Assessment and recommended in OMB Circular No. A-94.[40,61]

We derive costs and abatement intervals for each abatement measure based on multiple sources, including supporting information for recent EPA rule-makings and cost studies conducted by Carbon Limits, the Clearstone Group, the EPA Natural Gas Star Program, and ICF International.[37,38,40,54,55,57–59,62] Given that there is cost heterogeneity for each abatement measure, we fit simple distributions, reflective of reported cost ranges. Triangular distributions are employed when there is a clustering of central estimates, and uniform distributions are used if only a range is available.

There are several other factors and sources of uncertainty that we do not account for that may impact costs. There is locational heterogeneity in labor costs, but given that our model is not spatially explicit, we do not account for this. Emissions and abatement costs are also potentially correlated; for example, correlation may exist because of facility location (i.e., more remote areas may have higher emissions because maintenance is less frequent and more costly).

2.2.8 Natural gas prices and firm cost structure

Cost savings may be generated from reductions in losses of otherwise marketable natural gas. However, transmission firms often do not own the gas and fixed methane loss rates may be stipulated in long-term contracts; thus, firms may not realize cost savings, at least in the near-term. We do not have information to represent the distribution of firms that operate under different types of contracts; therefore, we perform a bounding analysis, modeling scenarios with and without savings. Given that the model is not spatially explicit, we calculate savings by broadly applying the Henry Hub spot prices reported in the EIA 2015 Annual Energy Outlook for the analysis year of 2020. To encapsulate the forecast range, we model three price scenarios – reference ($4.30/MMBTU), high oil and gas resource ($3.12/MMBTU), and high economic growth ($5.03/MMBTU); we convert to real 2014 USD per SCF, assuming an energy consumption-to-volume conversion factor of 1,015 BTU per SCF and reported inflation rates.[63–65]

2.2.9 Detection

The approach used to detect emissions is itself an important policy decision. For all of the modeled policies, we consider bottom-up or component-level quantification of emissions, rather than top-down, facility-level measurements. We further assume that detection is conducted at all facilities, rather than a subset of facilities. Given that some facilities already are performing detection in accordance with GHGRP, our detection cost estimate, which is based on full rather than incremental costs, may be overestimated.

We model two common detection methods that comply with NSPS requirement, optical gas imaging and EPA Method 21, assuming onsite surveys are conducted quarterly. We additionally model an emerging technology based on the design criteria of the Environmental Defense Fund's (EDF) Methane Detectors Challenge, which is a grant program that targets mature technologies for detecting larger leaks on a continuous basis that are deployable within the next two years.[66,67] Each detection method has associated costs and detection limits, which we incorporate in the optimization model (refer to SI Appendix B.3).

2.2.10 Social Cost of Methane

Social benefits associated with emissions reductions are estimated using the social cost of methane (SC-CH_4), a metric representing the present value of the anticipated future damages that would arise from an incremental unit of methane emissions in a given year.[68] SC-CH_4 accounts for climate change impacts, such as changes in agricultural productivity and human health, property damage from increased flood risk, and changes in heating and cooling costs. We do not account for benefits from concomitant reductions in volatile organic compounds and safety improvements.

We treat SC-CH_4 parametrically, employing four estimates (in units of 2014 USD per metric ton of methane) with differing discount rate assumptions, as reported in Marten et al. (2014; 2015): $601 (5% discount rate, mean), $1330 (3% discount rate, mean), $1780 (2.5% discount rate, mean), and $3560 (3% discount rate, 95th percentile).[69] We choose an abatement or analysis year of 2020, which is reflective of a realistic regulatory horizon; benefits would differ if the analysis year differed. SI Appendix E provides background information on direct modeling of the SC-CH_4 using integrated assessment models and a comparison to indirect estimates based on conversion of non-CO_2 emissions to carbon dioxide equivalents (CO_2eq) using global warming potentials and applying the social cost of carbon (SCC).

2.2.11 Optimization Model

To assess different abatement strategies for the U.S. T&S system, we employ an optimization framework that leverages the previously described sub-modules. To represent different system-wide and superemitter policies and associated policy instruments, we formulate integer linear programs with different objective functions, simulated inputs, and constraint sets. The models select the components in the system at which to perform abatement. Table 1 lists the basic structure of each policy; further information regarding the formulations is provided in SI Appendix F. The optimization models are coded in and solved using R and the General Algebraic Modeling System (GAMS).

We assume that the analysis period is a year, which has practical implications for policy implementation. The relative timing of detection, classification of facilities as superemitters (which may be ephemeral from year-to-year), and abatement would have to determined to implement a policy; however, we assume that these activities all occur within a year and do not attempt to order and further discretize.

2.2.11.1 Unconstrained System-Wide Policy

The objective function of this optimization is to minimize net social costs (or equivalently maximize net benefits), which includes abatement and detection costs (less cost savings in some formulations) and social costs of methane. The optimal level of abatement occurs where marginal benefits equate to marginal costs.

2.2.11.2 System-Wide Policy with Facility-Level Emissions Reduction Target

The objective function of this model is to minimize private costs, including abatement and detection costs (less cost savings in some formulations), subject to an emissions reduction target per facility (we looked at 10, 50, and 75% reduction targets). For each facility, the model selects the components to abate to just meet the emissions reduction target, if feasible. For example, given a target of 50%, facilities must quantify emissions and perform abatement to achieve a 50% reduction. Facilities for which it is not possible to achieve a given target, the maximum emissions reduction is selected.

2.2.11.3 Superemitter Policy with Absolute Emissions Threshold

Here, the objective function is the same as in the unconstrained system-wide policy, but abatement is only conducted at a subset of facilities with absolute annual emissions above a specified threshold. We model two absolute emissions thresholds: 50^{th} and 90^{th} percentiles of simulated facility-level emissions (i.e., the subsets are comprised of the highest emitting 50% and 10% of facilities, respectively).

2.2.11.4 Superemitter Policy with Proportional Loss Rate Threshold

In this scenario, the objective function is the same as in the unconstrained system-wide policy, but abatement is only conducted at a subset of facilities with proportional loss rates above a specified threshold. We model two thresholds: 50th and 90th percentiles of simulated facility-level proportional loss rates (i.e., the subsets are comprised of 50% and 10% of facilities with the highest proportional loss rates, respectively).

2.3 Results and Discussion

We first present intermediate results, including component- and system-level baseline emissions estimates and marginal abatement cost curves. Then, we describe findings for the unconstrained system-wide policy optimization modeling with respect to net benefits, private costs, and emissions reductions, and summarize various sensitivity analyses. We then compare facility- and system-level results across policy options and assess tradeoffs between different policy instruments.

2.3.1 Baseline Emissions

The modeled total methane emissions for the T&S sector (440,000 metric tons), accounting for some but not all emission sources, are similar to other recently published estimates. Using the existing suite of abatement technologies considered here, approximately 94% of emissions are from emission categories that have the potential to be abated, with a few emission categories (i.e., reciprocating compressor isolation valves and rod-packing) comprising approximately 57% of abatable emissions. SI Figure H-1 presents the estimated emissions for each emission category.

Based on the emissions simulation, 30% of the facilities account for 60% of the emissions (see SI Figure F-2). This skewness, which depends on the model structure and underlying data, may not be representative of the actual skewness and may not capture high-emitting facilities that are low in frequency across the T&S system. To contextualize the modeled emissions, methane emissions from the injection well leak in Aliso Canyon, California, were reported to be approximately 100,000 metric tons, which is much higher than the maximum emitting-facility that was modeled.[70]

2.3.2 Marginal Abatement Costs

For each emission category and corresponding abatement measure, we simulate marginal abatement cost (MAC) curves, accounting for variability in abatement costs, natural gas prices, emission factors, and operating hours (see SI Appendix H). We also aggregate MAC curves for each emission category to create system-wide curves. A MAC curve step function is presented in Figure 1a; each step represents the median simulated emissions reductions for an abatement measure, rank-ordered based on the median simulated marginal costs. Assuming no savings from recovered gas and fixed marginal costs for each abatement measure, nearly all potential emission reductions from the T&S system have marginal costs less than $50 per CO_2eq metric ton. For comparison, the Integrated Working Group (IWG) SCC estimate is $47 per CO_2eq metric ton (year = 2020, 3% discount rate).[71] Therefore, if a policy accounts for the IWG SCC estimates, this analysis indicates that most of the emissions should be abated.

Figure 1b displays system-wide marginal abatement cost curves for 100 realizations of the T&S system, with each curve representing a rank-ordering of every component in the system based on increasing marginal abatement costs. We observe that there is a distinct "knee" in the MAC curve at which point the marginal costs markedly increase. Employing the efficiency criterion, whereby net benefits are maximized when marginal benefits equate to marginal costs, the optimal level of abatement (i.e., subset of components in the system to abate), without accounting for savings and assuming a social cost of methane of around $1330 per metric ton of methane (~$46 per metric CO2eq ton), occurs at ~80% reduction in emissions. If we incorporate savings from marketing recovered gas, then an approximately 70% of emissions may be abated with positive net private benefits.

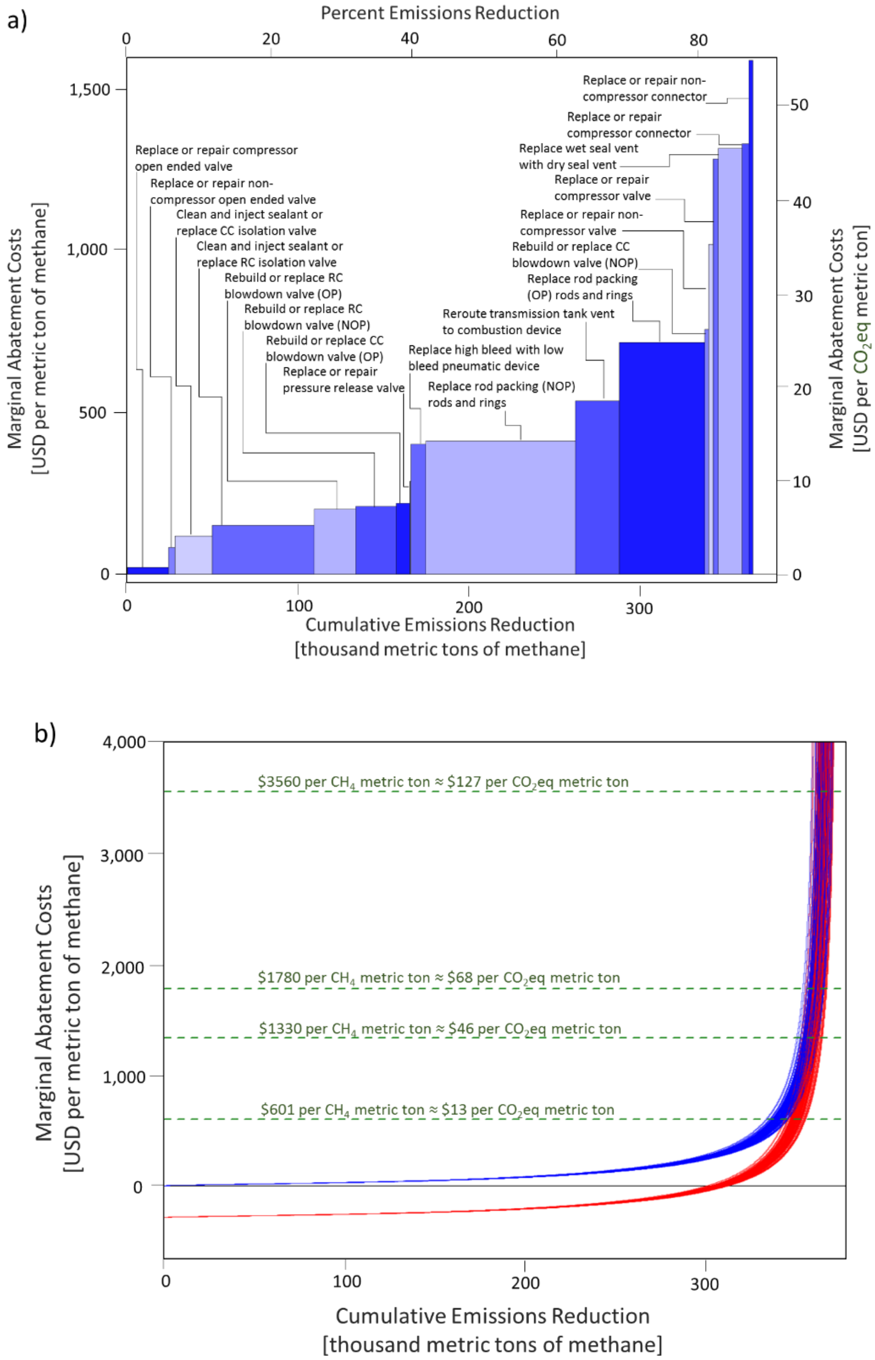

Figure 1. (a) System-level marginal abatement cost curve step function based on simulated median emissions and marginal costs, assuming no savings from natural gas. (b) Marginal abatement cost curve based on simulation of all components in the U.S. T&S system, accounting only for variability in emissions factors, operating hours, and abatement costs. Blue curves represent 100 realizations of the T&S system, assuming no savings; red curves assume savings (using the EIA reference case natural gas price projection). Green lines are marginal benefit curves at different assumed social costs of methane, based on modeling in Marten et al. (2015).

2.3.3 Policy Optimization

The results shown thus far assume omniscience in which the emissions from and abatement cost-effectiveness for each component in the T&S system are known *ex-ante*. For optimization modeling of the different policy scenarios, we also account for detection to identify emission sources, a necessary condition prior to abatement (in our stylization of policies), and consider policies in which only a subset of facilities perform abatement.

2.3.4 Unconstrained System-Wide Policy

For the unconstrained system-wide policy, when maximizing net benefits under base case modeling assumptions, most components in the T&S system (i.e., those with marginal costs less than the social cost of methane) are selected by the model for abatement (as depicted in Figure 2 inset).

Results of the optimization model, over a range of system-level emission reduction targets, are provided in Figure 2. Under scenarios both with and without gas savings, the optimal emissions reduction is approximately 80%, which is similar to the maximum level of abatement (83%). Figure 1b provides some insight into why this occurs; the marginal benefits curve intersects the vertical region of the MAC curve, in which marginal cost markedly increase with minimal emission reductions. Optimal net benefits are $420M without savings, and increase by 25% when accounting for savings (assuming a natural gas price of ~$5/MMBTU).

We performed sensitivity analyses, parametrically varying uncertain inputs, including social cost of methane, natural gas price, detection technology, discount rate, and abatement efficacy. Key results are shown in Figure 3 and additional details provided in SI Appendix H.4.

Figure 3 shows that net benefits are most sensitive to the social costs of methane. Across all of the single factor sensitivity analyses performed, the net benefits are positive, indicating that even with uncertainty in key inputs, regulation of methane emissions from the T&S sector is warranted. Changes in natural gas price, abatement efficacy, detection technology, and discount rate slightly shift the marginal cost curve, but the slope generally remains the same. Thus, across the parametric analyses performed, the marginal benefits curve intersects the vertical region of the marginal cost curve; as a result, the optimal subset of components to abate, as well as the optimal system-wide emissions reduction, remains relatively constant.

Since the optimal level of abatement occurs when marginal benefits are equivalent to marginal costs, the optimal emission reduction increases with increasing social costs of methane. Even under the most conservative scenario modeled ($600 per metric ton of methane ~ $13 per metric CO_2eq ton), mean net benefits at the optimal level of abatement are $160M. At the highest social cost modeled ($3,600 per metric ton of methane ~ $136 per metric CO2eq ton), mean net benefits of $1.2B are estimated.

Under scenarios assuming savings and varying natural gas prices, net social benefits are 16-25% higher than the scenario without savings, and private firms may experience net savings of upwards of $40M. The breakeven natural gas price, at which private firms experience no net costs, is approximately $3/MMBTU, which is in the range of EIA projections. As previously noted, transmission firms may not own the gas and their fee structure may be based on long-term contracts that stipulate fixed leak rates. Thus, a clear fiscal incentive for firms to minimize leakage may not exist. Even in other segments of the natural gas supply chain where there may be a fiscal incentive, firms have historically not adopted abatement practices, potentially because of a lack of information and cost uncertainty. However, with the recent advent of efforts to measure emissions and advancements in the NG STAR voluntary emission reduction program, which together signal impending regulation, it is possible that more firms will adopt abatement practices.

Although net benefits are relatively constant across detection technology scenarios, private costs are somewhat sensitive; private costs under the emerging technology scenario (based on the EDF Detectors Challenge) are 12-31% lower than for existing technology scenarios, which is solely attributable to differences in detection costs. Across the detection technologies modeled, detection costs contribute between 1-32% of total private costs. Additionally, at the breakeven point at which total social benefits are equivalent to total private costs, annual detection costs are estimated to be over $200,000 per facility, which is substantially higher than the cost of the detection technologies modeled. There are minimal differences in emissions reductions between technologies at both the optimal and maximum level of emissions reductions, despite the large range of detection limits modeled. Results further indicate that the detection limit does not substantially impact which components are selected by the model to abate; rather cost-effectiveness of abatement is typically limiting. This implies that improvements in the sensitivity of detection technologies may not contribute to emission reduction efforts because very low-emitting components typically have higher marginal abatement costs.

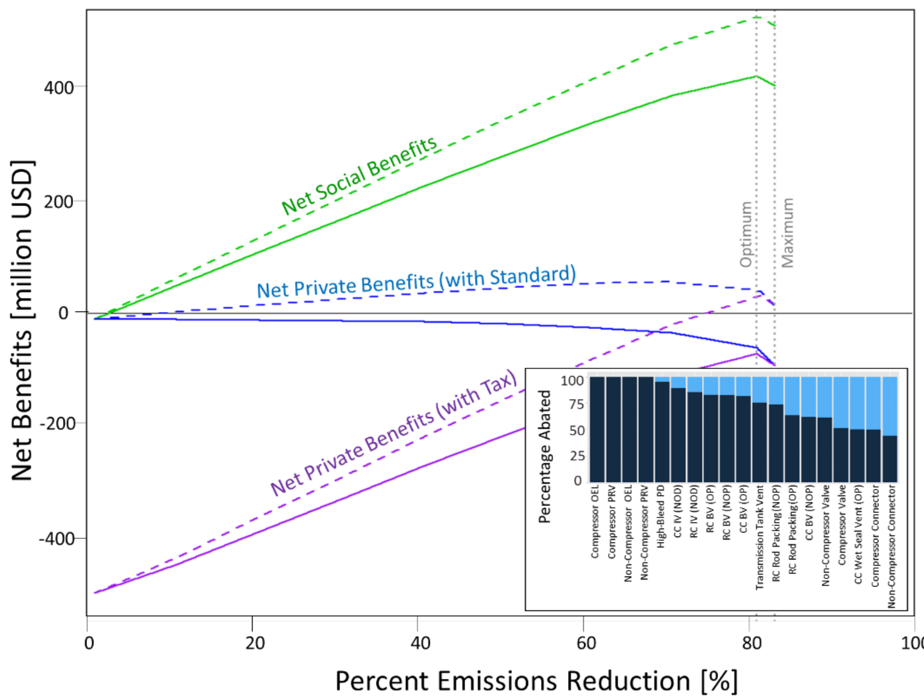

Figure 2. Net social benefits and net private benefits at different levels of system-wide emission reductions for the unconstrained system-wide policy. The results are expected values based on 100 realizations of the U.S. T&S system. Solid lines are for model runs assuming no savings from capturing natural gas. Dashed lines are for model runs assuming savings (using the EIA reference case natural gas price projection). The inset bar chart depicts the mean percentage of components for each emissions category that are selected for abatement at the optimum.

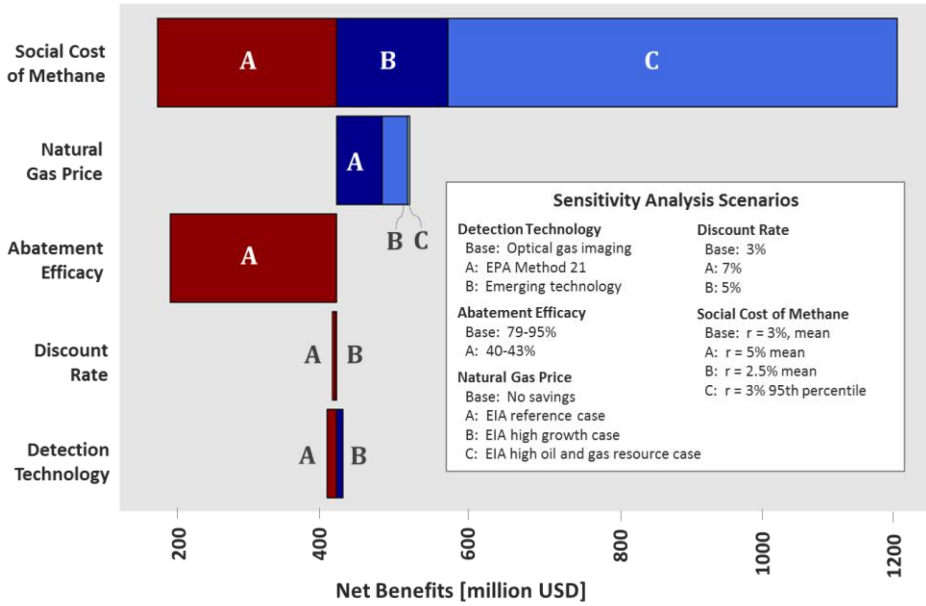

Figure 3. Tornado diagram representing expected value of net benefits across a range of parametrically varied inputs. The inset box maps the base and varied inputs to the tornado plot.

2.3.5 System-Wide Policy with Facility-Level Percent Reduction Target

Tradeoffs between private costs, social benefits, and emissions reductions across different system-wide and superemitter policies are depicted in Figure 4. We observe that the benefit-cost ratios (9 to 13) of all three system-wide policies with different facility-level targets (10, 50, or 75%) exceed the benefit-cost ratio (8) of the optimal unconstrained system-wide policy; this is attributable to conducting the most efficient abatement first (in aggregate across the system) relative to the unconstrained policy. While requiring all facilities to reduce emissions by an equivalent percentage may be equitable in a sense, some facilities may be required to conduct abatement at low-emitting components with high marginal abatement costs that exceed the social cost of methane (refer to SI Figure H-12). In addition, some facilities may be unable to achieve a target, given the current suite of abatement technologies.

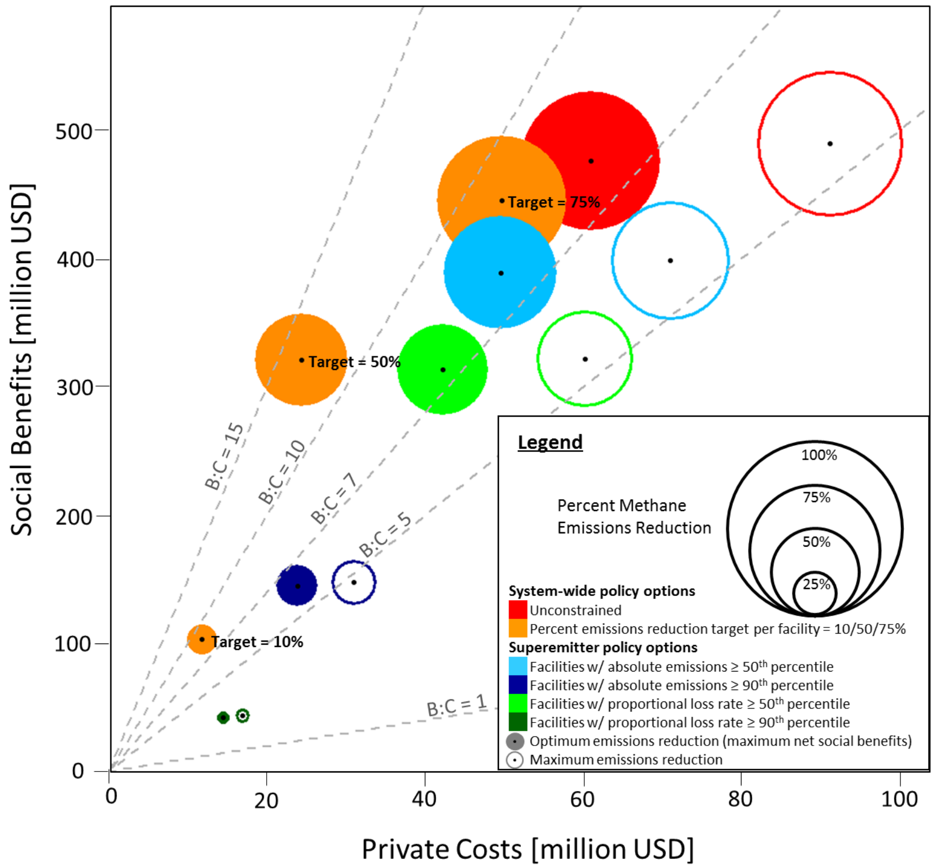

Figure 4. Tradeoffs between system-wide and superemitter policies with respect to social benefits, private costs, and emission reductions. Size of bubbles is proportional to the mean system-level percent emissions reduction, and the black center points represent the mean benefits and costs over 100 iterations of the model. Open bubbles represent maximum emission reductions and closed bubbles represent optimal emission reductions. Grey dotted lines indicate benefit-cost ratios (B:C).

2.3.6 Superemitter Policies with Absolute Annual Emission and Proportional Loss Rate Thresholds

In both types of superemitter policies, the maximum achievable and optimal emissions reductions are less than reductions under the unconstrained system-wide policy, given that abatement is only conducted at a subset of facilities with the highest emissions or loss rates. The superemitter policies with proportional loss rate thresholds are dominated by policies with absolute annual emission thresholds, which can achieve equivalent or greater emission reductions at lower cost. We also observe that both types of superemitter policies are dominated by system-wide policies, which may be attributed to the requirement to perform detection at all facilities to identify superemitters. The baseline emissions model may not capture low probability, catastrophic leaks from facilities; thus, if emissions across the system are more skewed than what has been modeled, there may be greater efficiency gains in targeting superemitters.

2.3.7 Policy Instruments

An additional dimension of this analysis is the assessment of different policy instruments, including component- and facility-level standards and a tax. There are other possible instruments which we do not model, such as best management practices, a voluntary emission reduction program, and a cap-and-trade program.

We computed the performance standards and tax rate required to achieve the optimal level of abatement under the unconstrained system-wide policy. SI Appendix H.6 describes the process we developed to estimate component-level standards (or emissions thresholds). The tax rate (applied to all unabated emissions that are detected and abatable) would have to be equivalent to the marginal benefits (or the social cost of methane). Regardless of the regulatory instrument under an unconstrained system-wide policy, net social benefits are unaffected because the tax paid by private firms is transferred to the public (ignoring transaction costs). Private costs under a performance standard policy include abatement and detection costs, whereas, under a tax instrument, private costs also include taxes paid on unabated emissions that are detected and potentially abatable. At the optimal level of abatement, private costs (without recovering the value of abated natural gas) are approximately $61M under a performance standard policy and $74M under a Pigouvian tax instrument. Total detection costs across the system range from $1M to $24M, depending on detection technology. To contextualize these private costs, the operating and maintenance costs and net income of all U.S. transmission facilities in 2012 were $6.2B and $4.8B, respectively; thus, the private costs under the modeled policy are relatively insignificant in aggregate and are roughly 1% of O&M costs and net income of private firms.[72] The breakeven system-level emission reductions, where private costs equal zero, are approximately 8% and 74% under a performance standard and tax instrument, respectively, which are lower than the optimal level of emissions reduction (80%).

Beyond differing private cost burdens, additional tradeoffs exist between taxes and performance standards. While both instruments (in theory) encourage innovation that may lower private costs, a tax also generates public revenue, which can be diverted to fund innovation or public services. Private firms also possess the property rights to emit under a standard, whereas, rights shift to the public under a tax. In the context of uncertain marginal costs and benefits, the policy instrument should, according to the Weitzman principle, mimic the slope of the marginal benefits curve. Methane is a stock pollutant – a pollutant that accumulates in the environment over time before damages occur – and small, incremental changes in emissions generate benefits that are equivalent at the margin; thus, the marginal benefit curves are flat, suggesting that a tax is preferable. However, current policy regulating emissions from natural gas facilities employs standards (e.g., NSPS), indicating that a standard may be more politically tractable.

Under a system-wide policy with a facility-level percent reduction target, facilities may choose a (least cost) abatement strategy to meet a given target; thus, a performance standard and tax instrument are unnecessary. For the superemitter policies, component-level standards, similar to those estimated for the unconstrained system-wide policy, may be employed to achieve the optimal level of abatement. An alternative to system-wide or superemitter policies is to require facilities to perform the maximum (achievable) level of abatement, which may be implemented through a component-level standard or technology requirement.

2.4 Policy Implications

Given the existing suite of abatement technologies, most methane emissions from the T&S system can be abated, with only a few emission categories comprising a majority of the abatable emissions. Under an unconstrained system-wide policy, most emissions (approximately 80%) are efficient to abate, and the private cost burden under both performance standard and tax instruments is modest relative to the potential societal benefits and current operating costs of firms. Furthermore, if firms market recovered gas, private benefits may result.

System-wide policies with facility-level percentage emissions reduction targets have the advantage of requiring equivalent percent emission reductions across facilities and result in higher benefit-cost ratios (at the targets modeled) than other policies; however, at a facility-level, some facilities would be required to perform abatement at the margin that is economically inefficient.

Superemitter policies, namely those in which only a targeted subset of facilities with the highest absolute annual emissions perform abatement, may reduce the system-level private cost burden and achieve high emissions reduction, especially if emissions across facilities are very skewed. However, a superemitter policy is unprecedented, and there may be practical limitations to implementation, including the relative timing of detection, superemitter classification, and abatement. In addition, detection across all facilities is necessary regardless of the policy option and there are non-trivial net social benefits resulting from abatement of relatively low-emitting sources.

3 Cumulative Air, Climate, and Employment Impacts of Natural Gas Systems

Natural gas has become the largest fuel source for electricity generation and accounts for a third of energy production and consumption in the United States. However, the cumulative environmental and socioeconomic impacts from extraction to end use and over the life of natural gas plays have not been comprehensively characterized to facilitate long-term decision making and comparisons to competing energy technologies. Here, we present an approach for robust estimation of spatially- and temporally-resolved cumulative air quality, climate change, and employment impacts through a case study of the shale gas boom in the Appalachian basin from 2004 to 2016. We find that short-lived air quality impacts (1200 to 4600 deaths; \$23B +99%/-164%) and employment impacts (467,000 job-years ±30%; \$21B ±30%) track with the boom-and-bust cycle, while climate impacts (\$12B to \$94B) persist for generations well beyond the period of natural gas activity. Employment effects are spatially concentrated in rural areas with thin labor markets where development is occurring. More than half of cumulative premature mortality is within source emissions states, while transboundary impacts are concentrated in populous coastal regions in the Northeast. Most premature mortality is associated with end uses, while upstream and midstream segments also account for a substantial portion of impacts. With respect to climate change impacts, the magnitude of methane emissions across the supply chain produces temperature impacts nearly equivalent to that of CO_2 over a 30-year time horizon, and over longer integration periods, the warming impact from CO_2 dominates that of CH_4. We additionally find that there is an implied tradeoff of approximately 200 job-years per death at a cumulative systems level. We estimate a tax on production of \$2 per thousand cubic foot (+172%/-76%) compensates for cumulative climate and air quality externalities across the supply chain.

3.1 Introduction

Rapid increases in U.S. natural gas production, resulting from advancements in horizontal drilling and hydraulic fracturing, have had reverberations on world energy markets and the domestic energy outlook. The U.S. has been the largest natural gas consumer and producer over the past decade, comprising 20% of the world market, and the domestic shale gas market has contributed to price volatility and a shift in global flows of natural gas.[31] Increases in domestic supply and reserves have contributed to the displacement of coal[32], but have also potentially delayed investment, development and deployment of renewables.[33]

The rapidly evolving energy landscape has presented new challenges in the areas of science and regulation.[73] There is expanding literature on the impacts of natural gas development on water quality[74–77], air quality[78–80], ecosystems[81,82], climate[83–85], labor markets[86–88], public health[89,90], and several other environmental and socioeconomic factors. However, the cumulative impacts over the boom-and-bust cycle and across the natural gas supply chain, as well as the spatial and temporal distribution of impacts, are still largely unexplored and unaccounted for in public and private decision making.

Here, we develop an integrated architecture to compute and characterize the cumulative impacts of natural gas systems, which we apply to the shale gas boom (and decline) in the Appalachian basin, the largest natural gas basin in the U.S. with respect to both reserves and production.[34] We examine air quality, climate change, and employment impacts across the supply chain from development to end use. We develop robust spatially- and temporally-resolved estimates using detailed natural gas activity, demographic, and emissions data and based on an integrated set of models, including process-level emissions inventories, reduced complexity source-receptor air quality and climate change models, and fixed effects models for assessing employment effects. Our approach highlights the attribution of impacts across the supply chain, the spatial distribution of impacts, and the evolution and accumulation of impacts over time with changing regulation, natural gas activity, and technological and operational efficiencies and practices. We compute estimates in physical units (i.e., premature mortalities, temperature change, and employment) and monetary units, which provide differential insight for decision making and policy design.

3.2 Methods

The intent of the study is to develop a framework with which we can comprehensively assess the temporal and spatial distribution and tradeoffs across multiple impact areas, each with differing magnitude, cumulative nature, and spatial and temporal dimensions. There are many impacts from natural gas development; we chose a subset as a basis for developing and demonstrating the methodology. These impacts include: premature mortality from primary fine particulate matter ($PM_{2.5}$) and secondary $PM_{2.5}$ formed from the atmospheric oxidation of nitrogen oxides (NO_X) and volatile organic compounds (VOCs) emissions, global mean temperature change from carbon dioxide (CO_2) and methane (CH_4) emissions, and employment effects associated with natural gas development.

We focus this analysis on the development of the Marcellus and Utica shale gas plays, which are part of the Appalachian basin. We model impacts from the rapid development of the plays beginning in 2004, the year in which the first shale gas well was drilled in the Marcellus play, to 2016. We only consider impacts that directly stem from natural gas activity (from production through consumption) within Pennsylvania, Ohio, and West Virginia. For example, we model emissions from activity within the tristate region and the corresponding mortalities which extend beyond tristate boundaries, but exclude emissions from gas exported to interstate and international markets and end use outside of the region [refer to Section 1 of the Supplementary Information (SI) for additional details].

The magnitude, timing, and geographic location of shale gas activity, including well development, production, transmission, distribution, and processed volumes, and upstream, midstream, and end use fuel consumption volumes, are publicly reported and fundamental inputs that cut across impact areas (see SI Section 3, Table S2 and SI Section 4, Table S9). Shale production volumes increased annually over the period of analysis, while drilling peaked in 2013 and has since rapidly declined (see SI Section 2, Figures S5 and S8). Despite rising electricity generation from natural gas, shale gas production (7.7 tcf in 2016) has exceeded natural gas end use demand (2.1 tcf in 2016) within Pennsylvania, Ohio, and West Virginia in recent years, leading to substantial exports to other regions of the country (SI Section 2, Figure B9).

Processing volumes have also increased (18-fold increase from 2004 to 2016) with planned processing capacity expansions slated over the next decade, including the development of a regional petrochemical industry.

Premature mortality and monetized damage estimates from air pollution emissions are derived for each segment of the natural gas supply chain (see SI Section 3). We derive emissions estimates at the county or finer spatial resolutions and for each year. We use both parametric and probabilistic process-level methods to characterize uncertainty in upstream emissions (i.e., drilling, hydraulic fracturing, well completion, trucking, wellhead compressors, condensate tanks, production fugitives, etc.), and we derive deterministic estimates of midstream (i.e., processing, transmission, distribution) and end use (i.e., electricity generation, residential, industrial, commercial) emissions based on operator-reported emissions and national inventories. Upstream emission model formulations account for time-varying parameters, such as changing regulation, natural gas activity, and water management practices and operational efficiencies associated with upstream activities. Premature mortalities are estimated by combining the emissions inventory with three source-receptor reduced complexity models (RCMs): Air Pollution Emission Experiments and Policy model (Version 3) (AP3)[91,92], the Air Pollution Social Cost Accounting tool (APSCA)[93], and Intervention Model for Air Pollution (InMAP)[94], which are functionally different but provide complementary insight. We use these three different models because they use distinct approaches to represent pollutant fate and transport, and thus, provide a representation of exposure uncertainty. Both AP3 and APSCA generate estimates of pollution-induced premature mortalities in downwind receptors associated with emissions from source locations. We use the source-resolved version of InMAP whereby pollution-induced mortality risk is attributed to source location. Premature mortality estimates are sensitive to the relationship between pollutant concentration and health response; therefore, we parametrically vary the concentration-response (C-R) relationship based on the American Cancer Society (ACS) and Harvard Six Cities (H6C) studies.[95,96] To develop monetized impact estimates, we use the value of a statistical life (VSL), a commonly used measure of the dollar value of small changes in mortality risk experienced by a large number of people.

We use a probabilistic VSL with a mean (and standard deviation) of \$8.5M (±\$5.7M), and we assume the VSL follows a Weibull distribution, following the approach used by the U.S. Environmental Protection Agency (in 2017 USD).[97]

Greenhouse gas (GHG) emissions, global temperature change, and monetized damage estimates are derived for each segment of the natural gas supply chain and over different integration periods (SI Section 4). We use both parametric and probabilistic methods to characterize uncertainty in process-level upstream and processing GHG emissions, and we derive deterministic estimates of transmission, distribution, and downstream emissions based on reported volumes. Global temperature change is estimated using a convolution of the emissions model and the average global temperature potential (AGTP).[98–103] AGTP is for a given time in the future and is due to a marginal pulse of emission. It is a function of the temperature response to radiative forcing which, in turn, is due to a pulse emission, both of which are parameterized based on more complex models that explicitly include physical and chemical processes.[99] We use a Monte Carlo simulation approach to reflect uncertainty in the AGTP values. To generate monetized estimates of climate damages, we employ the social cost of carbon (SCC) and social cost of methane ($SCCH_4$), metrics representing the present value of the anticipated future damages that would arise from an incremental unit of emissions in a given year. We assume values ranging from \$10 to \$126 per metric ton of CO_2 and \$319 to \$2773 per metric ton of CH_4, as reported in U.S. E.P.A. publications.[68,71] These metrics account for climate change impacts, such as changes in agricultural productivity and human health, property damage from increased flood risk, and changes in heating and cooling costs.

To isolate employment effects from shale gas activity, we compile a panel of county-level natural gas activity and demographic data (reported by state environmental agencies, U.S. Census Bureau, U.S. Bureau of Economic Analysis, U.S. Bureau of Labor Statistics) for Pennsylvania, West Virginia, Ohio, and New York over the period 2004 to 2016 (see SI Section 5). Using an approach similar to that of Paredes et al. (2015)[86], we specify fixed effects models, controlling for the diversity of local labor markets and including county fixed effects that control for observable and unobservable differences across counties and time fixed effects that control for common shocks across counties that vary over time. To estimate the models, we use ordinary least squares (OLS) regression, and to control for serial correlation, we estimate robust standard errors by clustering by county. We test the effect of lag and lead variables to capture the dynamic effect of natural gas development, natural gas activity variables such as spud and producing wells, natural gas development and time interactions, and alternative rate-based employment dependent variables. Aggregate county-level employment over time is estimated by combining the natural gas activity data with the marginal employment effects. To monetize employment effects, we apply the annual average earnings per county.

3.3 Air quality impacts

We model premature mortality and monetized damages from primary $PM_{2.5}$ and secondary $PM_{2.5}$ formed from the atmospheric oxidation of NO_X and VOCs emissions. Figure 5a-b show estimates of annual emissions and the spatial distribution of emissions. VOC emissions are largely associated with upstream processes (61%), and spatially concentrated in counties with the highest cumulative production (SI Section 3, Figures S15, S17). End use processes contribute a majority of NO_x (67%) and $PM_{2.5}$ (73%) emissions across the natural gas supply chain, and NO_x and $PM_{2.5}$ are relatively evenly distributed across counties (SI Section 3, Figures S14, S16, S18). Compared to total emissions within Pennsylvania, Ohio, and West Virginia in 2014 from all industries reporting in the National Emissions Inventory (NEI), shale gas activity accounted for approximately 10% of NO_x emissions, while the contribution to total $PM_{2.5}$ and VOC emissions was marginal (<2%). As noted in other studies, many counties within and adjacent to Appalachia are designated nonattainment or maintenance areas under the Clean Air Act National Ambient Air Quality Standards (NAAQS), and increasing emissions from shale activity contribute to and are projected to continue to contribute to regional noncompliance.[80,104] In isolation, most shale gas activities across the supply chain, including some natural gas electric generating units, do not constitute *major sources* under federal regulations, but in aggregate at the county-level may exceed thresholds.

We estimate 1200 to 4600 premature mortalities are associated with shale gas activity over the period 2004 to 2016; the range reflects uncertainty in the model functional form through the use of three RCMs (factor of up to 1.6X) and uncertainty in the concentration-response relationship (factor of up to 2.6X). As depicted in Figure 1c, more than half of the cumulative premature mortality is within source counties (54%), while transboundary impacts are concentrated in populous coastal regions in the Northeast and extend to the continental divide (SI Section 3, Figure B19). While the majority of the premature mortality is associated with residential, industrial, commercial, and electricity generation end uses (57 to 67%), upstream (16 to 21%) and midstream (17 to 22%) segments also contribute substantial cumulative air quality impacts (SI Section 3, Figure B20). There is also an urban and rural divide, with 76% of premature mortalities occurring in urban areas. Mean cumulative damages, based on mean mortality across six model specifications, amount to $23B (in 2017 USD), with a 95% confidence interval spanning two orders of magnitude ($2.3B to $61B) reflecting different VSL assumptions. Annual mortalities (439) and damages ($3.7B) peaked in 2014, the year after peak drilling activity. These estimates of impacts associated with air pollution are conservative, given that out-of-state consumption of Appalachian natural gas feedstock is not included and the set of processes, species, and health and environmental endpoints other than premature mortality is not all-encompassing. Furthermore, health impacts from natural gas extend beyond those associated with air pollution, such as health benefits associated with increased access to healthcare or costs associated with increased traffic accidents. To contextualize the estimated mortality, a study by Fann et al. (2018) projected $PM_{2.5}$-related premature mortalities attributable to emissions from the O&G sector in Pennsylvania and Ohio in 2025 to be 150 (95% CI 91 to 196). Although not directly comparable, given that our study has a vastly expanded scope, a previous study by Litovitz et al. (2013) estimated damages of $8-35M from preproduction, production, and compressor station emissions in Pennsylvania in 2011 only.[78]

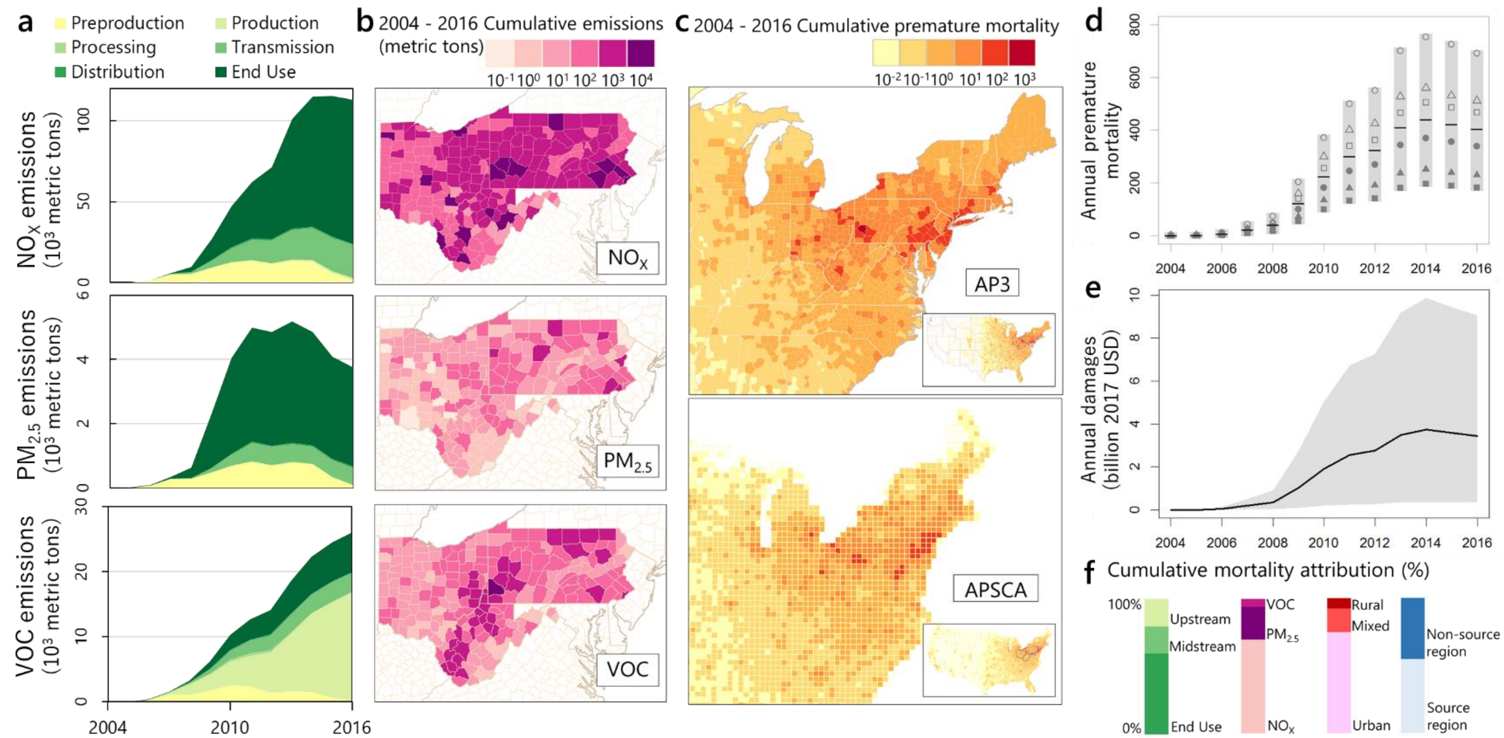

Figure 5. Air quality emissions and impacts across the natural gas supply chain from 2004 to 2016. (a) Annual NO_x, $PM_{2.5}$, and VOC emissions from sources within Pennsylvania, Ohio, and West Virginia. (b) Spatial distribution of cumulative NO_x, $PM_{2.5}$, and VOC emissions by county from 2004 to 2016. (c) Spatial distribution of cumulative premature mortality from 2004 to 2016 by receptor county. Estimates based on AP3 and APSCA source-receptor RCMs using the American Cancer Society (ACS) concentration-response (C-R) relationship. Larger figures depict Northeast U.S. and the insets depict the continental U.S. Blue lines border the source emission states. (d) Annual premature mortality using different RCMs and C-R relationships. Solid points represent estimates based on ACS C-R relationship, and open points represent estimates based on Harvard Six Cities (H6C) C-R relationship. Circle, triangle, and square points represent estimates based on AP3, APSCA, and InMAP, respectively. Black lines represent average annual mortality across all six specifications. Grey shaded regions represent range of annual estimates. (e) Annual damages associated with premature mortality from air pollution. The black line and grey shaded region represent the simulated mean and 95% confidence interval, respectively, and reflect uncertainty in the VSL. Based on average annual mortality across all six specifications, as shown in (d). (f) Attribution of cumulative mortality from 2004 to 2016 by segment of the natural gas supply chain, emission source/non-source regions, urban/mixed/rural geographic regions, and air pollutant species. Attribution is estimated using premature mortality based on AP3 and ACS C-R specification.

3.4 Climate change impacts

We estimate global mean temperature change and monetized damages from CO_2 and CH_4 emissions. As shown in

Figure 6a, CH_4 emissions are largely associated with production (63%) and gathering segments (15%), while processing, transmission, and distribution collectively account for the remaining emissions (SI Section 4, Table S14). With respect to the distribution of CH_4 emissions across the supply chain, our findings, which leverage recently collected emissions data, are consistent with recent estimation studies of emissions across the U.S. oil and gas (O&G) sector. Compared to national O&G sector estimates (13 MMT in 2015)[105], CH_4 emissions from natural gas-related sources within Pennsylvania, Ohio, and West Virginia (1.25 MMT in 2015) account for 10% of emissions.

End use processes contribute a majority of CO_2 (85%) emissions across the supply chain, with remaining emissions attributable to well development (2%) and fuel consumption for production, processing, transmission, and distribution (13%). Compared to CO_2 emissions across the entire U.S. natural gas sector (1502 MMT in 2016) as reported in the U.S. Greenhouse Gas Inventory (GHGI)[44], natural gas-related sources in Pennsylvania, Ohio, and West Virginia (134 MMT in 2016) account for 9% of emissions.

We translate emissions into impacts, including global temperature change and monetized damages, which provide differential insight. Impacts from climatically-relevant emissions are often described by the following illustrative chain: emissions → atmospheric concentrations → radiative forcing → climate change → societal and ecosystem impacts → monetized damages.[106–108] Stepping through the chain, there is cascading uncertainty and (arguably) increasing societal relevance.[107,109]

Using the social costs of carbon and methane, consistent with traditional benefit-cost analyses, we develop estimates of monetized damages from natural gas activity (Figure 6Figure 6d). Cumulative damages from natural gas activity over the period 2004 to 2016 range from \$12B to \$94B, depending on assumptions regarding social costs. A recent study by Ricke et al. (2018) estimates a SCC (\$177–805 per metric ton for 2020 emissions) much greater than the highest estimate in this study (\$126 per metric ton in 2016), suggesting our damage estimates may be conservative.[110] Although the use of physical metrics has been subject to criticism within the economics literature[111–113], metrics related to climate change phenomena, e.g., temperature, precipitation, sea level rise, and extreme events, provide additional information (e.g., temporal trace of climate impacts) and apply a different set of assumptions than monetization.

Translating emissions into global temperature change, the climate responses persists well beyond the period of shale gas activity, as shown in Figure 2b (see also SI Section 4, Figures S24-27). Short- and long-lived species emitted across the sector have differential effects on the trajectory of climate response. CH_4 has an atmospheric lifetime of 12 years, whereas CO_2 has multiple lifetimes, with 20% remaining for tens of thousands of years.[114] The magnitude of CH_4 emissions across the supply chain produces temperature impacts nearly equivalent to that of CO_2 over a 30-year time horizon from the year of first production, 2004. Over longer time periods (e.g., 100 years), the warming impact from CO_2 (77%) dominates that of CH_4 (Figure 6b). We additionally find that greater than 65% of temperature impacts integrated over a 100-year period are associated with end use, namely CO_2 emissions from natural gas combustion.

The relative atmospheric lifetimes of CO_2 and CH_4 have implications for policy design. CH_4 emissions can be treated as a flow, where reductions using the existing suite of cost-effective abatement technologies (e.g., replacing leaking components at compressor stations) can reduce near-term warming rates and result in climate benefits realized over relatively short time horizons. Furthermore, reducing CH_4 emissions may also be key in the avoidance or delay of reaching "tipping points" in the climate system, irreversible thresholds with drastic consequences.[115] In contrast, CO_2 emissions can be treated as a stock, and reductions can only be achieved through deep decarbonization or a fundamental transition of the energy system away from fossil fuels, including natural gas, the benefits of which are derived by future generations.

We only consider the temporal dimensions of climate change impacts from long lived GHGs. Global metrics, including global temperature change and monetized damages using the SCC, are useful for characterizing spatially and temporally smooth climate responses from emissions of well-mixed GHGs.

However, other short-lived chemically reactive gases (NO_x, CO, and VOCs) that indirectly lead to changes in climate forcers, as well as aerosols and precursors (black carbon, organic carbon, SO_2), react on very different time-scales and have regionally heterogeneous effects; in addition, climate forcings may be both negative and positive, thereby contributing to both warming and cooling.[101] Thus, a global metric provides a rather limited view of potentially nontrivial regional impacts when multiple-pollutant emission scenarios are considered.[116]

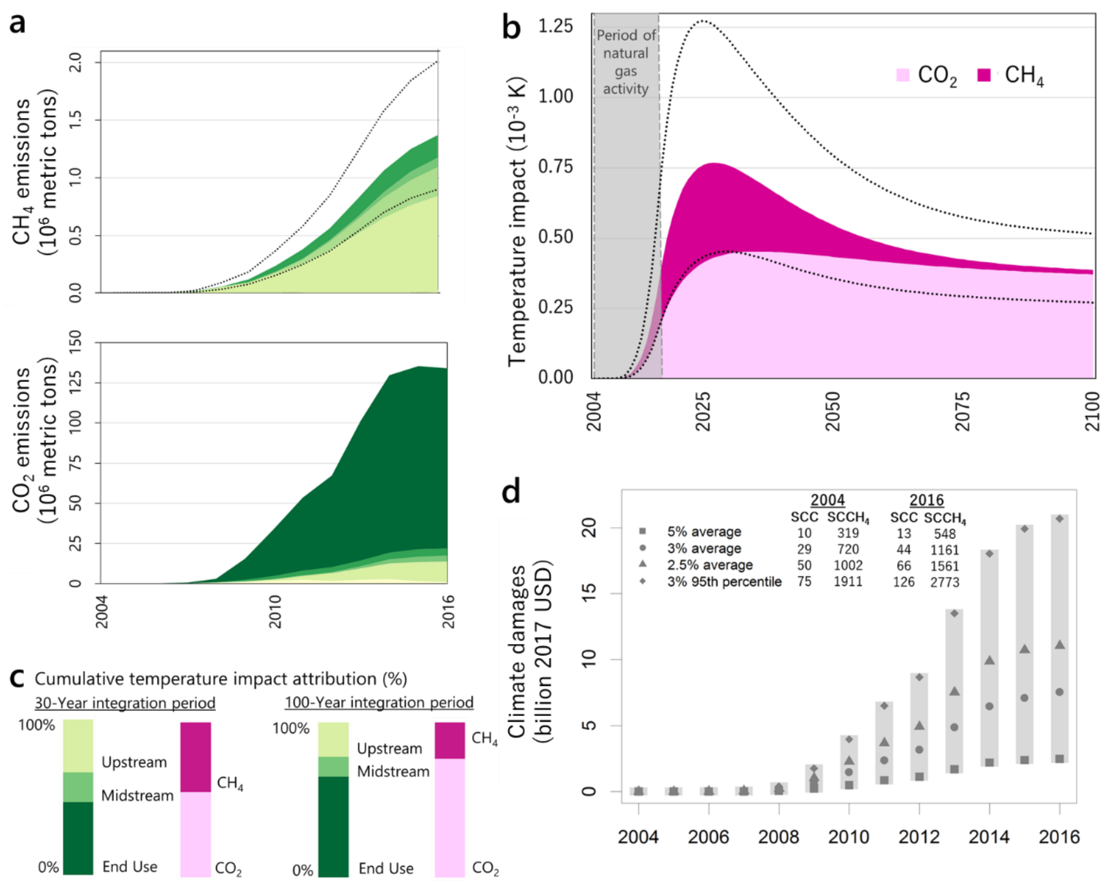

Figure 6. Climate change impacts across the natural gas supply chain from 2004 to 2016. (a) Annual CH_4 and CO_2 emissions from sources within Pennsylvania, Ohio, and West Virginia. Dotted black lines depict emissions under low and high scenarios. (b) Annual temperature impact indicating contributions from CH_4 and CO_2 emissions. Dotted black lines depict temperature impact under low and high scenarios. (c) Attribution of cumulative temperature impact over 30- and 100-year integration periods (beginning in 2004) by segment of the natural gas supply chain and GHG species. (d) Climate change damages under different social cost of carbon (SCC) and social cost of methane ($SCCH_4$) values. SCC and $SCCH_4$ estimates vary by year, and estimates for 2004 and 2016 (in units of $ per metric ton) are provided for reference. All values unless otherwise noted are based on baseline scenario assumptions.

3.5 Employment impacts

Countervailing the environmental and public health impacts are the effects of the natural gas sector on local economic conditions, including local labor markets. Several empirical studies demonstrate that natural gas activity may: impact local labor demand within the sector; have spillover effects on the non-resource economy[86,88]; and alter the distribution of income[117], poverty rates[88,118], and educational attainment[119]. Here, we focus on employment effects, including the marginal effect as it relates to upstream activity and the aggregate effect in producing counties over time.

We observe a positive and statistically significant employment effect from natural gas activity, a finding that is robust across multiple model specifications (SI Section 5, Tables S17-25). A model specification that includes both spud and producing wells as predictor variables, provides an intuitive result, differentiating the jobs directly or indirectly associated with drilling activities (16 jobs per spud well) and ongoing production operations (4 jobs per producing well). In an alternative model specification, we find a mean effect size of 5 job-years per billion cubic feet (bcf) of natural gas production. Other empirical labor market studies for the U.S. natural gas sector estimate slightly higher effects (6 to 16 job-years per producing well, 7 to 19 job-years per bcf), which potentially can be explained by the differing geographic scope and expanded number of years of data incorporated in this study.[86–88,120] We additionally observe decreasing marginal employment effects from natural gas activity over time, with 75% fewer job-years supported per bcf of production after 2012; the intuition is that learning occurs and the industry becomes more efficient and automated over time. [Refer to SI Section 5 for a detailed description of the fixed effects modeling including model formulation and additional specifications].

Employment effects are inclusive of not only those within the natural gas sector, but also spillover into other sectors, which can both positively and negatively impact local economies. We estimate that each natural gas industry job is associated with 1.9 jobs outside of the resource sector, consistent with other studies that similarly observe relatively minor multiplier effects at the county level (1 to 1.4).[86,88] A possible explanation for the low multiplier effect is that jobs associated with extractive industries, such as drilling crews, are often held by transient workers because long-term residents in rural communities may not have the requisite skills and training. In addition, production firms are largely based outside of the county of production and as such may source supplies and equipment elsewhere. A limitation to the empirical approach used is that regional employment effects may be greater than those observable at the county level.[121] Other studies show that there is minimal evidence of the negative implications of spillovers, including the phenomenon of the "natural resource curse," often observed in relation to extractive industries, wherein resource-rich areas tend to grow more slowly than resource-poor areas.[86,122]

We simulate aggregate employment over time, combining natural gas activity data with marginal employment effects. For clarity, we use the metric job-year, which is a full- or part-time job over a single year, not a sustained job over multiple years or a career. We find that the direct and induced job-years supported by the shale gas sector over the period 2004 to 2016 was 460,000 (95% CI ±30%). As shown in Figure 3a, the trace of employment over time largely mimics the transiency of jobs and the boom-and-bust cycles of other extractive resources, with the annual number of jobs-years peaking in 2014 at 76,000 (95% CI ±30%), reflecting both increasing production and rapid growth and subsequent decline in drilling activity. As highlighted in Figure 7c-d, the spatial distribution of jobs largely aligns with the intensity of drilling and production, with most jobs concentrated in rural (54%) and mixed rural-urban (31%) areas. The employment associated with shale gas activity comprised less than 1% of total employment in urban or low producing counties to over 60% of total employment in rural and high producing counties, and many jobs are in counties with thin labor markets. We also estimate cumulative earnings of $21B ($8B to $33B), based on aggregate employment effects and annual average per capita earnings by county (see Figure 7b); this method does not segment employment and average earnings by sector, account for executive compensation, and differentiate between full- and part-time employment.

These employment estimates only account for direct and spillover jobs associated with upstream activities and potentially co-located midstream and end use segments. A recent study of electricity generation related employment effects estimates 0.11 job-years are supported per Gigawatt hour (GWh), which includes direct natural gas generation-related employment associated with construction, installation, manufacturing, operating, maintenance, and fuel processing[123]; this suggests an additional 46,000 job-years may be associated with electricity generation using shale gas as a feedstock in the Appalachian basin. For comparison, wind (0.17 job-years/GWh) and solar photovoltaics (0.87 job-years/GWh) have higher generation-related employment.

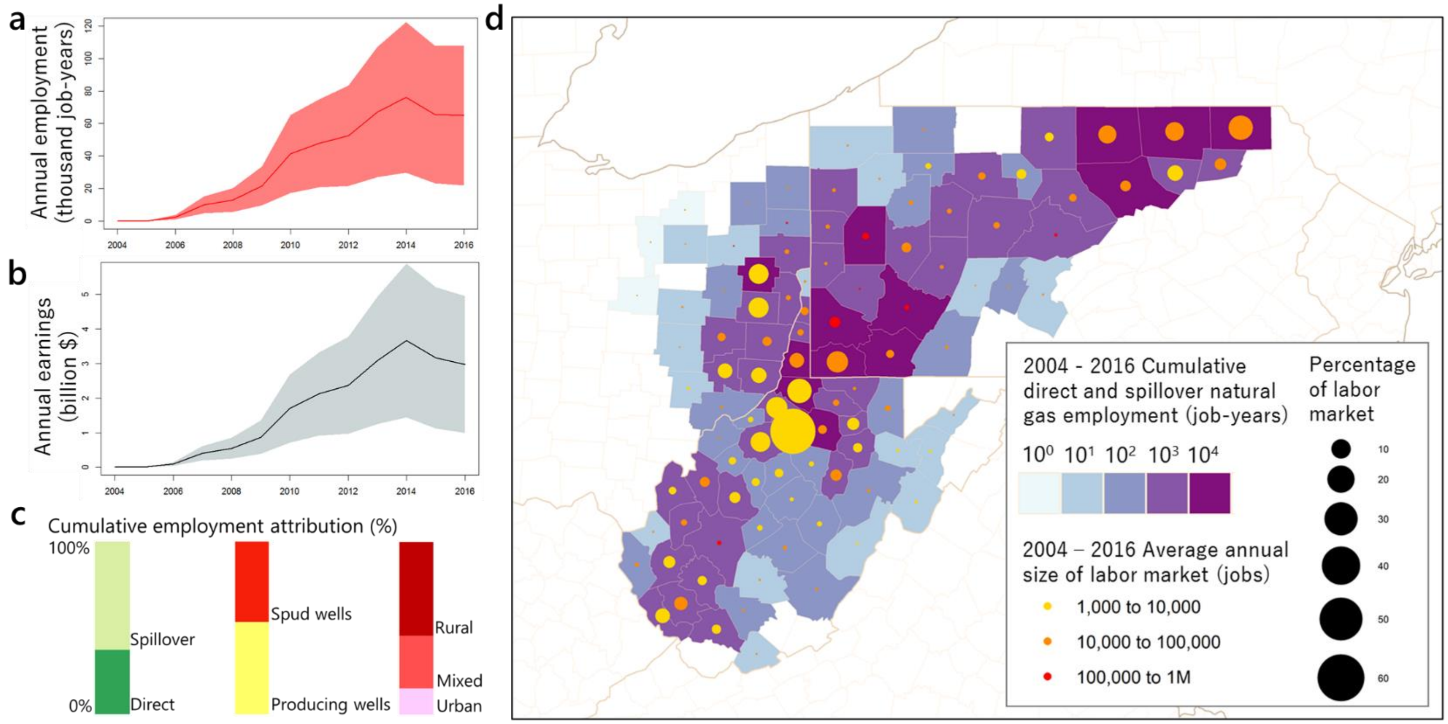

Figure 7. Employment impacts across the natural gas supply chain from 2004 to 2016. (a) Annual employment from natural gas activity, including direct effects within the natural gas sector and spillover effects into other sectors. Based on marginal employment effects and actual natural gas activity. Line represents mean simulated employment, and shaded regions are 95% confidence intervals (based on clustered standard errors). (b) Annual earnings from natural gas activity, including direct and spillover effects. Based on marginal employment effects, actual natural gas activity, and reported annual earnings per capita by county. Line represents mean simulated earnings, and shaded region is 95% confidence intervals. (c) Attribution of employment from natural gas activity by sector (direct and spillover), natural gas activity (spud and producing wells), and rural-urban regions. The sector attribution is based on fixed effects model 16 fit with alternative dependent variables, total employment and mining employment (a proxy for direct jobs within the natural gas sector). The natural gas activity attribution is based on 2004-2016 cumulative employment. The rural, mixed rural-urban, and urban regions were classified based on the 2010 U.S. Census county rurality level tertiles and based on 2004-2016 cumulative employment. (d) Spatial distribution of 2004-2016 cumulative employment by county. Color shading of counties represents cumulative employment from natural gas activity. The color of the dots at county centroids represents the average annual size of the labor market, and the size of the dots represents the percentage of employment from natural gas activity out of total employment. All results are based on mean marginal employment effects from fixed effects model specification 16.

3.6 Tradeoffs between air, climate, and employment impacts

We observe that after drilling peaked in 2013, air quality and employment impacts began to decline, whereas climate impacts are projected to continue to increase for another decade, all else equal, and persist over time horizons greatly exceeding the period of natural gas activity (Figure 8). Pairwise comparisons of physical impacts reveal the implied tradeoffs from natural gas development decisions. Job-years per premature mortality is both meaningful and understandable, whereas changes in average global temperature per job-year is less so. Based on mid-range cumulative estimates of premature mortality from air pollution and employment, the implied tradeoff is 217 job-years per premature mortality at a systems level. This tradeoff varies spatially among producing counties, ranging from 1 to 16,000 job-years per premature mortality.

Weighting between these impacts is (in part) normative, and monetization is a common weighting approach. We monetize impacts for comparison of an approach that relates physical impacts to a traditional benefit-cost analysis framing. Comparisons of monetized impacts, although subject to uncertainty, show a variable relationship between impacts from 2004 to 2016. As drilling activity declined, climate damages began exceeding air quality damages; although integrated climate damages obscure when impacts are realized. Based on mid-range estimates, cumulative employment impacts ($21B) are less than air quality ($23B) and climate change damages ($34B); however, no impact area stochastically dominates another, namely due to the vast range of VSL and social costs of methane and carbon that are considered. The breakeven VSL at which mean monetized employment equates to cumulative air quality damages over the development horizon is $7.7M. Similarly, the breakeven social cost of carbon at which mean monetized employment equates to cumulative climate damages over the development horizon is $25 per metric ton.

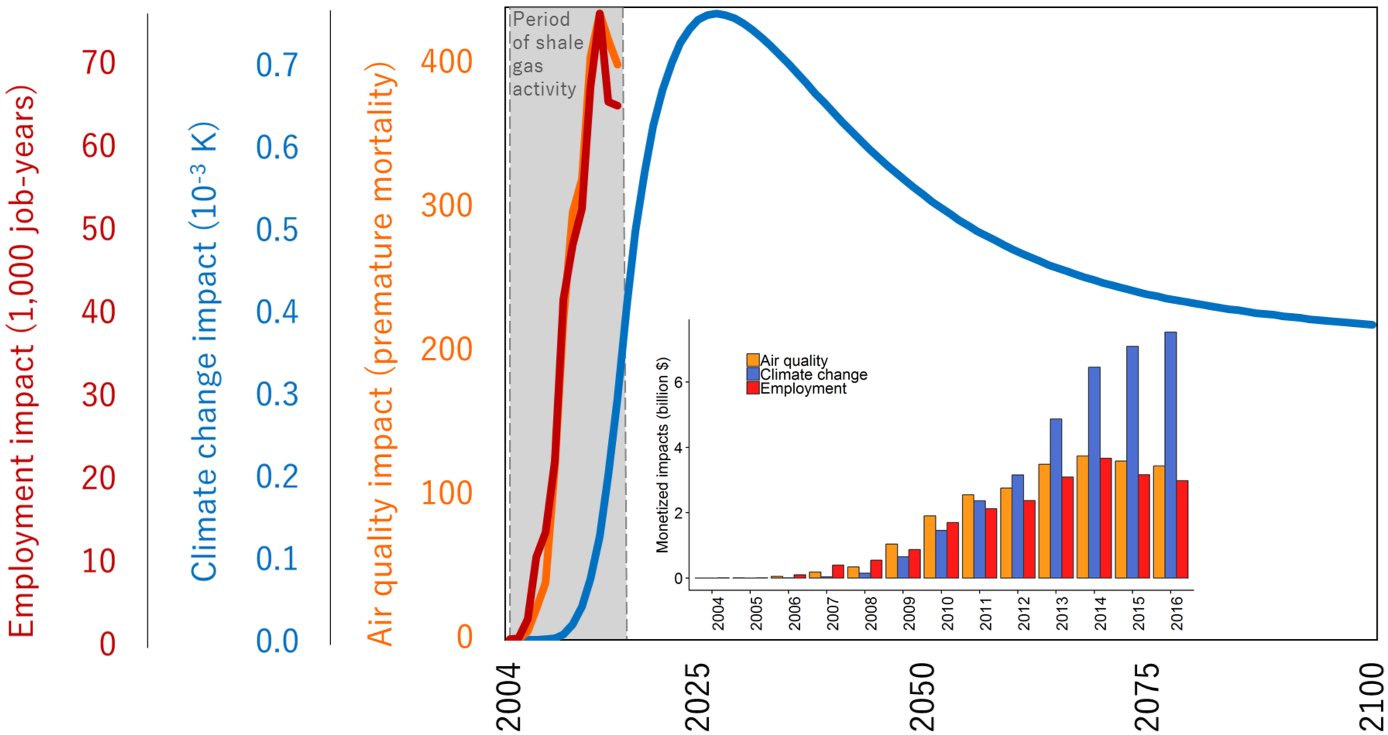

Figure 8. Air, climate, and employment impacts over time. Air quality impacts based mean annual mortality estimates across six model specifications (under different RCMs and C-R relationships). Climate impacts under baseline scenario assumptions. Employment impacts based on marginal effects from fixed effects specification 16. Vertical axes are standardized to range from 0 to the maximum impact value. The inset presents monetized impacts under mid-range assumptions: mean simulated VSL of $8.5M; SCC ranging from $29 (2004) to $44 (2016) per CO2 metric ton; and SCCH4 ranging from $720 (2004) to $1161 (2016) per CH4 metric ton.

3.7 Policy implications

Estimates of cumulative air, climate, and employment impacts and interdependencies have implications for decision making and policy design. Here, we focus on policies to address air quality and climate externalities. Absent comprehensive policy, private firms across the supply chain have not faced the full costs of natural gas development, and the public has effectively subsidized waste disposal costs, with local communities bearing the largest share of costs. Thus, production and activity across the supply chain were greater than it would have been had firms internalized environmental costs.

We find that a severance tax on production of $2 per thousand cubic feet (mcf) (+172%/-76%) should be levied to account for air quality and climate damages, as shown in Figure 9. The proposed Pigouvian tax is derived based on historical cumulative damages across the supply chain and production rates in the Appalachian basin. A similar approach can be used to derive appropriate severance tax rates in other producing regions, such as the Permian basin, Anadarko basin, Eagle Ford play, and Barnett play. The trajectory of production, changing system efficiencies, and other environmental policies are also relevant considerations in setting a future severance tax rate.

We may further derive spatially and temporally differentiated severance tax rates that account for the dispersed and changing nature of impacts, as well as, the differing mandates of federal, state, or local executive, judicial, and legislative authorities. While a tax that incorporates spatio-temporal variation may be optimal, the complexity of implementing such a policy may be prohibitive. Similarly, a tax rate can be estimated for different segments of the supply chain or based on consumption rather than production. Depending on where along the supply chain a tax is applied, ostensibly different price signals may be established for upstream and end use producers and consumers, in addition to other policy makers and investors that are considering competing energy technologies.

For context, wellhead prices, as reported in Security and Exchange Commission 10k filings of the top producing and publicly traded firms operating in the Appalachian basin, have been volatile (e.g., approximately $4 per mcf in 2014 and $2 per mcf in 2016) and similar in magnitude as the proposed severance tax rate. In addition, there is a large disparity between the proposed tax rate and the existing effective severance tax or impact fee rate ($0.08 per mcf), which we estimate based on past severance tax or impact fee revenue in Pennsylvania, Ohio, and West Virginia and production rates from 2004 to 2016; climate and air quality damages are seemingly not currently factored into existing state severance tax and impact fee structures.

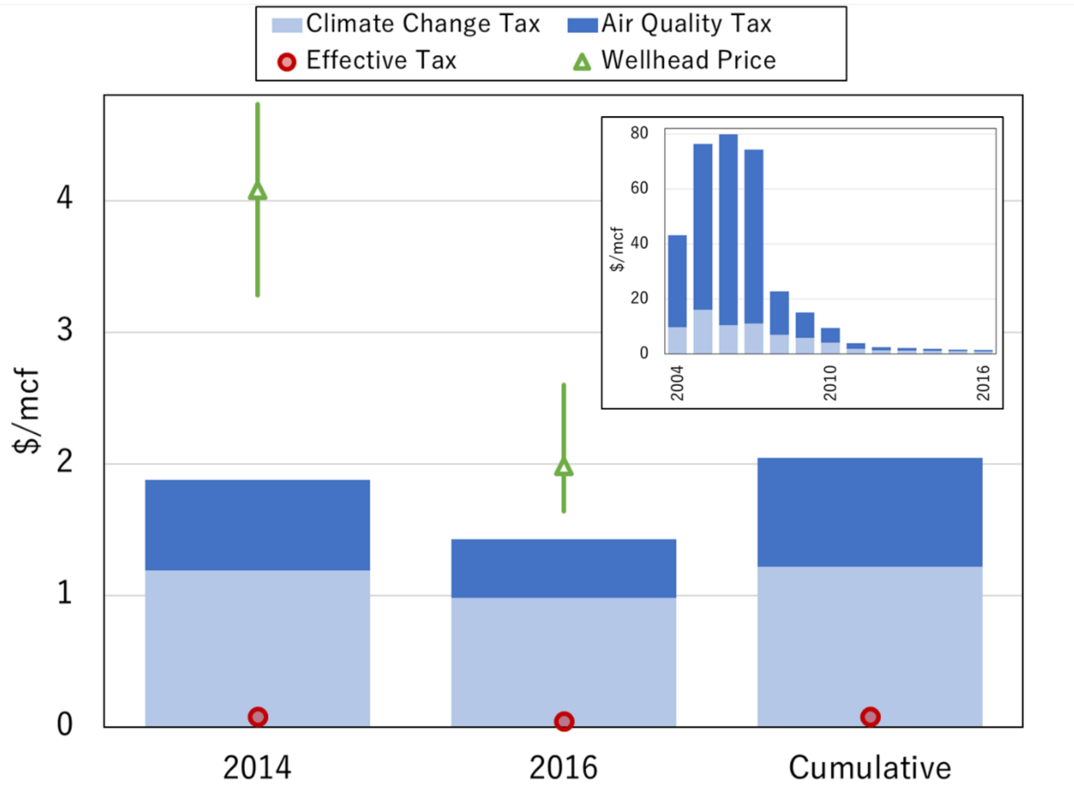

Figure 9. Climate change and air quality production tax rates. Climate change and air quality tax rates are estimated based on annual or 2004 to 2016 cumulative impacts and production. Climate change impacts are based on mid-range social cost of carbon and methane values (3% mean discount rate), and air quality impacts are based on the mean simulated VSL ($8.8M). The wellhead prices are based on Security and Exchange Commission 10-K filings of the top nine public producing firms in the Appalachian basin (as of 2017); the triangle points represent the production-weighted average price, and the lines represent the range. The effective "tax" is based on the aggregate annual Pennsylvania impact fee, Ohio severance tax, and West Virginia severance tax revenues divided by annual production.

3.8 Conclusions

Rapidly increasing natural gas activity in the Appalachian basin corresponds with increasing impacts, with short-lived air quality and employment impacts tracking the boom-and-bust cycle and climate impacts persisting for generations well beyond the period of natural gas activity. With evolving regulation and industry practices, there are efficiencies with respect to emissions and jobs supported by natural gas activity. We further find that more than half of cumulative premature mortality is within source emissions states, while transboundary impacts are concentrated in populous coastal regions in the Northeast. Most premature mortality is associated with end uses, while upstream and midstream segments also account for a substantial portion of impacts. With respect to climate change impacts, the magnitude of methane emissions across the supply chain produces temperature impacts nearly equivalent to that of CO_2 over a 30-year time horizon, and over longer integration periods, the warming impact from CO_2 dominates that of CH_4. Employment is concentrated in rural areas with thin labor markets with minimal spillover effects, but ancillary evidence suggests that jobs are held by transient, higher skilled workers.

This study focuses on some aspects of air, climate, and employment impacts, and it stands to be expanded to other outcomes and impact areas such as water quality and ecological health. Furthermore, the scope of this analysis is solely inclusive of the natural gas supply chain, and an analogous approach that captures cumulative impacts can be applied to and comparisons can be made with other energy sources and technologies. A prevailing comparison, especially with respect to climate change and air quality, has been natural gas versus coal. While the coal boom-and-bust cycles may provide a useful analogue, the observed decreases in CO_2 intensity of U.S. electricity production reflective of declining coal generation and corresponding increases in natural gas and wind[32], perhaps signal the fading relevance of natural gas' comparative advantage over coal with respect to the U.S. energy system. Arguably more relevant comparisons, both currently and prospectively, are between natural gas and renewables. Transitioning from natural gas to renewables would seemingly reduce or eliminate cumulative impacts with respect to climate change and air quality, while the relative employment effect is not as well understood. Furthermore, the argument that natural gas may serve as a bridge fuel, which is in part premised on its comparative climate advantage over coal and cost advantage over renewables and other energy technologies, is unsupported if natural gas prices are not reflective of the actual economics for producing firms or of climate and air quality damages.

This study supports a need for multivalent policies that consider interdependencies and cumulative impacts over the life of the play and across the supply chain. This includes cross-media policies that account for processes that emit species that may lead to air quality and climate impacts at varying spatial and temporal scales. This also includes both fiscal and environmental policy which may realign incentives to produce and develop the broader energy system in a way that considers longer time horizons.

Having a retrospective understanding of the cumulative impacts and implied tradeoffs of natural gas systems, we can ask larger normative and technical questions that are becoming more salient. Considering air, climate, and labor market objectives, are we producing and consuming too much? Are we approaching technological lock-in? What happens when the natural gas resources are played out?

4 Future Natural Gas System Development Pathways Incorporating Air, Climate, and Employment Objectives

Natural gas system development is driven by the complexity inherent in physical systems and the influence of a myriad of diverse, interacting stakeholders with heterogeneous preferences. The purpose of this study is to distill a portion of this complexity and provide insight into how the energy system theoretically could develop if other objectives that are often the subject of public discourse and concern, such as jobs, climate change, and health effects, influence the decision-making process. While environmental and employment objectives are conflicting if we follow a natural gas pathway consistent with the status quo, the collection of siting, emissions abatement, and renewable integration policies can resolve and reverse these conflicts. There are pathway dependencies between these policies, and delaying implementation only amplifies the cumulative, negative tradeoffs between employment and environmental objectives.

4.1 Introduction

Climate change has resulted in impacts to natural and human systems, and avoidance of serious impacts requires rapid and deep reductions in greenhouse gas (GHG) emissions and associated policies.[1–3] The Paris Agreement, adopted in December 2015, aims to reduce GHG emissions to a level consistent with limiting the average global temperature to well below 2°C above pre-industrial levels.[2,4] Studies have shown that a deeply decarbonized U.S. energy system can provide equivalent energy services as the status quo; however, deep decarbonization entails unprecedented and transformational changes in the energy system, including drastically increasing efficiency of end uses, decreasing the carbon intensity of electricity, and switching end uses from direct combustion of fossil fuels to electricity.[5–7]

Beyond technoeconomic challenges of decarbonization, accelerated transitions also depend upon widespread social acceptance and balancing potentially countervailing societal needs and objectives.[8,9] Given the coupling of air and climate impacts related to fossil fuel production and use, transitioning to a low-carbon energy also contributes to reducing near-term air pollution impacts.[10] Labor markets are also inherently linked to the transition of the energy system, with a redistribution of jobs away from carbon-intensive energy sources, such as coal, to green infrastructure development.[6,11] There are also distributional effects of climate change and mitigation.[12,13]

There is vast array of energy systems models that vary in complexity and sectoral, spatial, and temporal scope and application, but few models focus on energy system transitions and deep decarbonization.[5] There are several common elements of existing model architectures that limit their application to future energy system planning. First, many models use illustrative systems and networks that are not reflective of critical real world dynamics.[8] Understanding the costs of energy transitions is fundamental, and thus, models almost exclusively employ monetized objectives, adopt an economic efficiency or utilitarian framing, and do not consider impacts.[14–16] Many models focus on electric power systems, and do not consider the broader energy sector.[17] Finally, few energy system optimization model incorporate long time horizons, and multiperiod objectives.

In this study, we develop a multiobjective energy system optimization model that incorporates cumulative air, climate, and employment impacts objectives. This purpose of this model is to generate future natural gas system development pathways, with respect to the timing, location, and magnitude of natural gas activity and low-carbon interventions, which may provide insight for energy system planning and policy design. The modeling architecture is distinct from many existing energy systems optimization models because of the unique cumulative impact objectives, spatial resolution, long time horizons, air and climate impact (rather than emission and monetary) measures leveraging reduced complexity models, and parameterization and structure reflecting a real (rather than illustrative) natural gas system. We model alternative, theoretical natural gas development pathways by optimizing spatially and temporally explicit air quality, climate change, and employment impacts with respect to decisions regarding the magnitude, timing, and location of preproduction, production, and industrial, residential, commercial, and electric consumption. We additionally specify policy scenarios, including infrastructure siting and planning, emissions abatement, and renewable integration, and demonstrate the relative effect and pathway dependence of implementing these policies. We illustrate the approach by analyzing natural gas development pathways in the Appalachian basin.

4.2 Model formulation

To facilitate natural gas system planning, we formulate a multiobjective linear optimization model that incorporates socioeconomic and environmental objectives. Herein, we provide a general formulation for natural gas systems, as well as, parameterize the model for the case of natural gas development in the Appalachian basin. We model both retrospective and future cases, which provide insight regarding how the system could have been developed in the past and can be developed in the future accounting for air quality, climate change, and employment objectives. We additionally specify modifications in which we integrate emissions abatement and renewables to displace natural gas. The optimization model is implemented in the General Algebraic Modeling System (GAMS) and uses CPLEX to solve the linear program. Table 2 includes indices, decision variables, and parameters.

4.2.1 Decision variables

We incorporate decision variables related to six types of upstream and end use natural gas activities – the number of producing and spud wells, and residential, commercial, industrial, and electricity consumption volumes. We do not include midstream processes, including gathering, processing, transmission, storage, and distribution, in this version of the model.

The decision variables, which are non-negative and continuous, reflect the magnitude of natural gas activity for each county i and year t. We consider retrospective ($t = 2005, .., 2016$) and future ($t = 2017, .., 2030$) natural gas activity time horizons, and natural gas activity within 210 counties in Pennsylvania, Ohio, and West Virginia. Given that impacts may be spatially and temporally disperse from natural gas activity, we assume a longer impact horizon ($s = 2005, .., 2100$) to accommodate delayed and persistent climate change impacts, and we consider impact counties to include all of those within the continental U.S. to allow for spatial transport of air pollutant impacts.

In the retrospective case, we assume that cumulative production and consumption are fixed to historical levels, but that the timing and location of production and the location of additional electric generation beyond 2004 levels are variable. In the future case, we assume that consumption is spatially and temporally fixed to projected future levels, with the exception of the timing and location of additional electric generation and industrial end use beyond 2016 levels which are assumed to be variable. We also allow the magnitude, timing, and location of future production to be variable, while accounting for the residual production from historical wells.

Table 2. Indices, decision variables, and parameters.

Indices			
$i = \{1, \dots, \mathcal{I}\}$	county of natural gas activity	$t = \{1, \dots, \mathcal{T}\}$	year of natural gas activity
$j = \{1, \dots, \mathcal{J}\}$	county of impact	$u = \{1, \dots, \mathcal{U}\}$	well production year
$s = \{1, \dots, \mathcal{S}\}$	year of impact		
Decision variables			
$x_{i,t}^{commercial}$	commercial natural gas volume (mmcf)	$x_{i,t}^{producing}$	producing wells
$x_{i,t}^{electric}$	electric generation natural gas volume (mmcf)	$x_{i,t}^{residential}$	residential natural gas volume (mmcf)
$x_{i,t}^{industrial}$	industrial natural volume (mmcf)	$x_{i,t}^{spud}$	spud wells
Parameters (primary)			
$\mathrm{a}_{i,j,t}^{\mathrm{AIR,COMMERCIAL}}$	marginal premature mortalities from commercial end use (per mmcf)	$\mathrm{b}_{t}^{\mathrm{CLIMATE,PREPRODUCTION}}$	marginal climate damages from preproduction ($ per spud well)
$\mathrm{a}_{i,j,t}^{\mathrm{AIR,ELECTRIC}}$	marginal premature mortalities from electric generation end use (per mmcf)	$\mathrm{b}_{t}^{\mathrm{CLIMATE,PRODUCTION}}$	marginal climate damages from production ($ per producing well)
$\mathrm{a}_{i,j,t}^{\mathrm{AIR,INDUSTRIAL}}$	marginal premature mortalities from industrial end use (per mmcf)	$\mathrm{b}_{t}^{\mathrm{CLIMATE,RESIDENTIAL}}$	marginal climate damages from residential end use ($ per mmcf)
$\mathrm{a}_{i,j,t}^{\mathrm{AIR,PREPRODUCTION}}$	marginal premature mortalities from preproduction (per spud well)	$\mathrm{C}_{\mathrm{i,t}}^{\mathrm{COMMERCIAL}}$	commercial consumption (mmcf)
$\mathrm{a}_{i,j,t}^{\mathrm{AIR,PRODUCTION}}$	marginal premature mortalities from production (per producing well)	$\mathrm{C}_{\mathrm{i,t}}^{\mathrm{ELECTRIC}}$	electric generation natural gas consumption (mmcf)
$\mathrm{a}_{i,j,t}^{\mathrm{AIR,RESIDENTIAL}}$	marginal premature mortalities from residential end use (per mmcf)	$\mathrm{C}_{i,t}^{\mathrm{INDUSTRIAL}}$	industrial consumption (mmcf)
$\mathrm{a}_{s,t}^{\mathrm{CLIMATE,COMMERCIAL}}$	marginal temperature impact from commercial end use (milliKelvin per mmcf)	$\mathrm{C}_{i,t}^{\mathrm{RESIDENTIAL}}$	residential consumption (mmcf)
$\mathrm{a}_{s,t}^{\mathrm{CLIMATE,ELECTRIC}}$	marginal temperature impact from electric generation end use (milliKelvin per mmcf)	$\mathrm{p}_{i,u}$	annual production per well (mmcf)

Table 2. Indices, decision variables, and parameters (continued).

$a_{s,t}^{CLIMATE,INDUSTRIAL}$	marginal temperature impact from industrial end use (milliKelvin per mmcf)	$P^{CUMULATIVE}$	cumulative production (mmcf)
$a_{s,t}^{CLIMATE,PREPRODUCTION}$	marginal temperature impact from preproduction (milliKelvin per spud well)	R	reserves (mmcf)
$a_{s,t}^{CLIMATE,PRODUCTION}$	marginal temperature impact from production (milliKelvin per producing well)	VSL	value of a statistical life ($)
$a_{s,t}^{CLIMATE,RESIDENTIAL}$	marginal temperature impact from residential end use (milliKelvin per mmcf)	W_t^{MAX}	annual maximum number of spud wells
$a^{EMPLOY,PREPRODUCTION}$	marginal employment from preproduction (job-years per spud well)	W_i^{MAX}	county well density (wells)
$a^{EMPLOY,PRODUCTION}$	marginal employment from production (job-years per producing well)	$WAGES_{i,t}$	average wage ($ per job)
$b_t^{CLIMATE,COMMERCIAL}$	marginal climate damages from commercial end use ($ per mmcf)	$Z^{AIR\ QUALITY}$	cumulative air quality impacts (premature mortalities)
$b_{s,t}^{CLIMATE,ELECTRIC}$	marginal climate damages from electric generation end use ($ per mmcf)	$Z^{CLIMATE\ CHANGE}$	cumulative global temperature impact (milliKelvin-years)
$b_t^{CLIMATE,INDUSTRIAL}$	marginal climate damages from industrial end use ($ per mmcf)	$Z^{EMPLOYMENT}$	cumulative employment impacts (job-years)
	Parameters (intermediate)		
$AGTP_{s,t}^{CH4}$	absolute global temperature potential from CH_4 emissions (milliKelvin per metric ton)	EF_t^{VOC}	VOC emissions factor (metric tons per unit natural gas activity)
$AGTP_{s,t}^{CO2}$	absolute global temperature potential from CO_2 emissions (milliKelvin per metric ton)	$M_{i,j,t}^{PM2.5}$	marginal premature mortalities from PM2.5 emissions (per metric ton)
EF^{CH4}	CH_4 emissions factor (metric tons per unit natural gas activity)	$M_{i,j,t}^{NOX}$	marginal premature mortalities from NOx emissions (per metric ton)
EF^{CO2}	CO_2 emissions factor (metric tons per unit natural gas activity)	$M_{i,j,t}^{VOC}$	marginal premature mortalities from VOC emissions (per metric ton)
$EF_t^{PM2.5}$	$PM_{2.5}$ emissions factor (metric tons per unit natural gas activity)	SCC_t	social cost of carbon ($ per metric ton)
EF_t^{NOX}	NO_x emissions factor (metric tons per unit natural gas activity)	$SCCH4_t$	social cost of methane ($ per metric ton)

4.2.2 Air quality objectives

The air quality objective function is to minimize cumulative impacts, alternatively specified in terms of premature mortalities and monetized damages, resulting from primary fine particulate matter ($PM_{2.5}$) and secondary $PM_{2.5}$ formed from the atmospheric oxidation of nitrogen oxides (NO_x) and volatile organic compounds (VOCs) emissions. The general form of the objective function in terms of premature mortality is:

$$\text{Min } \mathrm{Z}^{\text{AIR QUALITY}} = \sum_{i,j,t} \Big(\mathrm{a}_{i,j,t}^{\text{AIR,PREPRODUCTION}} \cdot x_{i,t}^{spud} + \mathrm{a}_{i,j,t}^{\text{AIR,PRODUCTION}} \cdot x_{i,t}^{producing} + \mathrm{a}_{i,j,t}^{\text{AIR,INDUSTRIAL}} \cdot x_{i,t}^{industrial} + \mathrm{a}_{i,j,t}^{\text{AIR,COMMERCIAL}} \cdot x_{i,t}^{commercial} + \mathrm{a}_{i,j,t}^{\text{AIR,RESIDENTIAL}} \cdot x_{i,j,t}^{residential} + \mathrm{a}_{i,j,t}^{\text{AIR,ELECTRIC}} \cdot x_{i,j,t}^{electric} \Big) \qquad (1)$$

where the coefficients $\mathrm{a}_{i,j,t}^{\text{AIR}}$ are the premature mortalities in receptor county j associated with emissions in source county i per unit of natural gas activity. Objective coefficients ($\mathrm{a}_{i,j,t}^{\text{AIR}}$) are derived using emission factors ($\mathrm{EF}_t^{\text{PM2.5}}, \mathrm{EF}_t^{\text{NOX}}, \mathrm{EF}_t^{\text{VOC}}$) and marginal premature mortalities per unit of emission ($\mathrm{M}_{i,j,t}^{\text{PM2.5}}, \mathrm{M}_{i,j,t}^{\text{NOX}}, \mathrm{M}_{i,j,t}^{\text{VOC}}$). The coefficients $\mathrm{a}_{i,j,t}^{\text{AIR}}$, which we specify for each natural gas process or segment, are of the general form:

$$\mathrm{a}_{i,j,t}^{\text{AIR}} = \mathrm{M}_{i,j,t}^{\text{PM2.5}} \cdot \mathrm{EF}_t^{\text{PM2.5}} + \mathrm{M}_{i,j,t}^{\text{NOX}} \cdot \mathrm{EF}_t^{\text{NOX}} + \mathrm{M}_{i,j,t}^{\text{VOC}} \cdot \mathrm{EF}_t^{\text{VOC}} \qquad (2)$$

We account for air pollutant emissions from preproduction processes (including drilling, hydraulic fracturing, well completion, trucking), production processes (including wellhead compressors, condensate tanks, production fugitives, etc.), and end use combustion. Emission factors are derived based on Mayfield et al. (forthcoming), the US Environmental Protection Agency (EPA) National Emissions Inventory (NEI), and the US EPA Continuous Emissions Monitoring System (CEMS).[124–126]

We use the source-receptor reduced complexity model (RCM), Air Pollution Emission Experiments and Policy model (Version 3) (AP3)[91,92], which generates estimates of pollution-induced premature mortalities in downwind receptors associated with emissions from source locations. Premature mortality estimates are sensitive to the relationship between pollutant concentration and health response; therefore, we parametrically vary the concentration-response (C-R) relationship based on the American Cancer Society (ACS) (base assumption) and Harvard Six Cities (H6C) studies.[95,96] To account for time-varying marginal mortality, we apply an annual population adjustment based on historical, annual, county-level population reported by the U.S. Census Bureau, and state-level population projections reported by the US Centers for Disease Control and Prevention (CDC).[127,128]

The general form of the objective function in terms of monetized air quality damages (US$ 2017) is:

$$\text{Min } Z^{\text{AIR QUALITY}} = \text{VSL} \cdot \sum_{i,j,t}\left(a_{i,j,t}^{\text{AIR,PREPRODUCTION}} \cdot x_{i,t}^{spud} + a_{i,j,t}^{\text{AIR,PRODUCTION}} \cdot x_{i,t}^{producing} + a_{i,j,t}^{\text{AIR,INDUSTRIAL}} \cdot x_{i,t}^{industrial} + a_{i,j,t}^{\text{AIR,COMMERCIAL}} \cdot x_{i,t}^{commercial} + a_{s,t}^{\text{AIR,RESIDENTIAL}} \cdot x_{i,j,t}^{residential} + a_{i,j,t}^{\text{AIR,ELECTRIC}} \cdot x_{i,j,t}^{electric}\right) \quad (3)$$

where VSL is the value of a statistical life, a commonly used measure of the dollar value of small changes in mortality risk experienced by a large number of people. We use a VSL of $8.5M (base assumption), the mean value recommended in the U.S. Environmental Protection Agency guidance.[97] We also test the sensitivity of decisions to VSL, using a range of $0 to $100M.

4.2.3 Climate change objectives

The climate change objective function is to minimize cumulative climate change impacts, alternatively specified in terms of global temperature change and monetized damages, resulting from emissions of greenhouse gases (GHG), including methane (CH_4) and carbon dioxide (CO_2). The general form of the objective function in terms of global temperature change (milliKelvin-years) is:

$$\text{Min } Z^{\text{CLIMATE CHANGE}} = \sum_{i,s,t}\left(a_{s,t}^{\text{CLIMATE,PREPRODUCTION}} \cdot x_{i,t}^{spud} + a_{s,t}^{\text{CLIMATE,PRODUCTION}} \cdot x_{i,t}^{producing} + a_{s,t}^{\text{CLIMATE,INDUSTRIAL}} \cdot x_{i,t}^{industrial} + a_{s,t}^{\text{CLIMATE,COMMERCIAL}} \cdot x_{i,t}^{commercial} + a_{s,t}^{\text{CLIMATE,RESIDENTIAL}} \cdot x_{i,t}^{residential} + a_{s,t}^{\text{CLIMATE,ELECTRIC}} \cdot x_{i,t}^{electric}\right) \quad (4)$$

where the coefficients $\mathrm{a}_{s,t}^{\mathrm{CLIMATE}}$ are the global temperature change per unit of natural gas activity associated with emissions in year t and climate response in year s. Global temperature change per unit activity is estimated using a convolution of GHG emission factors ($\mathrm{EF}^{\mathrm{CH4}}$, $\mathrm{EF}^{\mathrm{CO2}}$) and the average global temperature potential ($\mathrm{AGTP}_{s,t}$).[98–103] $\mathrm{AGTP}_{s,t}$ is a function of the temperature response to radiative forcing which, in turn, is due to a marginal pulse emission, both of which are parameterized based on more complex models that explicitly include physical and chemical processes; see Mayfield et al. (forthcoming) for additional details.[99] The coefficients $\mathrm{a}_{s,t}^{\mathrm{CLIMATE}}$, which we specify for each natural gas process or segment, are of the general form:

$$\mathrm{a}_{\mathrm{s,t}}^{\mathrm{CLIMATE}} = \mathrm{AGTP}_{s,t}^{\mathrm{CH4}} \cdot \mathrm{EF}^{\mathrm{CH4}} + \mathrm{AGTP}_{s,t}^{\mathrm{CO2}} \cdot \mathrm{EF}^{\mathrm{CO2}} \tag{5}$$

We account for GHG emissions from preproduction processes (including drilling, hydraulic fracturing, well completion, trucking), production processes (including wellhead compressors, condensate tanks, production fugitives, etc.), and end use combustion. Emission factors are derived based on Mayfield et al. (forthcoming),and the US EPA CEMS.[124,126]

The general form of the objective function in terms of monetized damages (US$ 2017) is:

$$\mathrm{Min}\ \ \mathrm{Z}^{\mathrm{CLIMATE\ CHANGE}} = \mathrm{b}_{t}^{\mathrm{CLIMATE,PREPRODUCTION}} \cdot x_{i,j,t}^{spud} + \mathrm{b}_{t}^{\mathrm{CLIMATE,PRODUCTION}} \cdot x_{i,j,t}^{producing} + \mathrm{b}_{t}^{\mathrm{CLIMATE,INDUSTRIAL}} \cdot x_{i,j,t}^{industrial} + \mathrm{b}_{t}^{\mathrm{CLIMATE,COMMERCIAL}} \cdot x_{i,j,t}^{commercial} + \mathrm{b}_{t}^{\mathrm{CLIMATE,RESIDENTIAL}} \cdot x_{i,j,t}^{residential} + \mathrm{b}_{s,t}^{\mathrm{CLIMATE,ELECTRIC}} \cdot x_{i,j,t}^{electric} \tag{6}$$

where the coefficients $\mathrm{b}_{t}^{\mathrm{CLIMATE}}$ are the monetized climate damages per unit of natural gas activity. Monetized damages per unit activity are a function of the GHG emission factors, as well as, the social cost of carbon (SCC_{t}) and social cost of methane ($\mathrm{SCCH4}_{t}$), metrics representing the present value of the anticipated future damages that would arise from an incremental unit of emissions in a given year t. We apply mean estimated values assuming a 3% discount rate as reported in US EPA publications and historically used in federal regulatory analysis[68,71]; these values vary depending on emissions year from

2005 to 2030 – \$29 to \$58 per metric ton of CO_2 and \$720 to \$1865 per metric ton of CH_4 (base case assumption). These metrics account for climate change impacts, such as changes in agricultural productivity and human health, property damage from increased flood risk, and changes in heating and cooling costs. To explore the sensitivity of decisions to the social cost of carbon, we use a range of \$1 to \$400 per metric ton. The lower end of the range reflects the Trump administration's executive order in 2017 which lowered the social cost of carbon to be used in regulatory analysis to account for "domestic versus international impacts and the consideration of appropriate discount rates" and in order "to ensure sound regulatory decision making…[and] use estimates of costs and benefits in their regulatory analyses that are based on the best available science and economics."[129] The higher end of the range reflects mean estimates from Ricke et al. (2018).[110] The coefficients $\mathrm{b}_{s,t}^{\mathrm{CLIMATE}}$, which vary by natural gas process, are of the form:

$$\mathrm{b}_{s,t}^{\mathrm{CLIMATE}} = \mathrm{SCCH4}_t \cdot \mathrm{EF}^{\mathrm{CH4}} + \mathrm{SCC}_t \cdot \mathrm{EF}^{\mathrm{CO2}} \quad (7)$$

4.2.4 Employment objectives

The employment objective function is to maximize cumulative impacts, alternatively specified in terms of job-years and wages, resulting from upstream natural gas activity. The general form of the objective function in terms of job-years is:

$$\mathrm{Max}\ \ \mathrm{Z}^{\mathrm{EMPLOYMENT}} = \textstyle\sum_{i,t}\left(\mathrm{a}^{\mathrm{EMPLOY,PREPRODUCTION}} \cdot x_{i,t}^{spud} + \mathrm{a}^{\mathrm{EMPLOY,PRODUCTION}} \cdot x_{i,t}^{producing}\right) \quad (8)$$

where the coefficients $\mathrm{a}^{\mathrm{EMPLOY,PREPRODUCTION}}$ and $\mathrm{a}^{\mathrm{EMPLOY,PRODUCTION}}$ are the job-years per spud and producing well, respectively. The metric job-year is a full- or part-time job over a single year, not a sustained job over multiple years or a career, and includes direct jobs within the natural gas sector and spillover into the non-resource economy.

The general form of the objective function in terms of wages (USD 2017) is:

$$\text{Max } Z^{\text{EMPLOYMENT}} = \text{WAGES}_{i,t} \sum_{i,t}\left(a^{\text{EMPLOY,PREPRODUCTION}} \cdot x_{i,t}^{spud} + a^{\text{EMPLOY,PRODUCTION}} \cdot x_{i,t}^{producing}\right) \quad (9)$$

where $\text{WAGES}_{i,t}$ is the wage per job-year for each county i and year t, based on historical annual average wages per county.

4.2.5 Multiobjective functions

Multiobjective programming facilitates comparisons between conflicting objectives. Here, we use two approaches for aggregating objective functions, which provide differing insight regarding development decisions.

We evaluate pairwise tradeoffs between objectives in terms of their physical units (i.e., premature mortalities, Kelvin-years, job-years) using the ε-constraint method.[130] The ε-constraint method is based on formulating an auxiliary model in which a single objective is optimized subject to a secondary objective which is reformulated as a constraint. Are a projection on the employment and air quality of a three dimensional set This constraint imposes an upper limit ε on the value of the secondary objective and is iteratively solved to generate the Pareto set. As an example of the pairwise tradeoff between air quality and employment, it can be expressed as:

$$\text{Max } Z^{\text{EMPLOYMNENT}} \quad (10)$$

$$s.t. \quad Z^{\text{AIR QUALITY}} \leq \varepsilon_k \quad (11)$$

$$with\ \varepsilon_k = \varepsilon_1, \dots, \varepsilon_n \text{ and } E_{\text{LB}} \leq \varepsilon_k \leq E_{\text{LB}} \quad (12)$$

In this example, climate change impacts are treated as a projection on the employment and air quality Pareto set. The extreme points (E_{LB}, E_{LB}) are derived based on solving the single objective problems.

We alternatively develop an aggregate monetized objective as follows:

$$\text{Max } Z^{\text{TOTAL}} = Z^{\text{EMPLOYMENT}} - Z^{\text{AIR QUALITY}} - Z^{\text{CLIMATE CHANGE}} \quad (13)$$

4.2.6 Natural gas system constraints

We derive multiple natural gas activity constraints. Here, we present a representative formulation of these constraint sets for both retrospective and future cases.

Cumulative production constraint

In the retrospective case, we require that cumulative shale gas production over the development period must be equivalent to historical production ($\mathrm{P}^{\mathrm{CUMULATIVE}}$), as given by:

$$\sum_{i,t=1} \mathrm{p}_{i,u=1} \cdot x_{i,t}^{spud} + \cdots + \sum_{i,t=\mathcal{T}} (\mathrm{p}_{i,u=1} \cdot x_{i,t=\mathcal{T}}^{spud} + \cdots + \mathrm{p}_{u=1,t} \cdot x_{i,t=1}^{spud}) + \cdots = \mathrm{P}^{\mathrm{CUMULATIVE}} \tag{14}$$

where $\mathrm{p}_{i,u}$ is the annual well productivity. Well productivity profiles are derived from historical well production[131–133], differentiating sweet and non-sweet production regions of the Marcellus and Utica plays, as delineated by the U. S. Geological Survey[134,135]; specifically, we derive profiles by regressing annual well production on well production year and whether a location is a sweet spot or not (n=53,653). We additionally perform sensitivity analyses of well productivity, modifying the functional form. We set cumulative production ($\mathrm{P}^{\mathrm{CUMULATIVE}}$) based on historical production from 2005 to 2016 (27.8 tcf).

In the future case, we assume that cumulative shale gas production over the development period must not exceed reserves (R), as given by:

$$\sum_{i,t=1} \mathrm{p}_{i,u=1} \cdot x_{i,t}^{spud} + \cdots + \sum_{i,t=\mathcal{T}} (\mathrm{p}_{i,u=1} \cdot x_{i,t=\mathcal{T}}^{spud} + \cdots + \mathrm{p}_{u=1,t} \cdot x_{i,t=1}^{spud}) + \cdots \leq \mathrm{R} \tag{15}$$

Our base assumption is the 2016 shale gas reserve estimate (99.6 tcf) for the Appalachian basin reported by the US Energy Information Administration (EIA).[34] We additionally perform a sensitivity analysis to reflect a range of reserve estimates.

Annual wells constraints

We formulate a set of constraints that assign an upper limit on the number of wells that can be spudded in a given year, which reflect drilling crew and rig resource limits.

$$\sum_i x_{i,t}^{spud} \leq \mathrm{W}_t^{\mathrm{MAX}} \ \ \forall t = 1, \dots, \mathcal{T} \tag{16}$$

where $\mathrm{W}_t^{\mathrm{MAX}}$ is the maximum number of spud wells in a given year. In the retrospective case, we assume that the industry cannot develop more rapidly than what actually occurred. In the future case, we assume that drilling cannot exceed the annual peak observed from 2005 to 2016; it is possible that this drilling resource limit increases in the future as more drilling crews and rigs come online.

Well density constraints

We formulate a set of constraints that assign an upper limit on the well density in a given county.

$$\sum_t x_{i,t}^{spud} \leq \mathrm{W}_i^{\mathrm{MAX}} \ \ \forall i = 1, \dots, \mathcal{I} \tag{17}$$

where $\mathrm{W}_i^{\mathrm{MAX}}$ is the maximum number of spud wells in a given county. We estimate the maximum number of wells that can be drilled in each county based on the land area of each county and the maximum well density observed for any county over the period 2005 to 2016 (1.75 producing wells per square mile). The maximum well density may differ from this and vary by county based on the extent of the resource and other land uses.

Cumulative number of spud wells equivalent to producing wells constraints

We assume that the cumulative number of spud wells is equivalent to the number of producing wells. This formulation implicitly assumes that after a well is spud it continues to produce until productivity declines to zero.

$$x_{i,t=1}^{producing} - \sum_i \sum_{t=1} x_{i,t}^{spud} = 0 \tag{18}$$

.

.

.

$$x_{i,t=\mathcal{T}}^{producing} - \sum_i \sum_{t=1}^{\mathcal{T}} x_{i,t}^{spud} = 0 \quad (19)$$

Residential end use consumption constraints

In the retrospective and future cases, we assume that residential natural gas consumption must be equivalent to historical or projected future demand ($\mathrm{C}_{i,t}^{\mathrm{RESIDENTIAL}}$), respectively. We further assume that demand is spatially and temporally fixed, and it is largely unaffected by regional production, consistent with the historical residential consumption demand observations.

$$x_{i,t}^{residential} \leq \mathrm{C}_{i,t}^{\mathrm{RESIDENTIAL}} \;\; \forall i, t \quad (20)$$

$$\sum_{i,t} x_{i,t}^{residential} = \sum_{i,t} \mathrm{C}_{i,t}^{\mathrm{RESIDENTIAL}} \quad (21)$$

We formulate the constraint set as such to allow for readily incorporating displacement of natural gas with alternative energy technologies. To derive historical consumption by county and year, we use annual, state-level residential natural gas consumption reported by the EIA; state consumption is allocated to each county using the ratio of the number of housing units burning natural gas in a given county to the number of housing units burning natural gas in the state, as reported by the U.S. Census Bureau. To develop future residential consumption, we estimate annual residential consumption growth rates from 2017 to 2030 based on EIA regional consumption projections (under the EIA reference case). Then, we apply these growth rates to historical consumption by county in 2016.

Commercial end use consumption constraints

In the retrospective and future cases, we assume that commercial natural gas consumption must be equivalent to historical or projected future demand ($\mathrm{C}_{i,t}^{\mathrm{COMMERCIAL}}$), respectively. We further assume that demand is spatially and temporally fixed, and it is largely unaffected by regional production, consistent with the historical commercial consumption demand observations.

$$x_{i,t}^{commercial} \leq \mathrm{C}_{i,t}^{\mathrm{COMMERCIAL}} \;\; \forall \mathrm{i}, \mathrm{t} \quad (22)$$

$$\sum_{i,t} x_{i,t}^{commercial} = \sum_{i,t} \mathrm{C}_{i,t}^{\mathrm{COMMERCIAL}} \tag{23}$$

We formulate the constraint set as such to allow for readily incorporating displacement of natural gas with alternative energy technologies. To derive historical consumption by county and year, we use an approach similar to that used in development of the NEI. We use annual, state-level commercial natural gas consumption reported by the EIA; state consumption was allocated to each county using the ratio of employment in the commercial sector by county to commercial sector employment in the state, as reported by the U.S. Census Bureau. To develop future commercial consumption scenarios, we estimate annual commercial consumption growth rates from 2017 to 2030 based on EIA regional consumption projections (under the EIA reference case). Then, we apply these growth rates to historical consumption by county in 2016.

Industrial end use consumption constraints

In the retrospective case, we assume that industrial natural gas consumption must be equivalent to historical demand ($C_{\mathrm{i,t}}^{\mathrm{INDUSTRIAL}}$) and it is spatially and temporally fixed.

$$x_{i,t}^{industrial} \leq \mathrm{C}_{\mathrm{i,t}}^{\mathrm{INDUSTRIAL}} \;\; \forall \mathrm{i,t} \tag{24}$$

$$\sum_{i,t} x_{i,t}^{industrial} = \sum_{i,t} \mathrm{C}_{i,t}^{\mathrm{INDUSTRIAL}} \tag{25}$$

We formulate the constraint set as such to allow for readily incorporating integration of alternative energy technologies. We modify this constraint set for the future case, allowing for additional industrial load beyond 2016 levels to be sited in any county, as follows:

$$x_{i,t}^{industrial} \geq \mathrm{C}_{i,t=2016}^{\mathrm{INDUSTRIAL}} \; \forall i,t \tag{26}$$

$$\sum_{i} x_{i,t}^{industrial} = \sum_{i} \mathrm{C}_{i,t}^{\mathrm{INDUSTRIAL}} \;\; \forall \mathrm{t} \tag{27}$$

To derive historical consumption by county and year, we use an approach similar to that used in development of the NEI. We use annual, state-level industrial natural gas consumption reported by the

EIA; state consumption is allocated to each county using the ratio of employment in the industrial sector by county to industrial sector employment in the state, as reported by the U.S. Census Bureau. To develop future industrial consumption, we estimate annual industrial consumption growth rates from 2017 to 2030 based on EIA regional consumption projections (under the EIA reference case). Then, we apply these growth rates to historical consumption by county in 2016.

Electric power sector natural gas consumption constraints

In the retrospective and future cases, we assume that electric power sector natural gas consumption must be equivalent to historical or projected future demand ($\mathrm{C}_{i,t}^{\mathrm{ELECTRIC}}$). We allow for additional electric power load beyond 2004 or 2016 levels for the retrospective and future cases, respectively, to be sited in any county, as follows:

$$x_{i,t}^{electric} \geq \mathrm{C}_{i,t=2005\,\|\,2016}^{\mathrm{ELECTRIC}} \;\forall i,t \tag{28}$$

$$\textstyle\sum_i x_{i,t}^{electric} = \sum_i \mathrm{C}_{i,t}^{\mathrm{ELECTRIC}} \;\; \forall t \tag{29}$$

To derive historical consumption by county and year, we use plant-level natural gas consumption volumes from CEMS. To develop future electric power sector natural gas consumption, we estimate annual electric power sector consumption growth rates from 2017 to 2030 based on EIA regional consumption projections (under the EIA reference case). Then, we apply these growth rates to historical consumption by county in 2016.

4.3 Results and discussion

4.3.1 Retrospective tradeoffs between socioeconomic and environmental objectives

To demonstrate the mechanics of the optimization, we begin with a discussion of the retrospective model of natural gas activity from 2005 to 2016. In our presentation of results, we depict an approximation of the Pareto frontier based on the multiobjective formulation, as well as, the monetized optimum. We additionally denote the corner solutions, whereby each single objective is optimized, and define illustrative compromise solutions along the Pareto frontiers.

The air-climate Pareto frontier in Figure 10a demonstrates that there is a nominal tradeoff. We find that siting and planning policies with respect to upstream infrastructure and electricity generation capacity expansion, and assuming constant historical production and consumption volumes, may have decreased air quality impacts by greater than 15%, but such policies would have trivial effects on climate impacts (Figure 10b). We additionally observe that when optimizing for monetized air quality and climate change impacts under base case VSL and SCC assumptions, the solution is similar to the multiobjective corner solution where premature mortalities are minimized. The monetized formulation incorporating all three impact areas iterates towards the employment optimum and is suboptimal relative to the air-climate multiobjective scenario.

As shown in Figure 10c, there is an almost linear relationship between climate change and employment impacts absent additional climate mitigation policies. Actual impacts are near Pareto optimal, when considering only climate and employment objectives. In the monetized formulation, in which we consider either climate and employment only or all three impacts, the optimal solution is the same as the employment optimum.

The air-employment Pareto frontier in Figure 10e depicts a near linear tradeoff. There is an historical implied tradeoff of 84 job-years per premature mortality at a systems level. Infrastructure siting and planning policies improve this tradeoff to 100 job-years per premature mortality; however, there is spatial heterogeneity with respect to this tradeoff, and policies may have the undesirable effect of making some communities worse off.

Figure 11 shows the spatial distribution of impacts and natural gas activity for a compromise solution along the air-employment Pareto frontier, as well as, the actual distribution. For the compromise solution, upstream activity and additional electric generation capacity shift towards lower population areas to minimize air quality impacts. Although cumulative premature mortality is notably less for the compromise solution relative to what actually occurred, there is an almost imperceptible difference in the spatial distribution; this is largely because we assume that industrial, commercial, and residential end use demands are spatially and temporally fixed. We also observe that upstream activity shifts towards less productive portions of the Marcellus and Utica plays. This in part occurs because drilling and producing wells require the same marginal labor factors, regardless of well productivity. Thus, when maximizing employment, subject to fixed cumulative production, there is a tendency towards an outcome that is inefficient from the perspective of a private gas firm which ostensibly would seek to drill in the most productive regions and economize on labor. We further find that upstream activity and electric generation capacity expansion are spatially more concentrated than the actual distributions, which may have equity implications. Figure 12 traces impacts and natural gas activity over time. The shape of temporal profiles when optimizing for employment, air quality, climate change, and monetized impacts are similar.

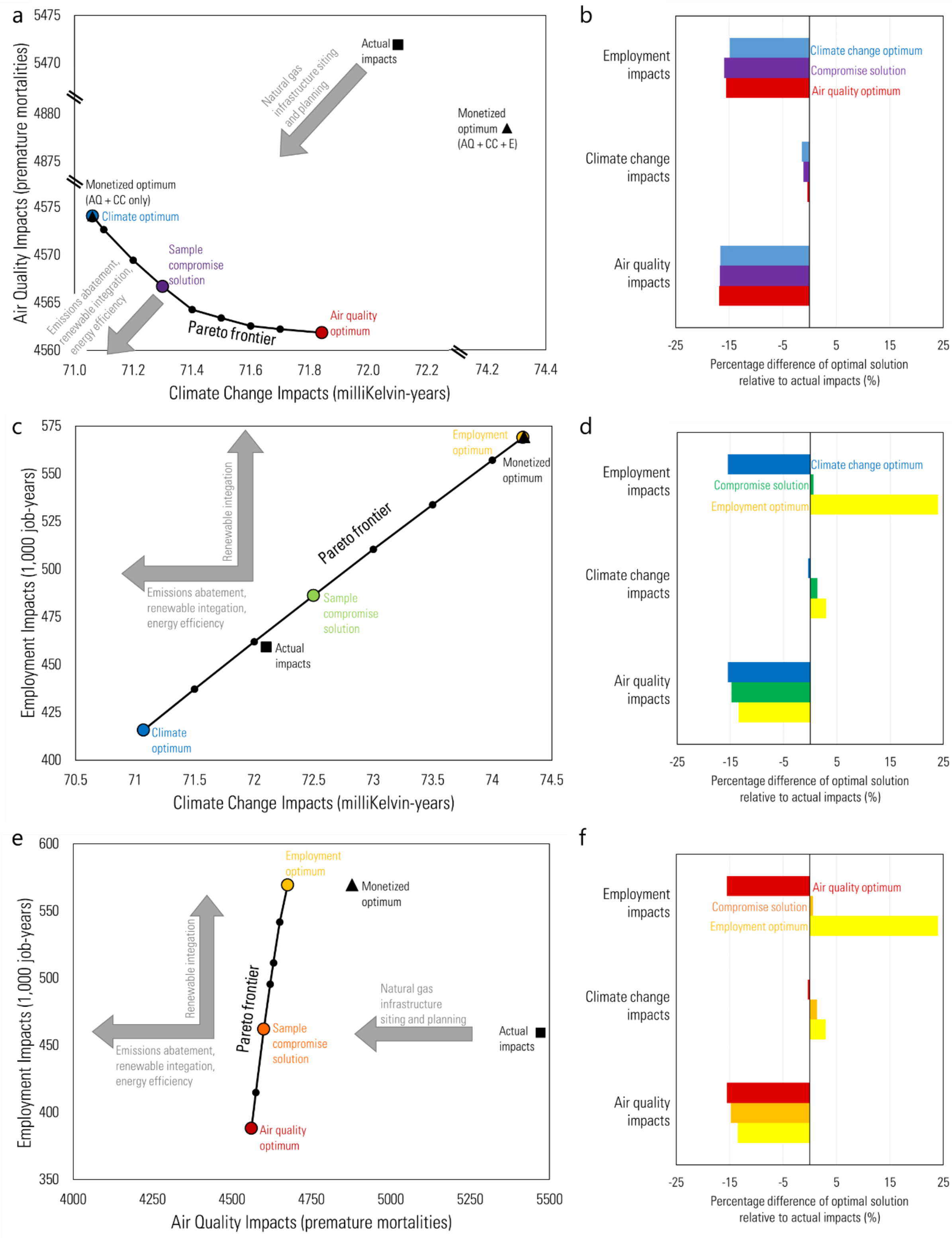

Figure 10. Tradeoffs between cumulative impacts for retrospective case.

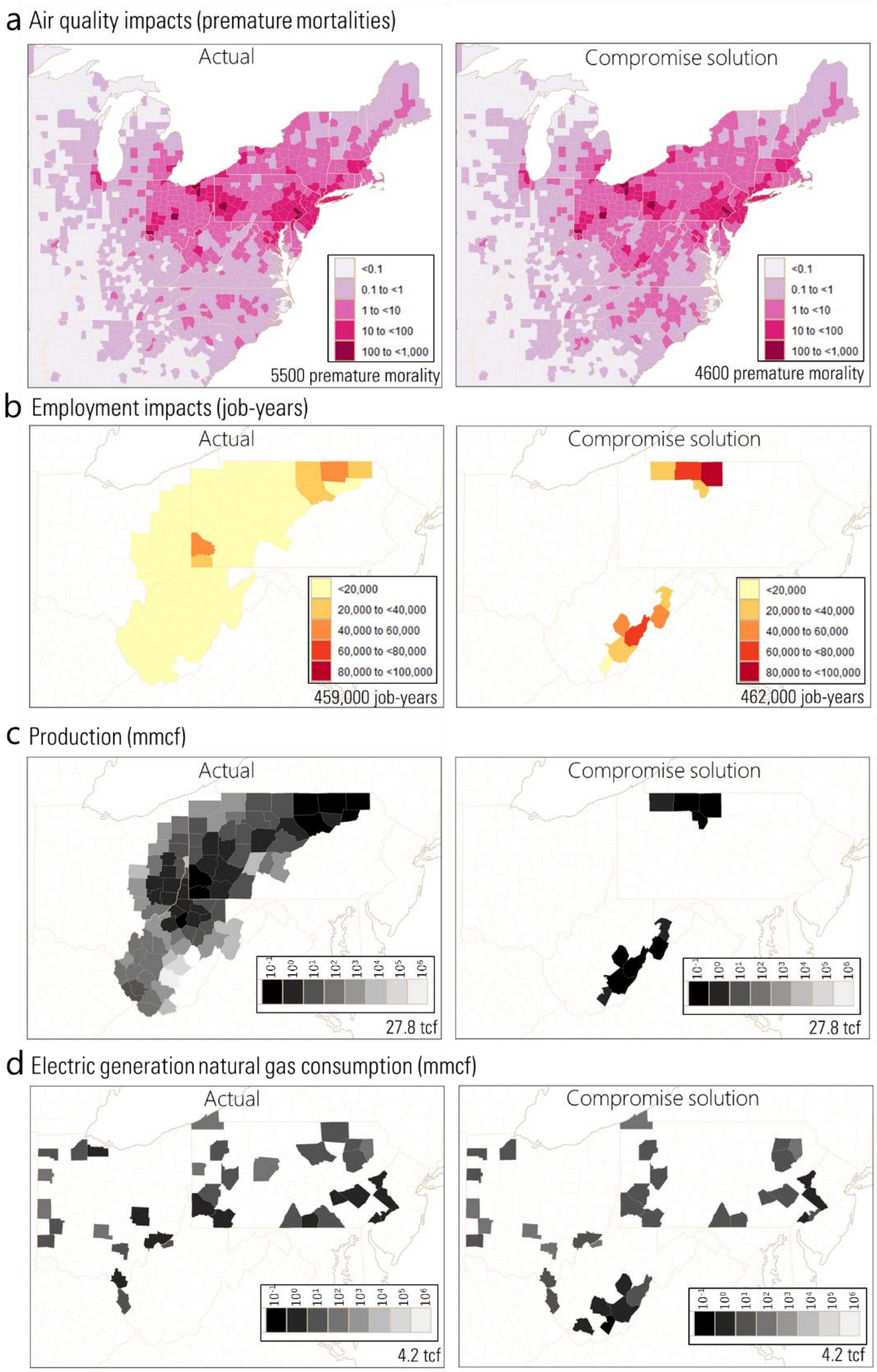

Figure 11. Spatial distribution of cumulative impacts, production, and electricity generation. Comparison of retrospective air quality-employment compromise solution and based on actual natural gas activity.

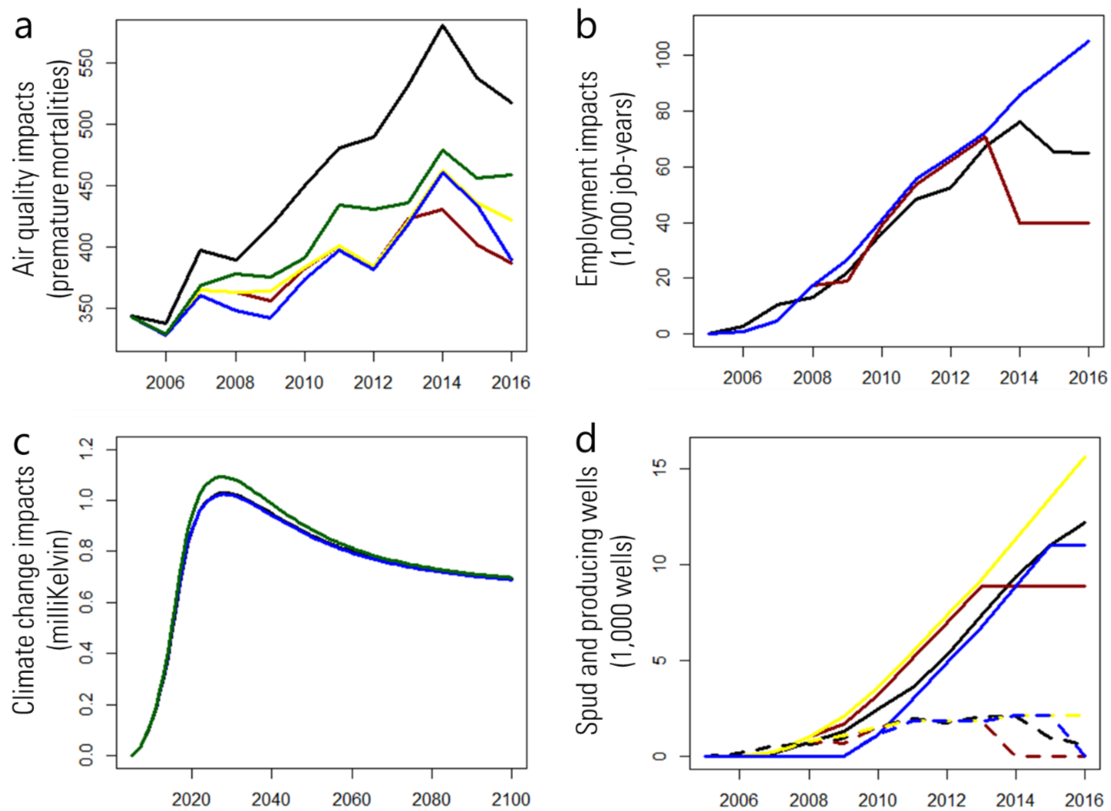

Figure 12. Impacts and natural gas activity over time for retrospective case. Red, blue, and yellow lines represent optimal solutions for pairwise multiobjective formulations in which all weight is on the air quality, climate change, or employment objective, respectively. Green lines represent the monetized optimum, considering air quality, climate change, and employment impacts. Black lines represent actual impacts or activity. In (d), the solid and dotted lines represent producing and spud wells, respectively.

4.3.2 Future natural gas system planning

In this section, we focus on future development pathways in the Appalachian basin from 2017 to 2030. In addition to siting and planning policies, we consider emissions abatement and renewable integration.

We find that air quality and climate change objectives are not in conflict, as shown in Figure 13. Under both objectives, the model selects to cease drilling activities, although there is residual production from wells drilled prior to 2017 (Figure 13d). While optimizing for these environmental impacts aligns with a fossil fuel divestment strategy or moratorium on new production, natural gas is still a fundamental part of the regional energy system. As such, consumption does not cease, and meeting continued demand would necessitate the region becoming a net importer of natural gas, as was the situation prior to the shale gas boom. Overall, there are still nontrivial climate change and regional air quality impacts, even at the optimum (Figure 13a, c), and much of the impact of upstream natural gas impacts are borne in other source fuel regions not included within the model.

Figure 14a depicts the pairwise tradeoff between conflicting climate change and employment objectives. We find that the employment objective pulls the model to produce more, but not to deplete reserves over the development period. We also observe that GHG emissions abatement decreases climate impacts. Abatement has the potential to increase the marginal number of jobs, an effect which we do not explicitly model; notwithstanding, abatement costs that make it inefficient for private firms to market and use natural gas may result in a loss of jobs within the industry and a labor shift into other sectors. The marginal climate impacts are vastly greater than marginal employment effects from natural gas activity; thus, we find that the monetized and climate optimums are equivalent under base SCC assumptions. Considering the monetized objective, the SCC would have to be less than $9 per metric ton to shift the decision away from the climate optimum.

Figure 14b depicts the pairwise tradeoff between conflicting air quality and employment objectives. Air pollutant emissions abatement decreases premature mortalities, while not impacting cumulative employment; as with GHG abatement, the marginal number of jobs may increase with abatement. The marginal air quality impacts are greater than marginal employment effects from natural gas activity; thus, we find that the monetized and air quality optimums are equivalent under base VSL assumptions. With respect to the monetized objective, the VSL would have to be less than $2.6M to shift the decision away from the air quality optimum.

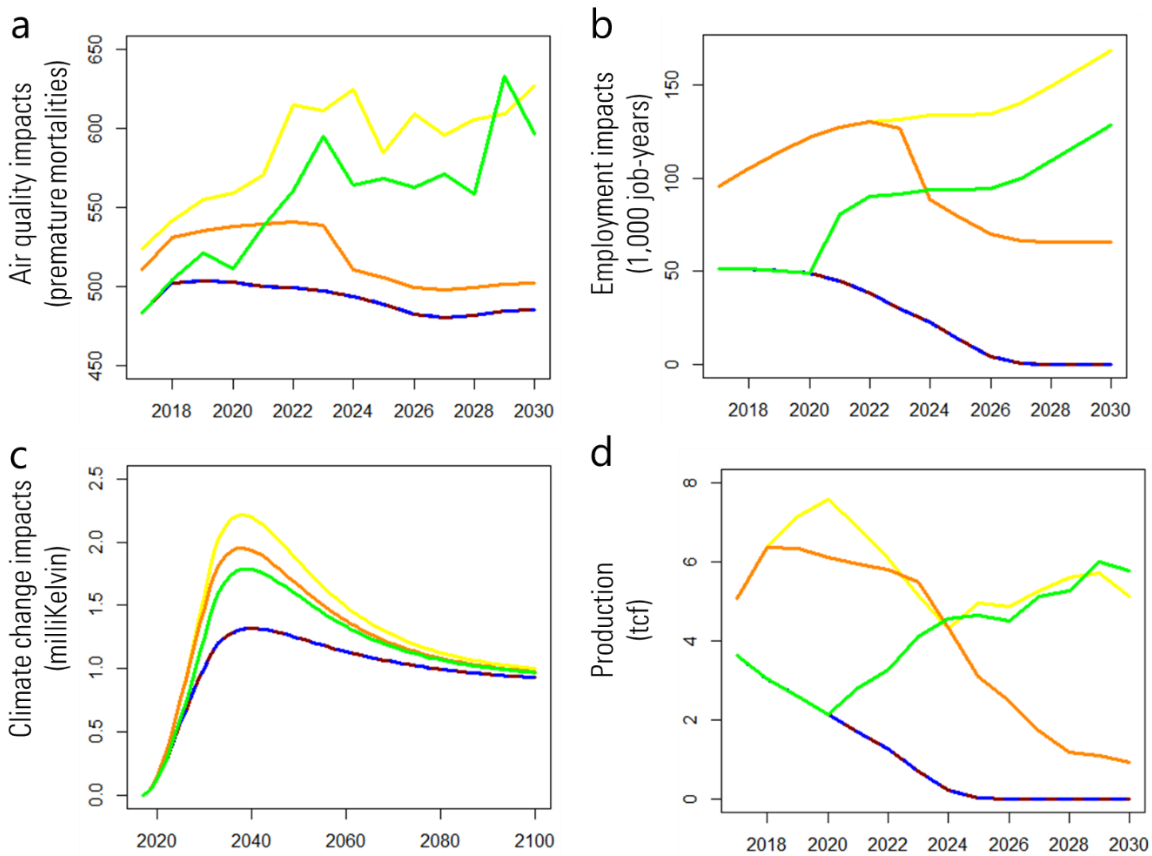

Figure 13. Impacts and natural gas activity over time for future case. Red, blue, and yellow lines represent optimal solutions for pairwise multiobjective formulations in which all weight is on the air quality, climate change, or employment objective, respectively. Green and orange lines represent the air quality-employment and climate change-employment compromise solutions, respectively.

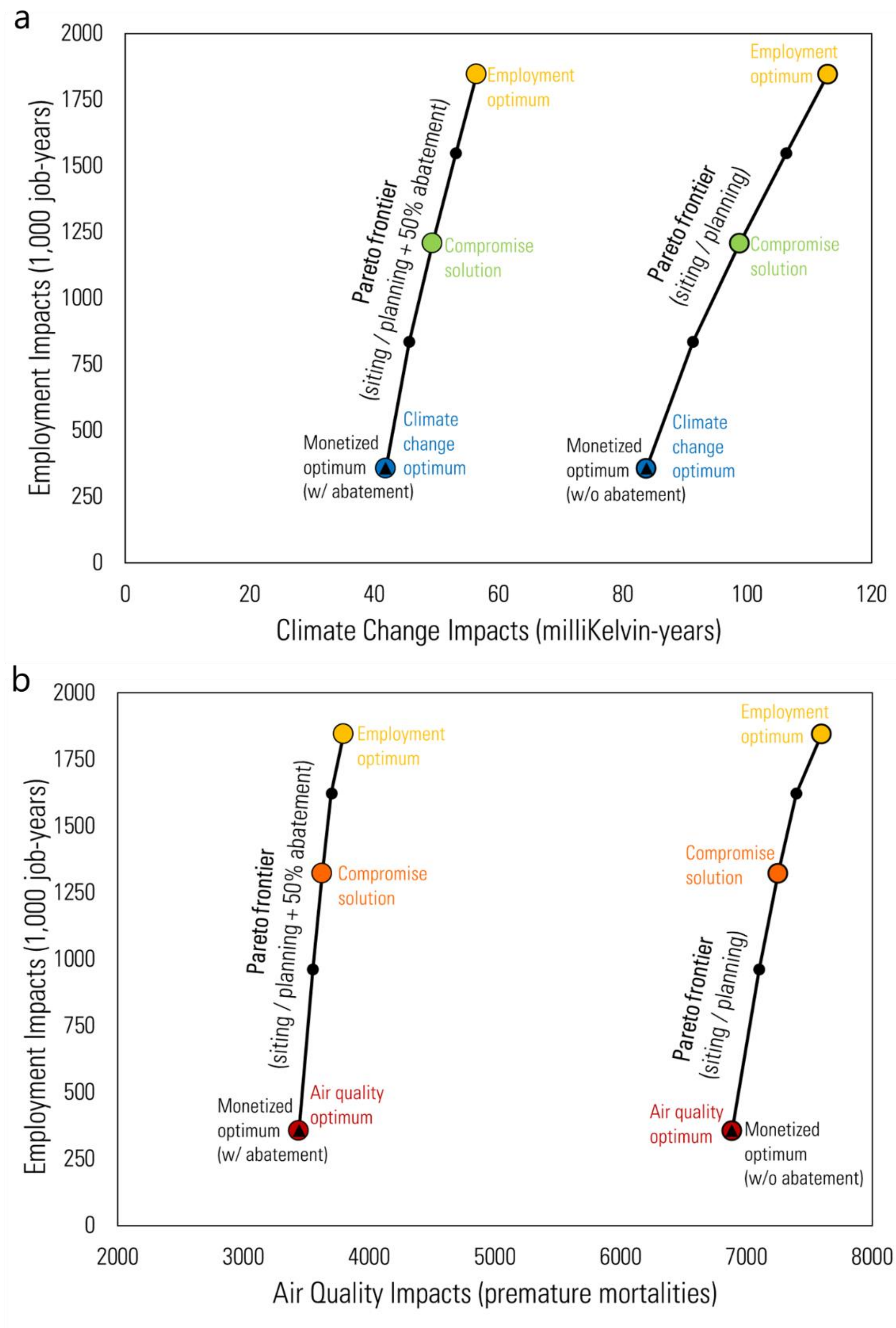

Figure 14. Tradeoffs between cumulative air quality, climate change, and employment impacts associated with natural gas activity from 2017 to 2030.

4.4 Timing and value of greenhouse gas emissions abatement

Implementing abatement has the potential to improve the extensive margin of climate impacts, which has the effect of reducing conflicts between environmental and socioeconomic objectives. In the previous section, we assumed a flat greenhouse gas emissions reduction of 50% beginning in 2017. Here, we evaluate the temperature impacts and monetized value of alternative CO2 and CH4 abatement strategies and delayed abatement.

We model two different development regimes – minimizing climate impacts and maximizing employment – which together provide a partial spread of development pathways. Assuming abatement begins in 2017, we find that policies in which methane is reduced by 50% have a larger cumulative temperature benefit by mid-century than policies in which carbon dioxide is reduced by 50%, as shown in Figure 15a-b. By the end of the century, carbon dioxide abatement has a larger cumulative temperature benefit than methane abatement. While methane abatement strategies have a relatively minimal effect on the absolute warming in 2100, they may contribute to minimizing interim losses and could have drastic benefits in the case of climate tipping points.

The value of abatement – expressed as the net monetized climate damages at the optimum without and with abatement – ranges from $0 to $58B, as depicted in Figure 16. While the value of abatement would seemingly be higher in an employment maximizing regime where there is more natural gas activity than in a climate impact minimizing regime, we find that there is almost no difference in the value of carbon dioxide abatement between regimes.

We also consider the effect of delaying abatement until 2025, as shown in Figure 15c-d. Delaying abatement results in cumulative temperature benefits that are less than half that derived from abatement beginning in 2017. We estimate that the cost of delay, or the forgone benefits of delay, (Figure 16) can be upwards of $28B.

In addition to or in lieu of the constraint sets derived, which provide a broad systems-level representation of natural gas production and consumption, it is possible to employ capacity expansion or dispatch constraint sets that may offer additional insight at higher temporal resolution and limits at the extensive margin of natural gas activity.

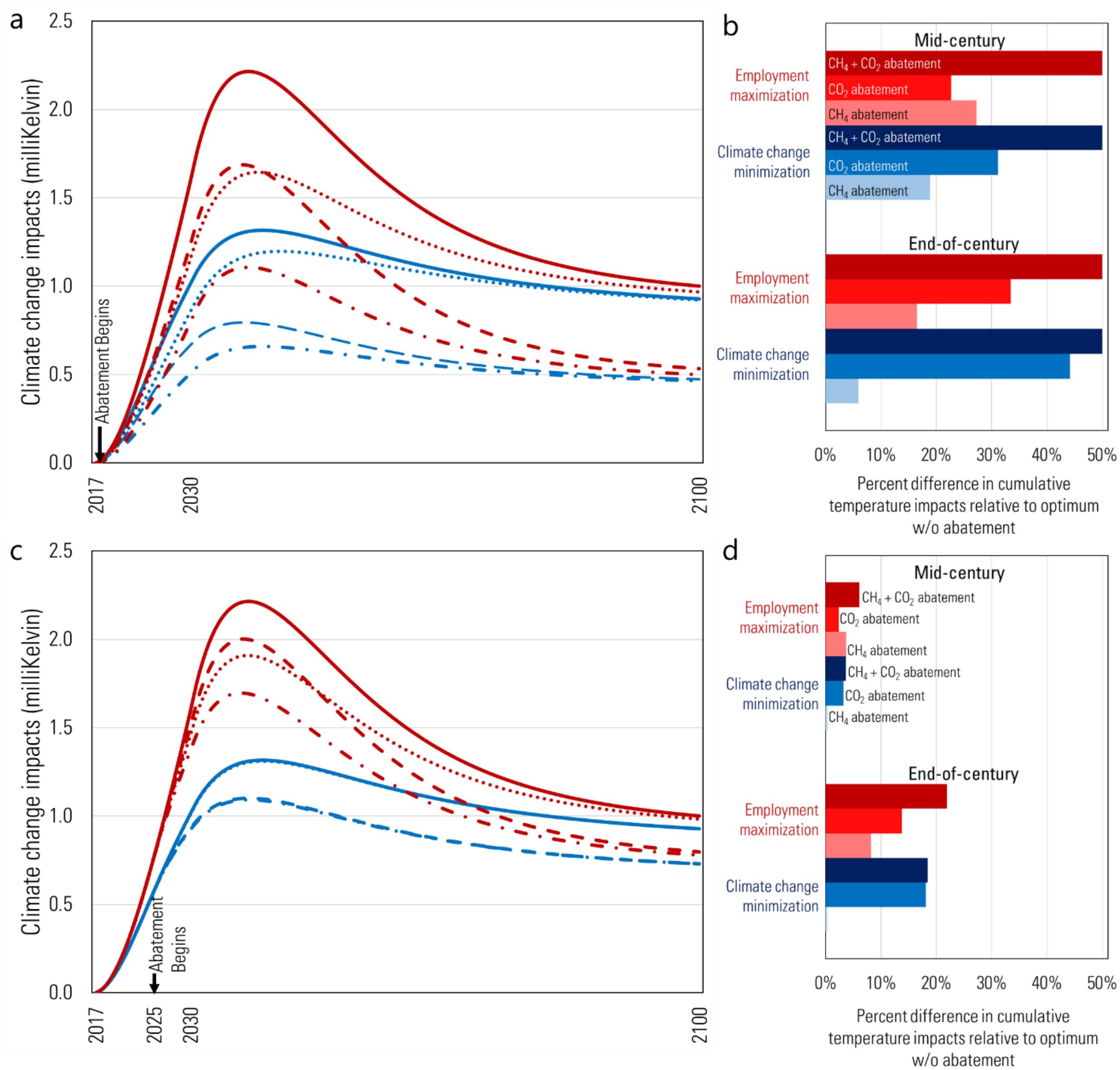

Figure 15. Effect of climate change abatement on global temperature. (a) and (c) depict temperature impacts to 2100. Blue and red lines represent temperature impact from minimizing climate impacts and maximizing employment impacts, respectively. Solid lines represent optimal solutions without abatement. Dotted, dashed, and dotted-dashed lines represent 50% reduction in CH4, CO2, or both, respectively. Abatement either begins in 2017 (a-b) or is delayed until 2025 (c-

d). (b) and (d) depict the percent difference between cumulative temperature impacts of abatement scenarios relative to scenarios without abatement. Mid- and end-of-century cumulative temperature impacts are provided.

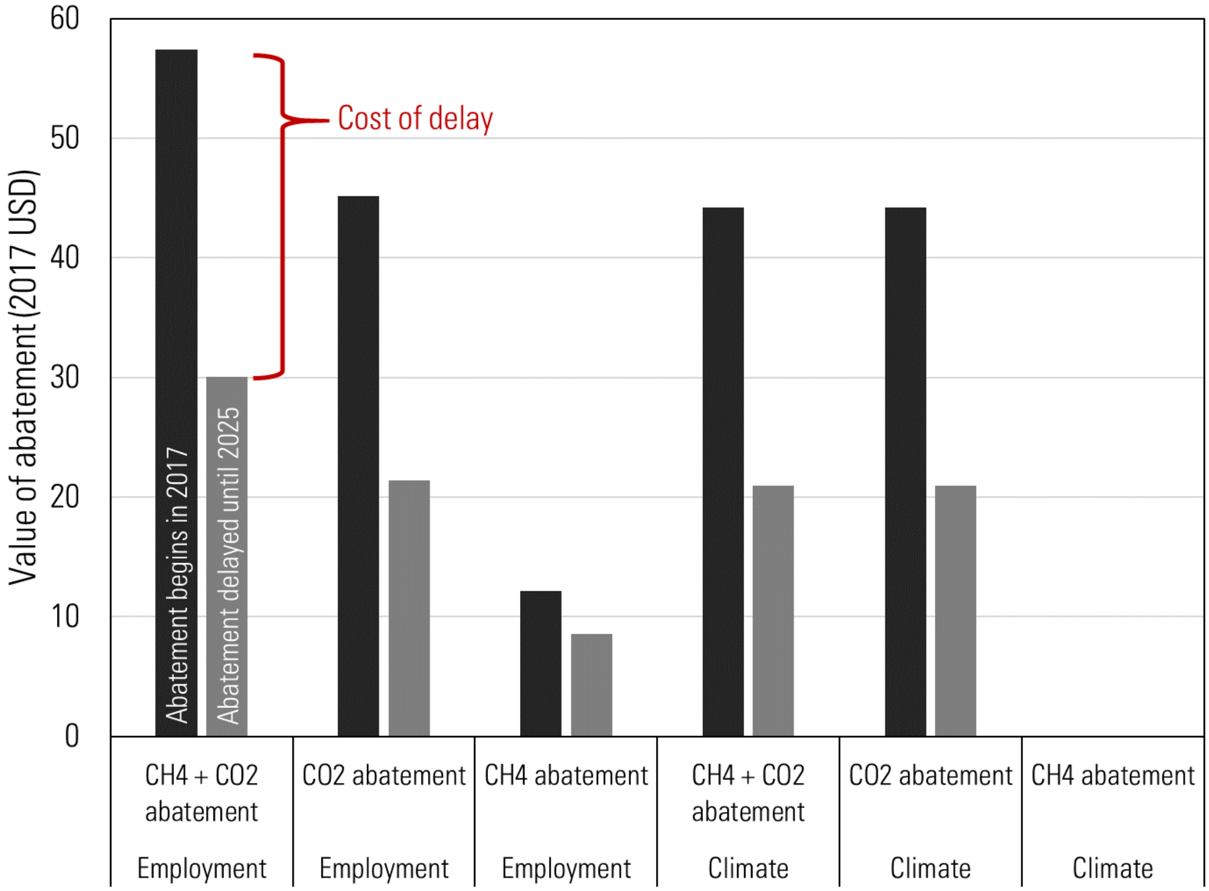

Figure 16. Value of climate change abatement and the cost of delaying abatement. The value of abatement is estimated under employment maximizing and climate change impact minimizing regimes. The value of abatement is the net monetized climate damages at the optimum without and with abatement.

4.5 Conclusions

Natural gas system development is driven by the complexity inherent in physical systems and the influence of a myriad of diverse, interacting stakeholders with heterogeneous preferences. The purpose of this study is to distill a portion of this complexity and provide descriptive insight into how the energy system theoretically could develop if other objectives that are often the subject of public discourse and concern, such as jobs, climate change, and health effects, influence the decision-making process. While environmental and employment objectives are conflicting if we follow a natural gas pathway consistent with the status quo, the collection of siting, emissions abatement, and renewable integration policies can resolve and reverse these conflicts. There are pathway dependencies between these policies, and delaying implementation only amplifies the cumulative, negative tradeoffs between employment and environmental objectives. Here, we demonstrate that future energy systems have the potential to meet multiple objectives.

5 Air, Climate, and Labor Market Equity of Natural Gas Systems

In this study, we take a systematic, but exploratory, approach to quantify and characterize the multi-dimensional equity of natural gas systems, considering the entire supply chain from production to end use and over the boom-and-bust cycle. We assess equity of natural gas development in the Appalachian basin at a systems level and focus on spatial, temporal, and distributional equity as it relates to air quality, climate change, and labor market impacts. We find that there are high temporal and spatial inequities with respect to cumulative air and employment impacts. With respect to distributional equity of air quality impacts, we do not observe a disparity in mortality rates across subpopulations on the basis of income and poverty; however, there is a trend of increasing income corresponding to decreasing damages, which demonstrates the higher health burden on lower income communities. With respect to distributional equity of labor markets, we find statistically significant declines in the income disparity and poverty rates in producing counties. Pairwise comparisons of air, climate, and employment impacts reveal that they are highly correlated.

5.1 Introduction

The U.S. energy landscape is rapidly evolving, and changes in the energy system over the past decade are largely associated with technological advancements enabling growth in domestic natural gas activity. Correspondingly, there is an expanding body of evidence regarding the impacts of natural gas activity across the supply chain on water quality[74–77], air quality[78–80], ecosystems[81,82], climate[83–85], labor markets[86–88], public health[89,90], and several other environmental and socioeconomic factors. However, the spatial, temporal, and distributional effects and equity of natural gas systems – i.e., how benefits and costs *are* and *should be* distributed spatially, temporally, or among sub-populations on the basis of demographics – are still largely unexplored.[29,42] With respect to policy research, there have been few studies which solely investigate distributional effects of upstream natural gas infrastructure by race[136], poverty[88], or income[117]. As part of the regulatory process, often in relation to siting energy infrastructure, environmental justice analyses are performed to identify and assist in addressing disparate impacts on vulnerable populations at an early stage in the public and private decision-making process; however, existing analytical instruments and processes may be inadequate for informing and facilitating public and private decision making by regulators and developers, as evidenced by such events as those surrounding the Dakota Access Pipeline project.[137] A systems-level approach that considers embodied energy injustices – the full spectrum of transboundary socio-environmental injustices across the supply chain – may better facilitate decision making.[138]

Here, we take a systematic, but exploratory, approach to quantify and characterize the multi-dimensional equity of natural gas systems, considering the entire supply chain from production to end use and over the boom-and-bust cycle. We assess equity at a systems level and focus on spatial, temporal, and distributional equity as it relates to air quality, climate change, and labor market impacts. To evaluate equity, it is necessary to decide and operationalize which distributive rules apply and which metrics to use.[139] We apply and interpret variants of standard methods and equity metrics, such as Gini coefficients and unit-hazard coincidence. And to further measure and highlight tradeoffs and the distribution of impacts, we develop new methods and metrics such as: job-years created per life-year lost (or premature mortality), premature mortalities by race, income, and poverty level, and cross-impact elasticities. We apply this approach to the shale gas boom (and decline) in the Appalachian basin from 2004 to 2016, leveraging previously derived temporally- and spatially-resolved estimates of impacts.[124] These impacts include: premature mortality from primary fine particulate matter ($PM_{2.5}$) and secondary $PM_{2.5}$ formed from the atmospheric oxidation of nitrogen oxides (NO_X) and volatile organic compounds (VOCs) emissions, global mean temperature change from carbon dioxide (CO_2) and methane (CH_4) emissions, and employment effects associated with natural gas development. In the following sections, we step through each dimension of equity, for which we estimate and interpret quantitative measures, describe the limitations, and evaluate the utility of such measures for policy design and decision making.

5.2 Methods

Equity has been variably and sometimes divergently defined, interpreted, and operationalized in legislation, judicial decisions, and public policy, as well as in scholarly research in natural sciences, economics, behavioral science, and operations research. For analytical purposes, we define equity as equality in the distribution of impacts across populations, a broad definition which provides the latitude for incorporating the nuances of scale, cross-disciplinary variation, and the spatial and temporal nature of each impact area.

To assess equity, we employ previously derived spatially- and temporally-resolved estimates of impacts both in physical and monetary units.[124] When selecting the appropriate scales of analysis, it is important to consider the nature of the impact, the area of a community or neighborhood based on cultural and social divisions, the scale at which data is aggregated and available, the scales at which statistical significance is observable, and the jurisdiction under which agencies or political agents have authority. In this analysis, we assess spatial and distributional equity at the county level, the resolution of quantitative air quality and employment impact estimates, and we define various temporal scales, namely in relation to climate impacts. In the following sections, we provide an overview of each equity analysis and underlying data.

5.2.1 Standard equity metrics

We estimate standard equity measures (e.g., Gini coefficient, Atkinson index, Thiel index, maximum difference, cumulative density functions, etc.), which provide information regarding the spatial and temporal variation of impacts across the system. These metrics of equity, including descriptions, mathematical properties, and axiomatics properties of provided in Table C1 of the Supplemental Information.

5.2.2 Premature mortality and mortality rates by race, income, and poverty level

We derive mortality and mortality rates associated with natural gas activity from 2004 to 2016 across the supply chain that are stratified by race, income level, and poverty level. We first modify receptor-resolved estimates of marginal mortality per metric ton of emissions, as specified in the reduced complexity air quality model Air Pollution Emission Experiments and Policy Analysis Model (Version 3) (AP3)[91,92], using mortality rates and population data stratified by age, race, income level, and/or poverty level.

The general form of the change in the number of premature mortality attributed to a change in concentration ($\Delta M_{c,t}$) is given by:

$$\Delta M_{c,t} = \alpha \cdot \sum_a \left(R_{a,c,t} \cdot P_{a,c,t}\right) \tag{30}$$

where $R_{a,c,t}$ and $P_{a,c,t}$ are the baseline mortality rate and population, respectively, for each age class a, county c, and year t. α is equivalent to $e^{\beta C_p} - e^{\beta C_b}$, where C_p and C_b are the perturbation and baseline emissions, respectively. Modifications of this formulation for each subpopulation are limited by the availability of annual, county-level mortality rate and population data stratified by age and subpopulation, as outlined in Table C5. We also systematically impute missing data.

We estimate the marginal mortality by race as follows:

$$\Delta M_{c,t,r} = \alpha \cdot R_{c,t,r} \cdot P_{c,t,r} \tag{31}$$

where $R_{c,t,r}$ and $P_{c,t,r}$ are the baseline mortality rate and population, respectively, for each race r, county c, and year t. We use county-level mortality and population estimates by race (i.e., white, black or African American, Asian or Pacific Islander, Native American or Native Alaskan) as reported by the Centers for Disease Control and Prevention (CDC).[140]

We estimate the marginal mortality by income as follows:

$$\Delta M_{c,t,i} = \alpha \cdot \frac{P_{c,t,i}}{P_{c,t}} \cdot \sum_a \left(R_{a,c,t} \cdot P_{a,c,t}\right) \quad (32)$$

where $P_{c,t}$ is the population by county c and year t, and $P_{c,t,i}$ is the population by income i, county c, and year t. We use annual, county-level population estimates disaggregated by income level (i.e., <\$15,000, >\$15,000 to <\$35,000, >\$35,000 to <\$75,000, >\$75,000 to <\$150,000, >\$150,000) reported by the US Census Bureau.[127]

We estimate the marginal mortality by poverty levels as follows:

$$\Delta M_{c,t,p} = \alpha \cdot \sum_a \left(R_{a,c,t} \cdot P_{a,c,t,p}\right) \quad (33)$$

where $P_{a,c,t,p}$ is the population by poverty level p, age class a, county c, and year t. We use annual, county-level population estimates by age and poverty levels reported by the US Census Bureau.[127]

Combining the previously described receptor-resolved marginal mortality estimates with the emissions model described in Mayfield et al. (forthcoming), we estimate total mortality by subpopulation associated with natural gas activity. In addition, we estimate mortality rates by subpopulation as follows:

$$Q_{t,r} = \frac{\sum_c M_{c,t,r}}{\sum_c P_{c,t,r}} \quad \forall t, r \quad (34)$$

$$Q_{t,i} = \frac{\sum_c M_{c,t,i}}{\sum_c P_{c,t,i}} \quad \forall t, i \quad (35)$$

$$Q_{t,p} = \frac{\sum_c M_{c,t,p}}{\sum_c P_{c,t,p}} \quad \forall t, p \quad (36)$$

where $M_{c,t,r}$, $M_{c,t,i}$, and $M_{c,t,p}$ are total mortality estimates associated with natural gas activity by race, income, or poverty level, respectively. $Q_{t,r}$, $Q_{t,i}$, and $Q_{t,p}$ are mortality rates associated with natural gas activity by race, income, or poverty level, respectively.

5.2.3 Spatial coincidence of upstream infrastructure with race, income, and poverty level

We assess the spatial coincidence of upstream infrastructure with racial and socioeconomic variables, applying standard methods used in the environmental justice literature to assess demographic disparities in the distribution of environmental hazards. Specifically, we compare producing counties and nonproducing counties using the unit-hazard coincidence method, as described in Mohai et al. (2007).[141] This method is limited, in part because it does not adequately control for proximity between environmental hazards and nearby populations.

We compile a dataset consisting of demographic and natural gas activity variables for all counties in Pennsylvania, Ohio, West Virginia, and New York over the period 2010 to 2016. We use demographic data (i.e., percent black, percent nonwhite, percent below poverty level, median household income, per capital personal income, population) reported by the U.S. Census Bureau and the U.S. Bureau of Economic Analysis, and natural gas activity data reported by state agencies.[131–133,142]

We first compute the mean and population-weighted mean values for each demographic variable for producing and non-producing counties. Then, we estimate whether there are statistically significant differences between these two populations using the Welch two-sample t-test.

To assess the relative importance of demographic characteristics in accounting for disparities, we also use binary logit regression, where the dependent variable is 1 if a county produced in a given year and 0 if not. The independent variables are a subset of the demographic variables used in the previous analysis (i.e., percent nonwhite, percent below the poverty level, median household income) to reduce multicollinearity. A statistically significant coefficient on demographic variables indicates a measurable difference between producing and nonproducing counites.

5.2.4 Income inequality and poverty level regression

We explore the effect of the natural gas boom on income inequality and poverty rates within local labor markets using an approach similar to that employed in Marchand (2013).[117] Changes in the measures of income inequality and poverty rates, before and after the boom and between treatment and comparison areas, are used to identify the distributional impacts of the natural gas boom through the local labor market variation using the following model specifications:

$$\Delta Outcome_c = \beta \cdot Treatment_c + \varepsilon_c \tag{37}$$

$$\Delta \ln(Outcome_c) = \beta \cdot Treatment_c + \varepsilon_c \tag{38}$$

where the change in the labor market between the pre-boom year of 2005 and the post-boom year of 2015 are given by $\Delta Outcome_c = Outcome_{c,t} - Outcome_{c,t-1}$ and $\Delta \ln(Outcome_c) = \ln\big(Outcome_{c,t}\big) - \ln\big(Outcome_{c,t-1}\big)$, and $Treatment_c$ is a binary variable indicating whether a county c is in the treatment or comparison group.

We compile a cross-sectional dataset comprised of poverty levels, Gini coefficients, and shale gas production of counties in 2005 and 2015 for Pennsylvania, Ohio, West Virginia, and New York. We use poverty levels reported by the U.S. Census Bureau in Small Area Income & Poverty Estimate datasets; poverty levels are the predicted percent of people of all ages in poverty based on the American Community Survey.[142] We develop county-level production estimates based on operator-reported production data for each shale well from state agencies.[131–133]

We specify two treatment sets – full and top treatment – in efforts to discern the effects associated with different intensities in natural gas activity. The full treatment set is comprised of all 90 counties with annual production exceeding 0 thousand cubic feet (mcf) in 2015. The top treatment set is comprised of 23 counties with annual production exceeding 44,000 mcf in 2015, representing the top 25% of producing counties. The comparison set is constructed of 191 counties with no shale gas production in 2015. We also exclude 14 counties which we identify as metropolitan based on a cross-county comparison of 2010 population estimates, whereby we classify counties as metropolitan if they are within the top 10 percentile based on population; including only nonmetropolitan counties creates a more homogenous sample, precluding counties with large cities from excessively influencing estimates.[122]

5.2.5 Air quality and employment tradeoff metrics

To further explore the tradeoff between air quality and employment, we derive the metrics, job-years created per premature mortality or per life-year lost. A job-year is a full- or part-time job within the natural gas sector or spillover into the non-resource economy and that is held over a single year (rather than a sustained job over multiple years or a career). Life-years lost – the years of life forgone due to dying premature of life expectancy – are estimated based on the number of premature mortalities and using the annual, county-level population distribution by age class reported by the U.S. Census Bureau and annual, national life tables of the life expectancy by age class reported by the Center for Disease Control (CDC). We calculate the simple quotient of the cumulative number of job-years created per premature mortality (or life-year lost), which provides information regarding the systems-level tradeoff across the supply chain and natural gas cycle. We additionally estimate the marginal effect based on linearly regressing impacts aggregated by county or year. To capture the spatial variability of this tradeoff, we also calculate the cumulative number of job-years created less the life-years lost for each county.

5.2.6 Cross-impact elasticity

We calculate cross-impact elasticities, which provide information with respect to the sensitivity of each impact to changes in other impacts. We first regress one impact (y) on another (x) to determine the marginal effect (β) associated with each pair of impacts, as follows:

$$y = \beta_0 + \beta x \tag{39}$$

Then, we calculate elasticities (ε), based on the following equation:

$$\varepsilon = \frac{\bar{x}}{\bar{y}} \cdot \frac{\Delta y}{\Delta x} = \frac{\bar{x}}{\bar{y}} \cdot \beta \tag{40}$$

where $\bar{x}$ and $\bar{y}$ are average impacts. We derive elasticities for pairs of the following impacts across the supply chain from 2004 to 2016: air quality impacts in units of premature mortality, employment impacts in units of job-years, and climate impacts (global temperature change) in units of kelvin-years integrated over 100 years. To estimate pairwise elasticities, we use impact estimates stratified by year, as well as, employment and (source-resolved) air quality impact estimates stratified by county.

5.3 Spatial equity of air quality and employment impacts

Spatial equity refers to the distribution of benefits and costs on the basis of geographic location. We apply several standard measures of equity to further characterize the spatial distribution of air quality and employment impacts. We focus on the Gini coefficient (η) – an aggregate measure of equity across a system that compares each equity unit (e.g., county) to all other equity units and ranges in value from 0 for a completely equal distribution to 1 for complete inequity. We also use Lorenz curves, which depict the proportion of the total impact that is cumulatively borne by a given segment of the population, and that are axiomatically equivalent to Gini coefficients. Additional equity metrics are provided in the Supplemental Information (SI) (see Tables C2-3). These standard equity measures are useful as aggregate comparative benchmarks of the state of the entire system, and may be used for comparing policies, such as siting and emission standards. However, they are limited in their explanatory value of the underlying mechanisms of air quality and employment inequity and general interpretability for prescriptive decision making.

As depicted in Figure 17a, over the shale boom from 2004 to 2016, we find that there are high spatial inequities of cumulative mortality across receptor counties within producing states (η = 0.69) and all receptor counties (η = 0.80), with 80% of mortalities concentrated in 20% of receptor counties. However, spatial air quality inequities across counties have marginally declined over time (η = 0.77 in 2010 and η = 0.66 in 2016) and the associated cumulative air quality burden has correspondingly increased over time, reflecting the expanding development of natural gas infrastructure within the basin (Table S2, Figure B1). Similarly, as depicted in Figure 1b, there are high spatial inequities of cumulative employment across producing counties (η = 0.88) and all counties within the tristate (η = 0.72), with 80% of employment concentrated in 10% of counties within the tristate (Figure 1a). Spatial inequities have remained relatively constant over time (Table C3, Figure C2). For context, observed spatial inequities of air quality and employment impacts exceed that of income disparity in the U.S. (η = 0.39 for 2016).[143] While these spatial inequities can in part be explained by the geological constraints of where natural gas is abundant, the siting of upstream infrastructure and general pattern of development across the supply chain are also influenced by local economic development decisions and industry learning. Changing patterns of spatial equity over time are additionally affected by the increasing rate of development during the boom, as well as, the abundant but ephemeral nature of exploration and production activity and the relatively longer-term end use activity.

In this analysis, we do not consider the spatial equity dimension of climate change impacts, given that we use global metrics including global temperature change and monetized damages using the social cost of carbon, which are useful for characterizing ubiquitous greenhouse gases but unsuitable for spatially allocating impacts. There is also a nascent but expanding literature on spatially explicit climate change damages[144] that provide important insight into the spatial dimensions of climate equity.

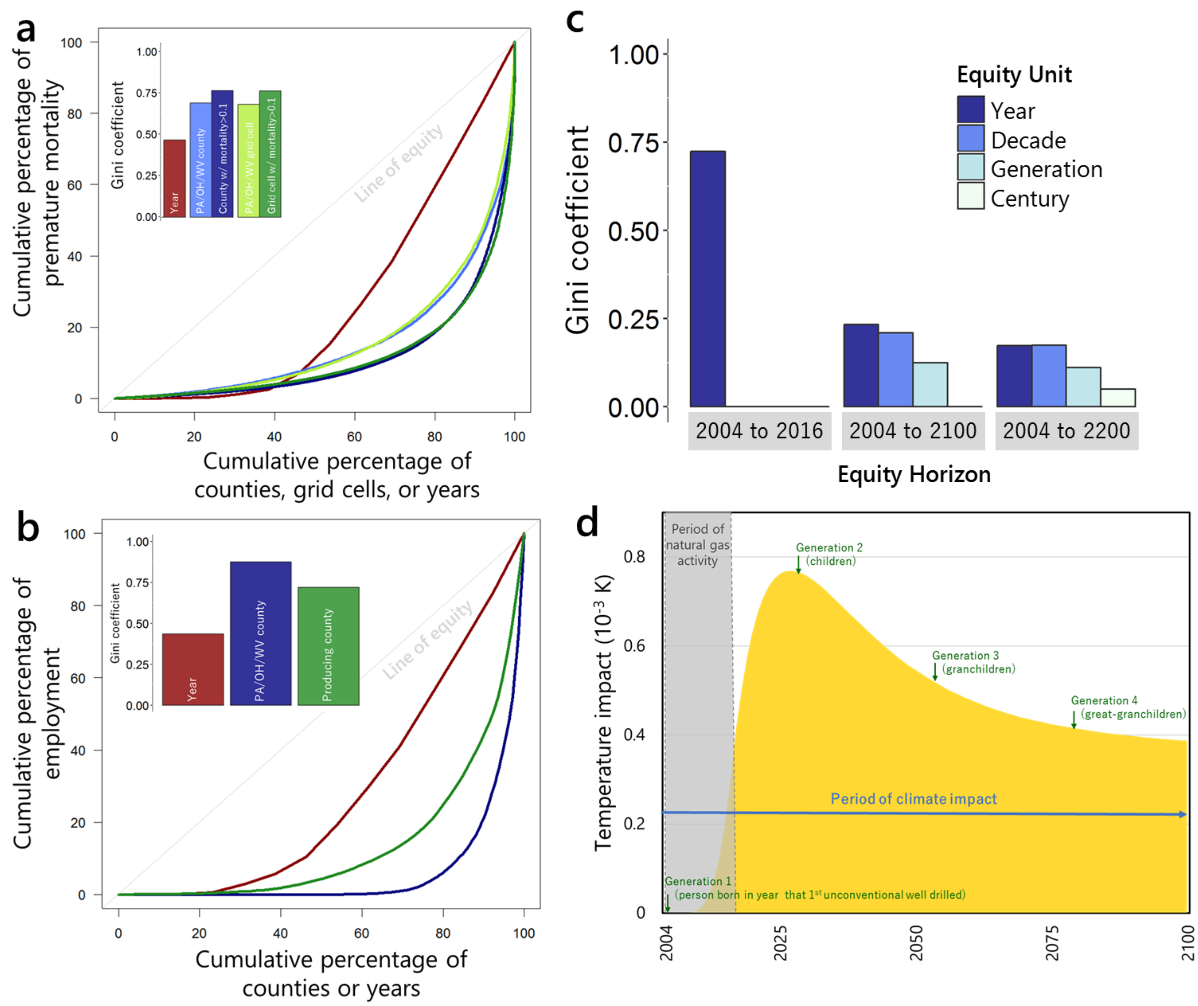

Figure 17. Spatial and temporal equity of air quality, employment, and climate change impacts across the natural gas supply chain from 2004 to 2016. Air quality equity across the natural gas supply chain from 2004 to 2016. (a) Lorenz curves and Gini coefficients representing premature mortality equity. The following are the equity units and samples considered (and the associated color used in the bar chart and Lorenz curves): temporal equity of premature mortality across years, based annual mean mortality estimates across six model specification [using three different source-receptor (S-R) models and two concentration-response (C-R) relationships] (red); spatial equity of cumulative mortality between source counties, based on AP3 S-R model and ACS C-R (light blue); spatial equity of cumulative mortality between all counties with mortality >0.1, based on AP3 S-R model and ACS C-R (dark blue); spatial equity of cumulative mortality between source 36 x 36 km grid cells, based on APSCA S-R model and ACS C-R (light green); and spatial equity of cumulative mortality between 36 x 36 km grid cells with mortality >0.1, based on APSCA S-R model and ACS C-R (dark green). (b) Lorenz curves and Gini coefficients representing employment equity. We consider the following equity units: years (dark red), Pennsylvania, Ohio, and West Virginia counties (dark blue), and producing counties (dark green). (c) Gini coefficients of climate change equity, assuming different equity units and equity horizons. (d) Climate impacts over time with an intergenerational equity framing.

5.4 Temporal equity of natural gas development

Temporal equity refers to the distribution of benefits and costs over time, and it is conceptually inclusive of intergenerational equity. Within the natural gas system, it has been observed that short-lived air quality and employment impacts track with the boom-and-bust cycle, whereas climate impacts persist for generations well beyond the period of natural gas activity.[124] Here, we apply a suite of systems-level equity metrics, similar to those used to evaluated spatial equity, as well as additional metrics to further elucidate temporal tradeoffs.

Key to operationalizing temporal equity is the use of impact metrics that explicitly estimate temporally lagged responses. Air quality and employment impacts in terms of premature mortality and jobs, respectively, temporally align with the natural gas activity, with a potential (but relatively minor) employment lag effect, and thus, are appropriate for use in equity analysis. A temporally-resolved climate change metric, such as the global mean temperature change that translates pulses of GHG emissions into future temperature response, is well-suited for equity analyses; this is in contrast to commonly employed time-integrated metrics that obscure when impacts are realized, such as global warming potential or monetized damages based on the social cost of carbon, which further assumes discounts rates and implies a modern economic and monetary structure. Characterization of temporal equity also requires normatively defining the equity horizon – the period over which equity is assessed (e.g., 15, 50, 100 years) – and the equity units – the period over which impacts are integrated (e.g., year, decade, generation, century).

As shown in Figures Figure 17a-b, over the period of natural gas activity from 2004 to 2016, we observe a high disparity between years with respect to air quality (η = 0.47) and employment (η = 0.44) impacts, reflecting the rapid increase in development over the boom period; depending on the future development pathway, temporal air quality and employment inequities will likely decrease as the industry contracts. As depicted in Figure 1c, the temporal dispersion of climate impacts decreases as we increase the equity horizon or equity unit (Table C4, Figure C4); for example, there is greater equity between years in the long-term (2004 to 2100) (η = 0.21) relative to the near-term (2004 to 2016) (η = 0.73) because of rapidly increasing natural gas activity in the near-term, the lagged climate response, the relatively short atmospheric lifetime of CH_4, and persistence of CO_2 in the atmosphere. Comparing the cumulative climate response in the near- to long-terms, we find that over an equity horizon out to 2100, the cumulative response in the long-term is 100 times that in the near-term (Figure C3). Figure 17d provides a potentially more salient representation of the residual impacts of near-term production and consumption on future generations, that is not otherwise captured by quantitative system measures; we depict an intergenerational framing of air quality, climate change, and employment impacts, whereby we trace the descendants of a child born in 2004, the year in which the first shale well was drilled in the Appalachian basin.

Although we use quantitative measures of temporal equity to capture the absolute state of the system, they are more interpretable when used as comparative measures of different states of the system, such as under different policy interventions. For example, changes in the Gini coefficient can capture the differential effect of relatively marginal CH_4 emissions abatement that may affect near-term warming rates, as compared to CO_2 emission reductions through more systemic interventions to transition the energy system away from fossil fuels that result in benefits derived largely by future generations. The Gini coefficient, however, does not capture the effect of reducing CH_4 emissions as a strategy to avoid or delay of reaching "tipping points" in the climate system, irreversible thresholds with drastic consequences.[115]

5.5 Distributional equity with respect to racial and socioeconomic subpopulations

Several studies have found evidence of racial and socioeconomic disparities in the distribution of environmental hazards and locally unwanted land uses.[145] Here, we perform analyses to elucidate the decomposition of impacts and natural gas activity across different subpopulations based on income, race, and poverty levels, which have implications for policy design and decision making in the realm of both environmental justice and local economic development.

5.5.1 Distributional equity of air quality impacts

We assess the distribution of premature mortality and monetized air quality damages across subpopulations on the basis of race, income level, and poverty level. Specifically, we estimate the subpopulation-weighted mortality and mortality rates (m) – the mortality induced by natural gas activity for each subpopulation relative to the total subpopulation in the contiguous U.S. (in units of premature mortality per 100,000).

We estimate mortality rates by income level, finding that there is not a noticeable spread across income levels, as shown in Figure 18c (Table C6). We also find that populations below the poverty line do not experience higher mortality rates induced by natural gas activity, but rather we observe marginally higher annual mortality rates (6 to 10%) among populations above the poverty line relative to those below the poverty line, as depicted in Figure 18b. This democratization of air quality impacts regardless of income and poverty level reflects the transport of air pollutants and the abundance of natural gas activity. A major caveat of these analyses is that they do not account for differences in baseline mortality rates nor access to healthcare across income and poverty levels, given data limitations. As provided in Figure 18d, however, we observe a trend of increasing income corresponding to decreasing damages (normalized by income) across counties within Pennsylvania, Ohio, and West Virginia, which demonstrates the higher health burden on lower income communities

We also estimate population-weighted premature mortality and mortality rates by race, accounting for race-based differences in baseline mortality. As shown if Figure 18a, we find that annual mortality rates induced by natural gas activity are marginally higher (13 to 20%) for white (m = 0.10) than black or African American (m = 0.10) populations, and annual mortality rates for white and black or African American populations are much higher (73 to 81%) than for Asian or Pacific Islander (m = 0.02) and American Indian or Alaska Native (m = 0.02) populations. This result is unsurprising, given that communities within the Appalachian basin are predominately white and other receptor communities in the Northeast region have a higher proportion of black or African American populations than many other regions in the contiguous U.S.

While transport of air pollutants results in a democratizing effect with respect to mortality rates, inequities on the basis of race and income may be borne out in infrastructure siting decisions, as evaluated in the subsequent section. There are noteworthy limitations to these finding regarding distributional equity of air quality impacts. The county-level resolution of this analysis may not reveal inequities that are observable only at finer spatial resolutions, and a population-level analysis only demonstrates average effects and not inequities experienced by individual communities. We also only evaluate equity as it relates to premature mortality from primary $PM_{2.5}$ and precursors for secondary $PM_{2.5}$, and we would anticipate differing equity implications for other species and health outcomes. Additionally, there are nontrivial methodological limitations associated with conducting equity based on health outcomes, such as premature mortality, rather than exposure; this is due to data limitations and a lack of robust evidence in the broader literature regarding differences in baseline mortality rates by race, income, and poverty level and the underlying mechanisms which account for such differences.

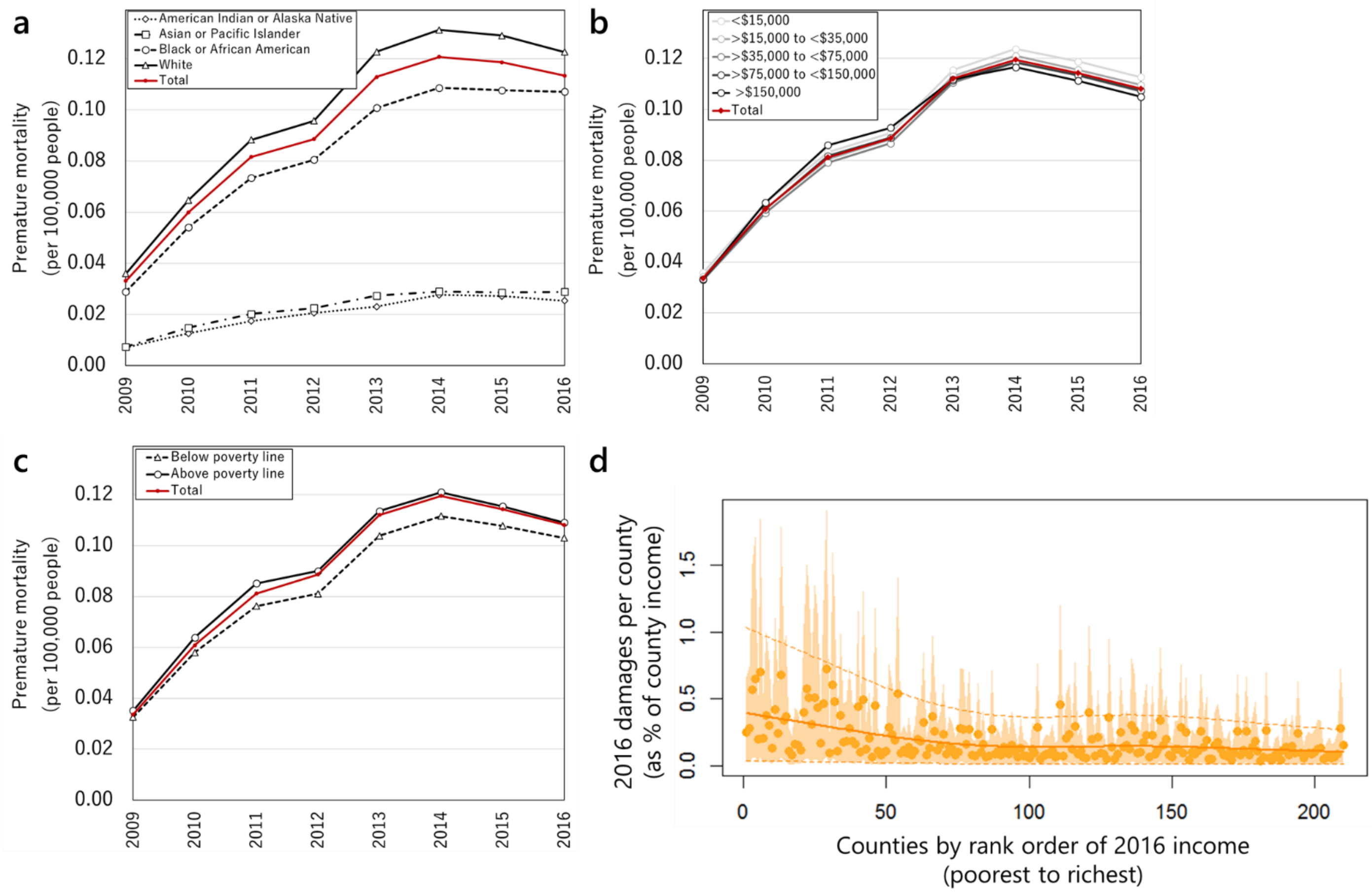

Figure 18. Distributional equity of air quality impacts by race, income, and poverty levels. Mortality rates associated with natural gas development across the supply chain from 2009 to 2016 by (a) race, (b) income, and (c) poverty level. (d) 2016 mortality-related damages as a percentage of county personal income ranked by income from poorest to richest. Includes counties within Pennsylvania, Ohio, and West Virginia. Dots and shaded regions represent mean and 95% confidence interval, respectively, and reflect uncertainty in the value of statistical life while assuming fixed baseline scenario assumptions regarding mortality. Solid and dashed lines are spline regressions through mean and 95% CI values.

5.5.2 Distributional equity of labor market impacts

We explore the distributional effects of the natural gas boom on income inequality and poverty rates within local labor markets. Regionally, in both producing and non-producing counties, there have been declines in income disparities and the percentage of the population below the poverty line between 2005 and 2015. However, we find statistically significant mean differences between producing and non-producing counties in the change in poverty rates and income Gini coefficient.

Table 3 displays the marginal effects from the natural gas boom on poverty and income disparity measures. To discern the effects associated with different intensities in natural gas activity, we specify two treatment sets – a full treatment set comprised of all 90 producing counties and a top treatment set comprised of the top 25% of producing counties. We find that the shale boom is associated with an absolute decline in the percentage below the poverty line of 1.08 (SE ± 24%) among all producing counties and 1.72 (SE ± 23%) among the top producing counties. This is equivalent to a 9.9% (SE ± 17%) and 14.1% (SE ± 20%) decline in the poverty rates among all and top producing counties, respectively. Our findings are consistent with several other studies showing that energy booms lower the poverty rate, at least in the short-run.[88,117] It has further been shown in other studies that poverty rates increase during resource declines, and our findings are not reflective of long-run effects over the natural gas boom-and-bust cycle; as an analogue, the 1970s coal mining boom in the Appalachian region decreased poverty, but the 1980s bust reversed this reduction.[146]

We additionally find that the shale boom is associated with a minimal, but statistically significant, absolute decline in the income Gini coefficient of 0.01 (SE ± 40%) or a 3.5% (SE ± 35%) decline in income disparity among all producing counties. However, we do not observe a statistically significant change in income disparity among the top producing counties. There is mixed evidence in the labor market literature with respect to the effect of energy booms on income inequality; for example, the recent energy boom in Western Canada generally increased local inequality with a U-shaped growth across income distributions.[117] The distribution of the gains from energy booms depends on the skills of local residents and where they fall in the income distribution, the extent of integration between local and regional labor markets, and the extent of spillover.[147] Our finding of a minimal or insignificant change may indicate that income effects are almost equally distributed across the population. Countervailing evidence suggests that labor demand from the shale gas boom is filled by transient workers because long-term residents in rural communities in Appalachia may not have the requisite skills and training. Based on Security and Exchange Commission (SEC) filings in 2017 of the top publicly-traded producing firms in Appalachia, the median employee compensation ranged from $76,000 to $160,000, further suggesting a more skilled labor force, and there is evidence of vertical inequities within producing firms (Table C11).

Table 3. Distributional equity of labor market impacts. Changes in poverty level and income Gini coefficients between 2005 and 2015 from natural gas activity.

	Full Treatment				Top Treatment			
Variable	Estimate	Std. Error	P(>t)		Estimate	Std. Error	P(>t)	
△ % below poverty level	-1.084	(0.257)	0.000	***	-1.723	(0.396)	0.000	***
Intercept	1.683	(0.118)	<2e-16	***	1.683	(0.118)	<2e-16	***
Log △ % below poverty level	-0.099	(0.017)	0.000	***	-0.141	(0.028)	0.000	***
Intercept	0.143	(0.009)	<2e-16	***	0.143	(0.009)	<2e-16	***
△ Gini coefficient	-0.011	(0.004)	0.009	**	0.001	(0.006)	0.925	
Intercept	-0.080	(0.002)	<2e-16	***	-0.080	(0.002)	<2e-16	***
Log △ Gini coefficient	-0.035	(0.012)	0.002	**	-0.014	(0.02)	0.499	
Intercept	-0.169	(0.006)	<2e-16	***	-0.169	(0.006)	<2e-16	***
Observations	258				191			

Significance codes: ***p<0.001, **p<0.01, *p<0.05, .p<0.1
Huber-White robust standard errors are reported.

5.5.3 Racial and socioeconomic disparity in the geographic distribution of natural gas activity

We assess racial and socioeconomic disparities in the distribution of natural gas activity and infrastructure, applying standard spatial coincidence methods used in the environmental justice literature. We focus on upstream activity; however, an analogous method can be applied to evaluate disparities in the siting of other infrastructure, such as pipelines and power plants.

Comparing mean estimates of several demographic variables, we find that there are statistically significant differences between producing and non-producing counties (Table C10). Population-weighted differences between producing and nonproducing counties with respect to racial variables (i.e., percentage nonwhite, percentage black) are nontrivial (e.g., nonwhite percentages in 2016 for producing and nonproducing counties are 8.5% and 25.7%, respectively); however, differences in per capita income, median household income, and poverty rates are minimal. Assessing the relative importance of demographic characteristics in accounting for these disparities, as provided in Table 4, we find that over the period from 2010 to 2016, the percentage nonwhite (odds ratio = 0.926, p-value ≤ 0.00) and the log median household income (odds ratio = 0.005, p-value ≤ 0.00), and are statistically significant predictors of the geographic location of production, but the percentage below the poverty line (odds ratio = 0.964, p-value ≤ 0.15) is not. As the nonwhite percentage decreases or the median household income decreases, there is an increasing probability that a county is producing natural gas. Observed disparities based on this spatial coincidence approach may have both environmental justice and local economic development implications, but only if environmental risks or economic benefits are reasonably correlated with the geographic location of production. When benefits or costs are dispersed and are not fully borne by local communities, such as in the transport of air pollutants, a simplistic spatial coincidence approach is limited.

Table 4. Racial and socioeconomic disparity in the geographic distribution of natural gas activity. Spatial coincidence of production and demographic variables, using a logistic regression approach.

	2010			2016			2010 to 2016		
Variable	Estimate	Odds Ratio	P-value	Estimate	Odds Ratio	P-value	Estimate	Odds Ratio	P-value
Percent nonwhite	-0.132	0.876	0.060.	-0.046	0.955	0.108	-0.077	0.926	0.000***
Log median household income	-8.447	0.000	0.009**	-8.685	0.000	0.000***	-5.213	0.005	0.000***
Percent below poverty line	-0.026	0.975	0.800	-0.118	0.888	0.100.	-0.036	0.964	0.147
Constant	30.672	2.1E+13	0.022.	35.069	1.7E+15	0.000***	20.031	5.0E+08	0.000***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.6 Pairwise air quality, climate change, and labor market tradeoffs

We are interested not only in the spatial, temporal, and distributional equities of individual impacts, as described in previous sections, but also pairwise comparisons of impacts that reveal the implied tradeoffs of natural gas development decisions.

5.6.1 Cross-impact elasticity

We estimate cross-impact elasticities (ε), which provide information regarding the sensitivity of each impact to changes in other impacts, as provided in Table 5. At a systems level, based on cumulative impacts across the supply chain over the development period, each pairwise cross-impact elasticity is near unit elastic, which is an intuitive result given that all impacts are some function of the intensity of natural gas activity. The employment elasticity of premature mortality (ε = 1.08) and cumulative global temperature change over a 100-year integration period (ε = 1.26) can be interpreted as a 1% increase in employment is associated with a 1.08% increase in air quality impacts and a 1.26% increase in climate impacts. Similarly, the premature mortality elasticity of global temperature change over a 100-year integration period (ε = 1.17) is slightly elastic, with a 1% increase in air quality impacts corresponding to a 1.17% increase in climate impacts. We would expect that as zero- or low-carbon energy technologies increasingly displace natural gas in the energy system, the employment elasticity of premature mortality and global temperature change would iterate towards becoming inelastic, with the decoupling of emissions and employment.

Table 5. Cross-impact elasticities.

Observations aggregated by year (n=13)	ε
Employment elasticity of air quality	1.08
Employment elasticity of climate change	1.26
Air quality elasticity of climate change	1.17
Observations aggregated by county (n=210)	
Employment elasticity on air quality	0.48

5.6.2 Air quality and employment tradeoffs

We derive additional metrics to further explore and provide salience to the tradeoff between air quality and employment impacts experienced in the near-term. This tradeoff varies by supply chain segment, temporally, and spatially, and is subject to uncertainty with respect to air quality and employment modeling specifications.

At a systems level across the supply chain from 2004 to 2016, the implied tradeoff, based on the simple quotient of employment and air quality impacts, is 217 job-years per premature mortality, with a range of 100 to 410 job-years per premature mortality reflecting uncertainty in the air quality model functional form and concentration-response (C-R) relationship, as depicted in Figure 19a. The mean marginal effect of air pollution on employment, whereby we regress employment on premature mortality based on annual average estimates, is 157 (95% CI 146 to 167) job-years per premature mortality (Table C13, Figure C6-10). Translating premature mortality into life-years lost, the tradeoff can also be expressed as 3 job-years created per life-year lost, with a range from <1 to 7 job-years per life-year reflecting uncertainty in the C-R relationship and employment estimates, as shown in Figure 19b.

Air quality impacts are more spatially dispersed than employment effects, with communities in closest proximity to natural gas infrastructure experiencing the highest mortality rates. The air quality and employment tradeoff varies spatially among producing counties, ranging from 1 to 16,000 job-years per premature mortality.[124] To further explore this spatial tradeoff, as depicted in Figure 19c, we estimate the number of life-years lost minus the number of job-years created by county, finding a range from -1,100 to 4,400. In most producing counties, more job-year years are created than life-years lost, whereas, in all other counties there are more life-years lost than job-years created.

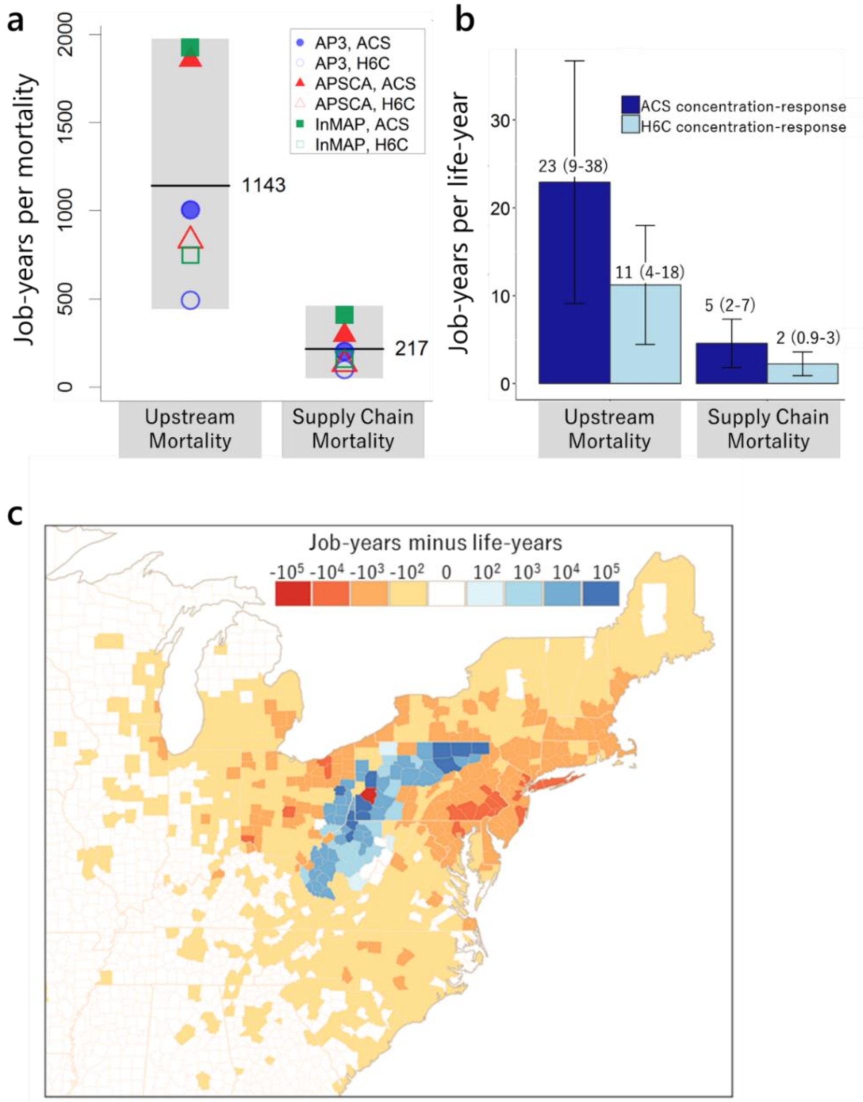

Figure 19. Air quality and employment tradeoffs across the natural gas supply chain from 2004 to 2016. (a) Job-years per premature mortality. Based on job-year (mean) and premature mortality estimates from 2004 to 2016 reported in Mayfield et al. (forthcoming). Solid points represent estimates based on American Cancer Society (ACS) C-R relationship, and open points represent estimates based on Harvard Six Cities (H6C) C-R relationship. Circle, triangle, and square points represent premature mortality estimates based on AP3, APSCA, and InMAP, respectively. Black lines represent average annual mortality across all six specifications. Grey shaded regions represent range of annual estimates. (b) Job-years per life-year lost. Life-years lost are based on premature mortality estimates using AP3. Dark and light blue bars represent life-year lost estimates based on ACS and H6C C-R relationships, respectively. The error bars represent the 95% confidence interval, reflecting uncertainty in the job-year estimates. (c) Spatial distribution of air quality and employment tradeoff in units of job-years created minus life-years lost. Based on cumulative impacts across the supply chain from 2004 to 2016. Life-years lost are based on premature mortality estimates using AP3 and ACS C-R relationship, and job-years are based on mean estimates.

5.7 Conclusions

The intent of this study is to provide a descriptive evaluation of the state of systems-level equity. We find that there are high temporal and spatial inequities with respect to cumulative air and employment impacts. With respect to distributional equity of air quality impacts, we do not observe a disparity in mortality rates across subpopulations on the basis of income and poverty; however, there is a trend of increasing income corresponding to decreasing damages, which demonstrates the higher health burden on lower income communities. With respect to distributional equity of labor markets, we find statistically significant declines in the income disparity and poverty rates in producing counties. Pairwise comparisons of air, climate, and employment impacts reveal that they are highly correlated.

6 Operationalizing Distributional, Spatial, and Temporal Equity Objectives in Energy System Planning

Analytical approaches for evaluating and planning future energy system pathways often adopt efficiency or utilitarian objectives. In this study, we develop a multiobjective optimization model, incorporating several objective functions which instill different concepts of spatial, temporal, and distributional equity. We apply this framework to future natural gas development decisions and pathways in the Appalachian basin. We find that there are a range of conflicts and agreements between different equity objectives, as well as, between equity and cumulative air, climate, and employment impact objectives in a fossil-fuel dominated energy system. For example, equity objectives which seek to minimize air quality impacts on populations below the poverty line, minority populations, rural communities, and communities with existing environmental burdens all generate different optimal natural gas system pathways.

6.1 Introduction

Expanding upon an existing multiobjective energy optimization model presented in Chapter 4, we formulate equity objectives that reflect additional societal values, beyond the traditional economic efficiency maximization or utilitarian criteria. The purpose of this study is to operationalize concepts of distributional, spatial, and temporal equity for future energy system planning and policy design. We model alternative, theoretical natural gas development pathways by optimizing spatially and temporally explicit air quality, climate change, and employment impacts and equity with respect to decisions regarding the magnitude, timing, and location of preproduction, production, and industrial, residential, commercial, and electric consumption. We additionally specify policy scenarios, including infrastructure siting and planning, emissions abatement, and renewable integration, and demonstrate the relative effect and pathway dependence of implementing these policies. For contextualization, we focus on natural gas development decisions and pathways in the Appalachian basin.

6.2 Model formulation

To facilitate energy systems planning and environmental policy evaluation, we formulate a multiobjective linear optimization model that incorporates two types of objectives, equity and cumulative impacts. We build upon an existing formulation in which cumulative air, climate, and employment impact objectives are optimized (see Chapter 4). Herein, we provide a general formulation for energy systems, as well as parameterize the model for the case of natural gas development in the Appalachian basin. We additionally specify modifications in which we integrate emissions abatement and renewables to displace natural gas. The optimization model is implemented in the General Algebraic Modeling System (GAMS) and uses CPLEX to solve the linear program. Table 2 includes indices, decision variables, and parameters.

Table 6. Indices, decision variables, and parameters.

Indices			
$i = \{1, \dots, \mathcal{I}\}$	county of natural gas activity	$jpov \subset j$	county with high population below poverty line
$ia, ib = \{1, \dots, \mathcal{I}\}$	aliases for i	$jrace \subset j$	county with high non-white population
$j = \{1, \dots, \mathcal{J}\}$	county of impact	$jrural \subset j$	county with high rural population
$ja, jb = \{1, \dots, \mathcal{J}\}$	aliases for j	$s = \{1, \dots, \mathcal{S}\}$	year of impact
$jlow \subset j$	county with high population that is low income	$t = \{1, \dots, \mathcal{T}\}$	year of natural gas activity
$jnonattainment \subset j$	county classified as NAAQS non-attainment	$ta, tb = \{1, \dots, \mathcal{T}\}$	aliases for t
Decision variables			
$x_{i,t}^{commercial}$	commercial natural gas volume consumed (mmcf)	$x_{i,t}^{producing}$	number of producing wells
$x_{i,t}^{electric}$	electric generation natural gas volume consumed (mmcf)	$x_{i,t}^{residential}$	residential natural gas volume consumed (mmcf)
$x_{i,t}^{industrial}$	industrial natural volume consumed (mmcf)	$x_{i,t}^{spud}$	number of spud wells
$\sigma_{1,ia,ib}$, $\sigma_{2,ia,ib}$, $\sigma_{1,ja,jb}$, $\sigma_{2,ja,jb}$	spatial slack variables	$\phi_{ia,ib}$, $\phi_{ja,jb}$	spatial binary variables
$\sigma_{1,ta,tb}$, $\sigma_{2,ta,tb}$	temporal slack variables	$\lambda_{ta,tb}$	temporal binary variables
Parameters (primary)			
$\mathrm{a}_{i,j,t}^{\mathrm{AIR}}$	marginal premature mortalities per unit natural gas activity (per mmcf or well)	$\mathcal{M}$	a very large number
$\mathrm{a}_{s,t}^{\mathrm{CLIMATE}}$	marginal temperature impact per unit natural gas activity (milliKelvin per mmcf or well)	m	number of counties
$\mathrm{a}^{\mathrm{EMPLOY}}$	marginal job-years per unit natural gas activity (per mmcf or well)	n	number of years
$\mathrm{G}^{\mathrm{MAX}}$	maximum Gini coefficient	$\mathrm{p}_{i,j,t}^{\mathrm{AIR}}$	marginal premature mortalities for population below the poverty line per unit natural gas activity (per mmcf or well)
$\mathrm{l}_{i,j,t}^{\mathrm{AIR}}$	marginal premature mortalities for low income population per unit natural gas activity (per mmcf or well)	$\mathrm{r}_{i,j,t}^{\mathrm{AIR}}$	marginal premature mortalities for non-white populations per unit natural gas activity (per mmcf or well)

6.2.1 Decision variables

We incorporate decision variables related to six types of upstream and end use natural gas activities – the number of producing and spud wells, and residential, commercial, industrial, and electricity consumption volumes. We do not include midstream processes, including gathering, processing, transmission, storage, and distribution, in this version of the model.

The decision variables, which are non-negative and continuous, reflect the magnitude of natural gas activity for each county i and year t. We consider a future natural gas activity time horizon ($t = 2017, .., 2030$), and natural gas activity within 210 counties in Pennsylvania, Ohio, and West Virginia. Given that impacts may be spatially and temporally dispersed from natural gas activity, we assume a longer impact horizon ($s = 2017, .., 2100$) to accommodate delayed and persistent climate change impacts, and we consider impact counties to include all of those within the continental U.S. to allow for spatial transport of air pollutant impacts.

We assume that consumption is spatially and temporally fixed to projected future levels, with the exception of the timing and location of additional electric generation and industrial end use beyond 2016 levels which are assumed to be variable. We also allow the magnitude, timing, and location of future production to be variable, while accounting for the residual production from historical wells.

6.2.2 Energy system constraints

We specify multiple energy system constraints related to operational and technological efficiencies, consumption patterns, and natural gas, wind, and solar resources, which are further described in Chapter 4.

6.2.3 Cumulative air, climate, and employment impact objectives

We adopt three types of cumulative impact objectives; the reader is directed to Chapter 4 which describes the formulations and parameterization of these objectives. The purpose of defining cumulative impact objectives is to explore and demonstrate the potential conflict between efficiency and equity, which are fundamentally different decision-making frameworks. Herein, we refer to air, climate, and employment regimes to mean the energy system pathways under which air quality, climate change, or employment impacts are optimized, respectively.

The air quality objective function is to minimize cumulative impacts, alternatively specified in terms of premature mortalities and monetized damages, resulting from primary fine particulate matter ($PM_{2.5}$) and secondary $PM_{2.5}$ formed from the atmospheric oxidation of nitrogen oxides (NO_x) and volatile organic compounds (VOCs) emissions. The climate change objective function is to minimize cumulative climate change impacts, alternatively specified in terms of global temperature change and monetized damages, resulting from emissions of greenhouse gases (GHG), including methane (CH_4) and carbon dioxide (CO_2). The employment objective function is to maximize cumulative impacts, alternatively specified in terms of job-years and wages, resulting from upstream natural gas activity.

6.2.4 Equity objective functions

Equity, which is inherently subjective, may be variably and divergently defined and operationalized. Here, we define six broad categories of equity: spatial, income / poverty, race, rural, existing environmental burden, and temporal. Across these categories, we specify fifteen objective functions or equity rules. While we tailor these categories and objectives for the regional case of the Appalachian basin, variants are more broadly applicable to the U.S. energy system. Key inputs for defining some equity objectives are provided in Figure 20.

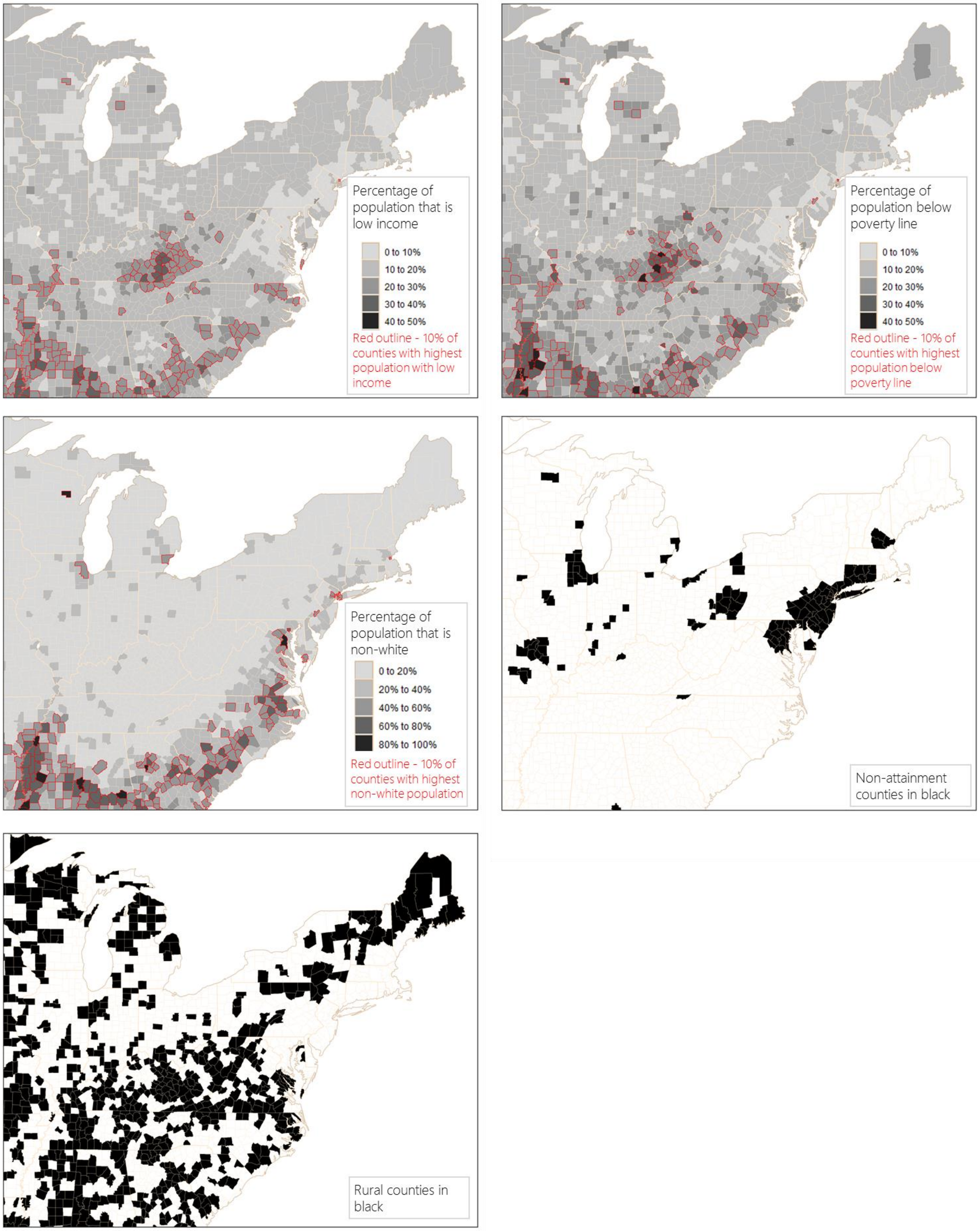

Figure 20. Maps of low income, below the poverty line, non-white, nonattainment, and rural subpopulations. Low income populations are people with an income of <$15,000. Non-attainment counties are those classified under the National Ambient Air Quality Standards. Rural counties are the third of counties with the highest percentage of rural populations.

1. Minimize temporal inequity

We formulate four temporal inequity objectives: minimize air quality temporal inequity (1a), minimize employment temporal inequity (1b), minimize combined air-employment temporal inequity (1c), and minimize combined air-climate-employment temporal inequity (1d).

$$\text{Equity objective 1a:} \qquad \text{Min}\left\{\frac{\sum_{ta,tb}\left|\mathrm{Z}_{ta}^{\mathrm{AIR}}-\mathrm{Z}_{tb}^{\mathrm{AIR}}\right|}{2\mathrm{n}\cdot\mathrm{Z}^{\mathrm{AIR}}}\right\} \tag{41}$$

$$\text{Equity objective 1b:} \qquad \text{Min}\left\{\frac{\sum_{ta,tb}\left|\mathrm{Z}_{ta}^{\mathrm{EMPLOY}}-\mathrm{Z}_{tb}^{\mathrm{EMPLOY}}\right|}{2\mathrm{n}\cdot\mathrm{Z}^{\mathrm{EMPLOY}}}\right\} \tag{42}$$

$$\text{Equity objective 1c:} \qquad \text{Min}\left\{\frac{\sum_{ta,tb}\left|\mathrm{Z}_{ta}^{\mathrm{AIREMPLOY}}-\mathrm{Z}_{tb}^{\mathrm{AIREMPLOY}}\right|}{2\mathrm{n}\cdot\mathrm{Z}^{\mathrm{AIREMPLOY}}}\right\} \tag{43}$$

where ta and tb are indices that are aliases for activity year t, and n is the number of activity years.

These objective functions (1a-1c) are based on the Gini index, an aggregate measure of equity across the system that compares each year to all other years. We reformulate the objectives into sets of linear constraints. Here, we provide an example for employment temporal equity.

$$\mathrm{Z}_{ta}^{\mathrm{EMPLOY}} - \mathrm{Z}_{tb}^{\mathrm{EMPLOY}} = \sigma_{1,ta,tb}^{\mathrm{EMPLOY}} - \sigma_{2,ta,tb}^{\mathrm{EMPLOY}} \qquad \forall\, ta = 1,\dots,\mathcal{T};\ tb = 1,\dots,\mathcal{T} \tag{44}$$

$$\sigma_{1,ta,tb}^{\mathrm{EMPLOY}} \le \mathcal{M}\cdot\lambda_{ta,tb}^{\mathrm{EMPLOY}} \qquad \forall\, ta = 1,\dots,\mathcal{T};\ tb = 1,\dots,\mathcal{T} \tag{45}$$

$$\sigma_{2,ta,tb}^{\mathrm{EMPLOY}} \le \mathcal{M}\cdot(1-\lambda_{ta,tb}^{\mathrm{EMPLOY}}) \qquad ta = 1,\dots,\mathcal{T};\ tb = 1,\dots,\mathcal{T} \tag{46}$$

$$\textstyle\sum_{ta,tb}(\sigma_{1,ta,tb}^{\mathrm{EMPLOY}} - \sigma_{2,ta,tb}^{\mathrm{EMPLOY}}) \le \mathrm{G}^{\mathrm{MAX}}\cdot 2\mathrm{n}\cdot\sum_t \mathrm{Z}_t^{\mathrm{EMPLOY}} \tag{47}$$

where $\sigma_{1,ta,tb}^{\mathrm{EMPLOY}}$ and $\sigma_{2,ta,tb}^{\mathrm{EMPLOY}}$ are slack variables, $\lambda_{ta,tb}^{\mathrm{EMPLOY}}$ are binary variables, and $\mathcal{M}$ is a very large number.

We additionally specify an objective function which accounts for near- and long-term tradeoffs.

2. Minimize impacts on poor populations

We formulate four poverty and low income objectives: minimize air quality impacts on populations in the lowest income class (2a), on populations below poverty line (2b), in 10% of counties with highest proportion of population in lowest income class (2c), and in 10% of counties with highest proportion of population below poverty line (2d).

$$\text{Equity objective 2a:}\quad \mathrm{Min}\Big\{\sum_{i,j,t}\big(\mathrm{l}_{i,j,t}^{\mathrm{AIR,PREPRODUCTION}} \cdot x_{i,t}^{spud} + \mathrm{l}_{i,j,t}^{\mathrm{AIR,PRODUCTION}} \cdot x_{i,t}^{producing} + \mathrm{l}_{i,j,t}^{\mathrm{AIR,INDUSTRIAL}} \cdot x_{i,t}^{industrial} + \mathrm{l}_{i,j,t}^{\mathrm{AIR,COMMERCIAL}} \cdot x_{i,t}^{commercial} + \mathrm{l}_{s,t}^{\mathrm{AIR,RESIDENTIAL}} \cdot x_{i,j,t}^{residential} + \mathrm{l}_{i,j,t}^{\mathrm{AIR,ELECTRIC}} \cdot x_{i,j,t}^{electric}\big)\Big\} \tag{48}$$

where the coefficients $\mathrm{l}_{i,j,t}^{\mathrm{AIR}}$ are the premature mortalities for populations in the lowest income class (<$15,000) in receptor county *j* associated with emissions in source county *i* per unit of natural gas activity.

$$\text{Equity objective 2b:}\quad \mathrm{Min}\Big\{\sum_{i,j,t}\big(\mathrm{p}_{i,j,t}^{\mathrm{AIR,PREPRODUCTION}} \cdot x_{i,t}^{spud} + \mathrm{p}_{i,j,t}^{\mathrm{AIR,PRODUCTION}} \cdot x_{i,t}^{producing} + \mathrm{p}_{i,j,t}^{\mathrm{AIR,INDUSTRIAL}} \cdot x_{i,t}^{industrial} + \mathrm{p}_{i,j,t}^{\mathrm{AIR,COMMERCIAL}} \cdot x_{i,t}^{commercial} + \mathrm{p}_{s,t}^{\mathrm{AIR,RESIDENTIAL}} \cdot x_{i,j,t}^{residential} + \mathrm{p}_{i,j,t}^{\mathrm{AIR,ELECTRIC}} \cdot x_{i,j,t}^{electric}\big)\Big\} \tag{49}$$

where the coefficients $\mathrm{p}_{i,j,t}^{\mathrm{AIR}}$ are the premature mortalities for populations below the poverty line per unit activity.

$$\text{Equity objective 2c:}\quad \mathrm{Min}\Big\{\sum_{i,j\in jlow,t}\big(\mathrm{a}_{i,j,t}^{\mathrm{AIR,PREPRODUCTION}} \cdot x_{i,t}^{spud} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,PRODUCTION}} \cdot x_{i,t}^{producing} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,INDUSTRIAL}} \cdot x_{i,t}^{industrial} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,COMMERCIAL}} \cdot x_{i,t}^{commercial} + \mathrm{a}_{s,t}^{\mathrm{AIR,RESIDENTIAL}} \cdot x_{i,j,t}^{residential} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,ELECTRIC}} \cdot x_{i,j,t}^{electric}\big)\Big\} \tag{50}$$

where the coefficients $\mathrm{a}_{i,j,t}^{\mathrm{AIR}}$ are the total premature mortalities per unit activity. The subset $jlow$ are the 10% of receptor counties with the highest proportion of their populations in the lowest income class

(<\$15,000); we define the subset based on county-level populations by income class as reported by the US Census Bureau for 2016.

$$\text{Equity objective 2d:} \quad \mathrm{Min}\Big\{\textstyle\sum_{i,j\in jpov,t}\big(\mathrm{a}_{i,j,t}^{\mathrm{AIR,PREPRODUCTION}} \cdot x_{i,t}^{spud} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,PRODUCTION}} \cdot x_{i,t}^{producing} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,INDUSTRIAL}} \cdot x_{i,t}^{industrial} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,COMMERCIAL}} \cdot x_{i,t}^{commercial} + \mathrm{a}_{s,t}^{\mathrm{AIR,RESIDENTIAL}} \cdot x_{i,j,t}^{residential} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,ELECTRIC}} \cdot x_{i,j,t}^{electric}\big)\Big\} \quad (51)$$

where the subset $jpov$ is the 10% of receptor counties with the highest proportion of their populations below the poverty line; we define the subset based on county-level poverty rates as reported by the US Census Bureau for 2016.

3. Minimize impacts on minority populations

We formulate two race objectives: minimize air quality impacts on non-white populations (3a) and in the 10% of counties with the highest proportion of their population that is non-white (3b).

$$\text{Equity objective 3a:} \quad \mathrm{Min}\Big\{\textstyle\sum_{i,j,t}\big(\mathrm{r}_{i,j,t}^{\mathrm{AIR,PREPRODUCTION}} \cdot x_{i,t}^{spud} + \mathrm{r}_{i,j,t}^{\mathrm{AIR,PRODUCTION}} \cdot x_{i,t}^{producing} + \mathrm{r}_{i,j,t}^{\mathrm{AIR,INDUSTRIAL}} \cdot x_{i,t}^{industrial} + \mathrm{r}_{i,j,t}^{\mathrm{AIR,COMMERCIAL}} \cdot x_{i,t}^{commercial} + \mathrm{r}_{s,t}^{\mathrm{AIR,RESIDENTIAL}} \cdot x_{i,j,t}^{residential} + \mathrm{r}_{i,j,t}^{\mathrm{AIR,ELECTRIC}} \cdot x_{i,j,t}^{electric}\big)\Big\} \quad (52)$$

where the coefficients $\mathrm{r}_{i,j,t}^{\mathrm{AIR}}$ are the premature mortalities for non-white populations per unit activity.

$$\text{Equity objective 3b:} \quad \mathrm{Min}\Big\{\textstyle\sum_{i,j\in jrace,t}\big(\mathrm{a}_{i,j,t}^{\mathrm{AIR,PREPRODUCTION}} \cdot x_{i,t}^{spud} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,PRODUCTION}} \cdot x_{i,t}^{producing} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,INDUSTRIAL}} \cdot x_{i,t}^{industrial} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,COMMERCIAL}} \cdot x_{i,t}^{commercial} + \mathrm{a}_{s,t}^{\mathrm{AIR,RESIDENTIAL}} \cdot x_{i,j,t}^{residential} + \mathrm{a}_{i,j,t}^{\mathrm{AIR,ELECTRIC}} \cdot x_{i,j,t}^{electric}\big)\Big\} \quad (53)$$

where the subset $jrace$ is the 10% of receptor counties with the highest proportion of their populations that is non-white; we define the subset based on county-level non-white populations as reported by the US Census Bureau for 2016.

4. Minimize impacts on rural populations

We formulate a single rural objective: minimize air quality impacts on counties with the highest rural populations (4a).

$$\text{Equity objective 4a:} \quad \text{Min}\Big\{\textstyle\sum_{i,j\in jrural,t}\Big(\mathrm{a}_{i,j,t}^{\text{AIR,PREPRODUCTION}} \cdot x_{i,t}^{spud} + \mathrm{a}_{i,j,t}^{\text{AIR,PRODUCTION}} \cdot x_{i,t}^{producing} + \mathrm{a}_{i,j,t}^{\text{AIR,INDUSTRIAL}} \cdot x_{i,t}^{industrial} + \mathrm{a}_{i,j,t}^{\text{AIR,COMMERCIAL}} \cdot x_{i,t}^{commercial} + \mathrm{a}_{s,t}^{\text{AIR,RESIDENTIAL}} \cdot x_{i,j,t}^{residential} + \mathrm{a}_{i,j,t}^{\text{AIR,ELECTRIC}} \cdot x_{i,j,t}^{electric}\Big)\Big\} \tag{54}$$

where the subset $jrural$ is the third of receptor counties with highest rural population; we define the subset based on county-level rurality rates as reported by the US Census Bureau for 2016.

5. Minimize impacts on populations with existing environmental burden

We formulate a single environmental burden objective: minimize air quality impacts on counties which are classified as nonattainment under the National Ambient Air Quality Standards (NAAQS) (5a).

$$\text{Equity objective 5a:} \quad \text{Min}\Big\{\textstyle\sum_{i,j\in jnonattainment,t}\Big(\mathrm{a}_{i,j,t}^{\text{AIR,PREPRODUCTION}} \cdot x_{i,t}^{spud} + \mathrm{a}_{i,j,t}^{\text{AIR,PRODUCTION}} \cdot x_{i,t}^{producing} + \mathrm{a}_{i,j,t}^{\text{AIR,INDUSTRIAL}} \cdot x_{i,t}^{industrial} + \mathrm{a}_{i,j,t}^{\text{AIR,COMMERCIAL}} \cdot x_{i,t}^{commercial} + \mathrm{a}_{s,t}^{\text{AIR,RESIDENTIAL}} \cdot x_{i,j,t}^{residential} + \mathrm{a}_{i,j,t}^{\text{AIR,ELECTRIC}} \cdot x_{i,j,t}^{electric}\Big)\Big\} \tag{55}$$

where the subset $jnonattainment$ is the receptor counties that were classified as nonattainment in 2017.

6. Minimize spatial inequity

We formulate three spatial equity objectives: minimize air quality spatial inequity (6a), employment spatial inequity (6b), and combined air-employment spatial inequity (6c).

$$\text{Equity objective 6a:} \quad \text{Min}\left\{\frac{\sum_{ja,jb}\left|\mathrm{Z}_{ja}^{\mathrm{AIR}}-\mathrm{Z}_{jb}^{\mathrm{AIR}}\right|}{2\mathrm{m}_{air}\cdot\mathrm{Z}^{\mathrm{AIR}}}\right\} \tag{56}$$

where ja and jb are indices that are aliases for receptor counties j, and m_{air} is the number of receptor counties.

$$\text{Equity objective 6b:} \quad \text{Min}\left\{\frac{\sum_{ia,ib}\left|\mathrm{Z}_{ia}^{\mathrm{EMPLOY}}-\mathrm{Z}_{ib}^{\mathrm{EMPLOY}}\right|}{2\mathrm{m}_{employ}\cdot\mathrm{Z}^{\mathrm{EMPLOY}}}\right\} \tag{57}$$

where ia and ib are indices that are aliases for source counties i, and m_{employ} is the number of source counties.

$$\text{Equity objective 6c:} \quad \text{Min}\left\{\frac{\sum_{ja,jb}\left|\mathrm{Z}_{ja}^{\mathrm{AIREMPLOY}}-\mathrm{Z}_{jb}^{\mathrm{AIREMPLOY}}\right|}{2\mathrm{m}_{airemploy}\cdot\mathrm{Z}^{\mathrm{AIREMPLOY}}}\right\} \tag{58}$$

where $\mathrm{Z}^{\mathrm{AIREMPLOY}} = \mathrm{Z}^{\mathrm{EMPLOY}} - \mathrm{Z}^{\mathrm{AIR}}$, and $\mathrm{Z}^{\mathrm{AIR}}$ is a modification of the previously described impact objective function, where we convert premature mortalities to life-years lost.

These objective functions are based on the Gini index, an aggregate measure of equity across the system that compares each equity unit (i.e., county) to all other equity units. We reformulate the objectives into sets of linear constraints, which transforms the optimization into a mixed integer linear program. Here, we provide an example for employment spatial equity.

$$\mathrm{Z}_{ia}^{\mathrm{EMPLOY}} - \mathrm{Z}_{ib}^{\mathrm{EMPLOY}} = \sigma_{1,ia,ib}^{\mathrm{EMPLOY}} - \sigma_{2,ia,ib}^{\mathrm{EMPLOY}} \quad \forall\, ia = 1, \dots, \mathcal{J};\; ib = 1, \dots, \mathcal{J} \tag{59}$$

$$\sigma_{1,ia,ib}^{\mathrm{EMPLOY}} \le \mathcal{M} \cdot \phi_{ia,ib}^{\mathrm{EMPLOY}} \quad \forall\, ia = 1, \dots, \mathcal{J};\; ib = 1, \dots, \mathcal{J} \tag{60}$$

$$\sigma_{2,ia,ib}^{\mathrm{EMPLOY}} \le \mathcal{M} \cdot (1 - \phi_{ia,ib}^{\mathrm{EMPLOY}}) \quad \forall\, ia = 1, \dots, \mathcal{J};\; ib = 1, \dots, \mathcal{J} \tag{61}$$

$$\sum_{ia,ib}(\sigma_{1,ia,ib}^{\mathrm{EMPLOY}} - \sigma_{2,ia,ib}^{\mathrm{EMPLOY}}) \le \mathrm{G}^{\mathrm{MAX}} \cdot 2\mathrm{m}_{employ} \cdot \mathrm{Z}^{\mathrm{EMPLOY}} \tag{62}$$

where $\sigma_{1,ia,ib}^{\text{EMPLOY}}$ and $\sigma_{2,ia,ib}^{\text{EMPLOY}}$ are slack variables, $\phi_{ia,ib}^{\text{EMPLOY}}$ are binary variables, and $\mathcal{M}$ is a very large number.

6.2.5 Multiobjective functions

Multiobjective programming facilitates comparisons between conflicting objectives. We evaluate pairwise tradeoffs between impact and equity objectives in terms of their physical units (i.e., premature mortalities, Kelvin-years, job-years, etc.) using the ε-constraint method.[130] The ε-constraint method is based on formulating an auxiliary model in which a single objective is optimized subject to a secondary objective which is reformulated as a constraint. This constraint imposes an upper (lower) limit ε on the value of the secondary maximization (minimization) objective and is iteratively solved to generate the Pareto set. An example of the pairwise tradeoff between cumulative employment impacts and rural air quality impacts can be expressed as:

$$\text{Max}\, \text{Z}^{\text{IMPACT}} \tag{63}$$

$$s.t.\;\; \text{Z}^{\text{EQUITY}} \leq \varepsilon_k \tag{64}$$

$$with\; \varepsilon_k = \varepsilon_1, \ldots, \varepsilon_n \text{ and } E_{LB} \leq \varepsilon_k \leq E_{LB} \tag{65}$$

Z^{IMPACT} and Z^{EQUITY} represent generic impact and equity objective functions, respectively. In this example, other impact and equity objectives are treated as projections on the Pareto set based on the given impact and equity objective pair. The extreme points (E_{LB}, E_{LB}) are derived based on solving the single objective ε-constraint problems.

6.3 Interpreting equity rules for energy systems planning

6.3.1 Minimize impacts on poor populations

The poorest counties in Appalachia are primarily concentrated in West Virginia, southern Ohio, and Kentucky, which largely aligns with the location of historical extractive energy booms (refer to Figure 20). For all equity objectives in which air quality impacts on poor populations are minimized, the model selects a natural gas system pathway in which production is spatially dispersed with somewhat greater siting within northern Pennsylvania or otherwise away from poor populations, as shown in Figure 22. However, there is a noticeable difference in the spatial distribution of production and impacts with respect to whether an equity objective targets the aggregate population or counties with the highest rates of low income populations or populations below the poverty line (Figure 23). We additionally find that end use demand, including electric generation, is largely fixed and unresponsive to the equity objectives.

Figure 21 depicts the Pareto frontier representing the subset of solutions which are non-inferior – i.e., no other feasible solution will yield an improvement in one objective without degradation in at least one other objective. The marginal rate of substitution (MRS), which is the slope along the Pareto frontier, can be interpreted as the rate at which some amount of cumulative impact can be exchanged for equity. In climate and air quality optimization regimes, there are minor conflicts (if any) between impact and equity objectives, given that both types of objectives iterate to solutions which tend towards minimal natural gas activity; whereas, in an employment regime, there is a non-trivial, decreasing MRS between employment and equity objectives.

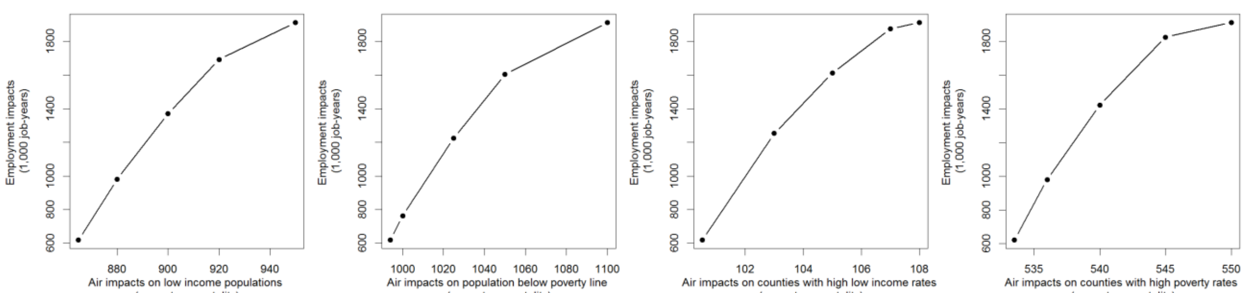

Figure 21. Pareto frontiers of pairwise equity and employment impact tradeoffs for equity objectives in which air quality impacts are minimized for poor populations.

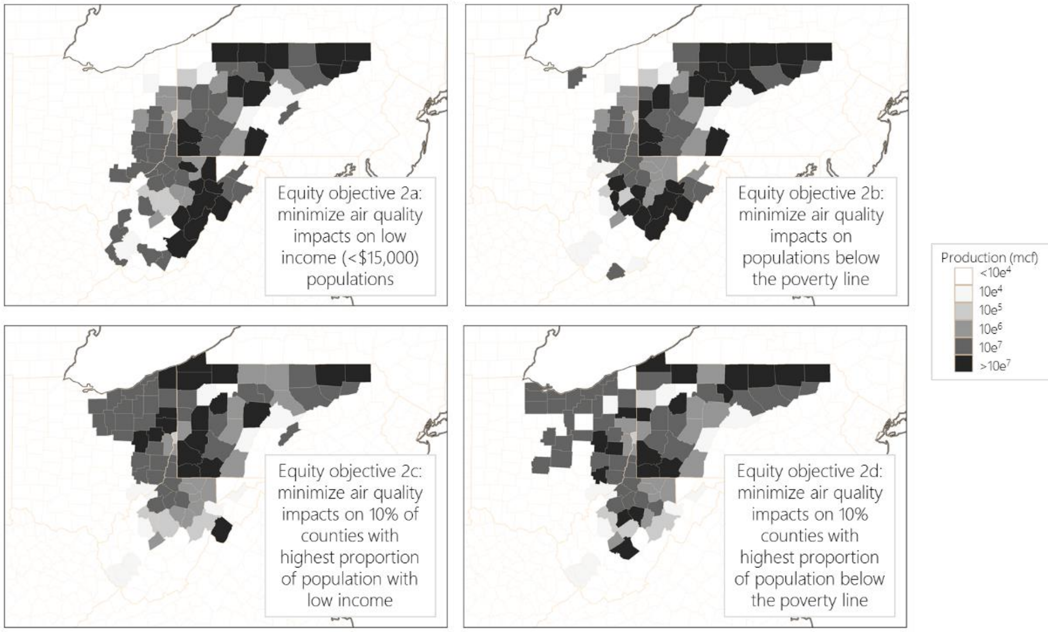

Figure 22. Spatial distribution of 2017 to 2030 production for equity objectives in which air quality impacts are minimized for poor populations. Production values are based on the employment optimal solution along the relevant employment-equity Pareto frontier.

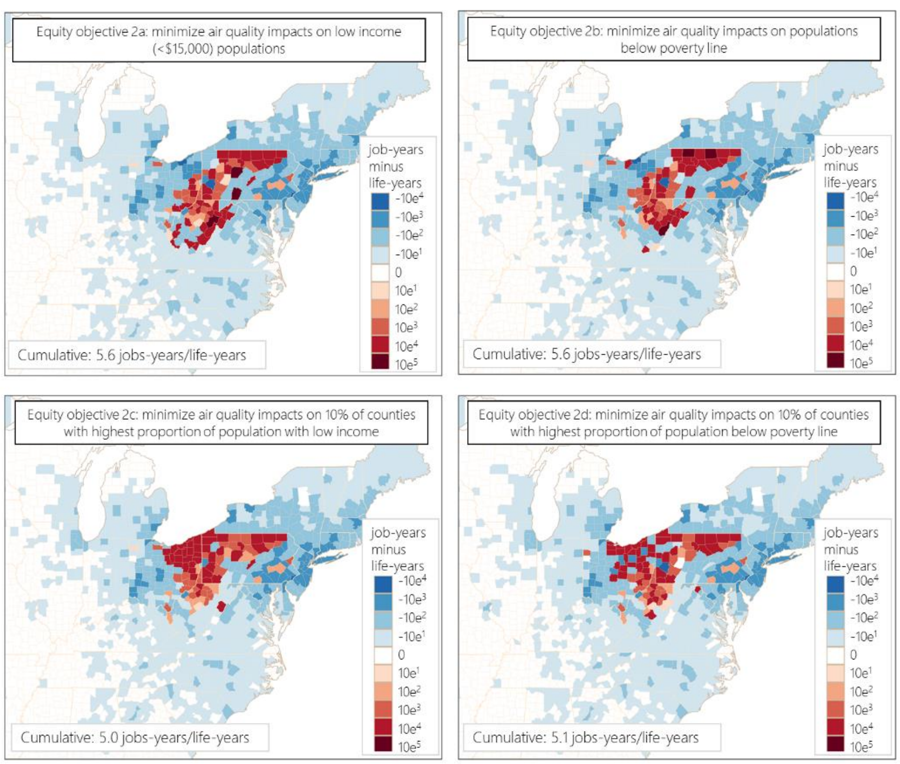

Figure 23. Spatial distribution of 2017 to 2030 impact for equity objectives in which air quality impacts are minimized for poor populations. Impact values are based on the employment optimal solution along the relevant employment-equity Pareto frontier.

6.3.2 Minimize impacts on minority populations

Counties with high non-white populations are primarily concentrated in urban and coastal states, outside of the Appalachian basin (refer to Figure 20). We find that there is a noticeable difference in the spatial distribution of production and impacts with respect to whether an equity objective targets the aggregate non-white population or counties with the highest rates of non-white populations, as shown in Figure 25 and Figure 26. In climate and air quality optimization regimes, there are minor conflicts (if any) between impact and equity objectives, given that both types of objectives iterate to solutions which tend towards minimal natural gas activity; whereas, in an employment regime, there are non-trivial conflicts between employment and equity objectives, as shown in Figure 24.

While there are intersectionalities between poor and non-white subpopulations, the associated equity objectives do not iterate towards similar development pathways, given the differing general spatial distribution of these subpopulations. There is not an observable systematic difference in spatial production and impact distributions between equity objectives minimizing impacts to poor and non-white subpopulations.

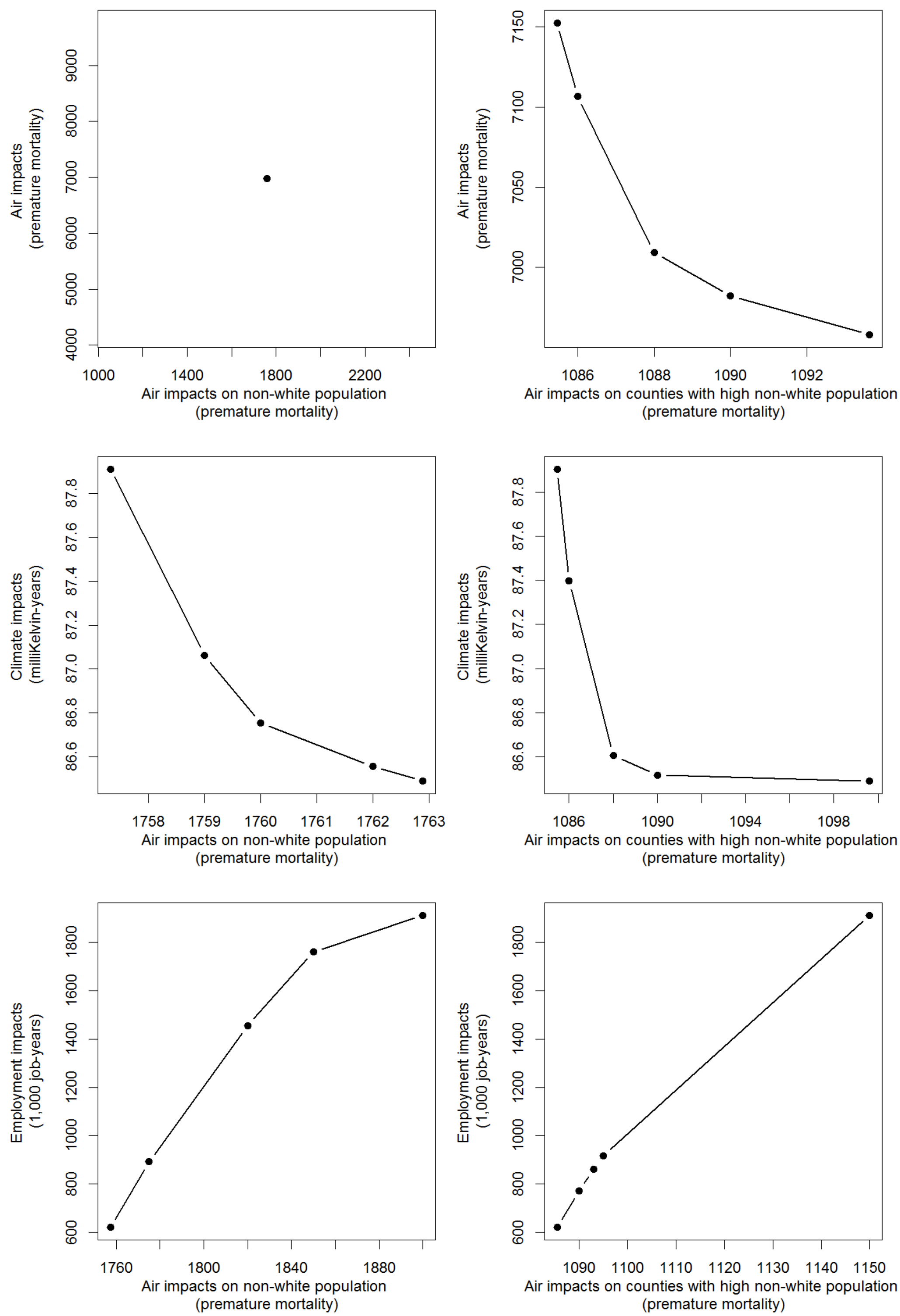

Figure 24. Pareto frontiers of pairwise equity and impact tradeoffs for equity objectives in which air quality impacts are minimized for non-white populations.

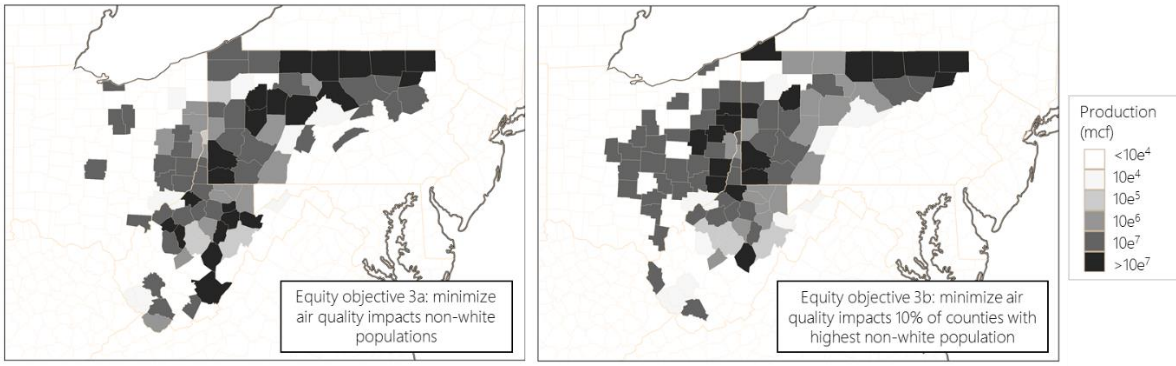

Figure 25. Spatial distribution of 2017 to 2030 production for equity objectives in which air quality impacts are minimized for non-white populations. Production values are based on the employment optimal solution along the relevant employment-equity Pareto frontier.

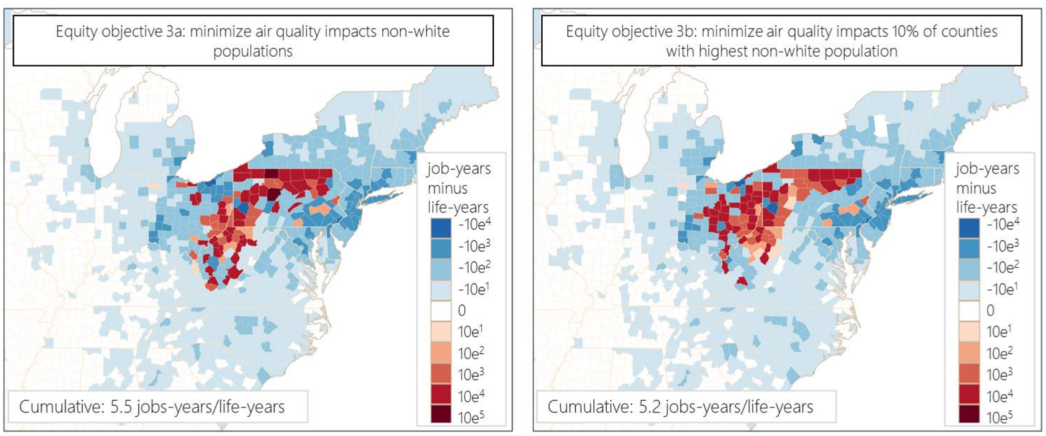

Figure 26. Spatial distribution of 2017 to 2030 impacts and natural gas activity for equity objectives in which air quality impacts are minimized for non-white populations. Impact values are based on the employment optimal solution along the relevant employment-equity Pareto frontier.

6.3.3 Minimize impacts on rural populations

We assign rural counties in Appalachia as a proxy for communities with historical extraction and concerns. We observe that the model selects a natural gas system pathway with more concentrated production and impacts in Pennsylvania and western Ohio, as shown in Figure 28 and Figure 29.

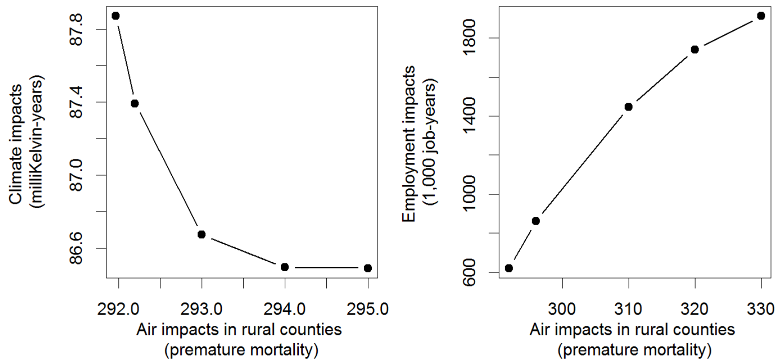

Figure 27. Pareto frontiers of pairwise equity and impact tradeoffs for equity objectives in which air quality impacts are minimized for rural populations.

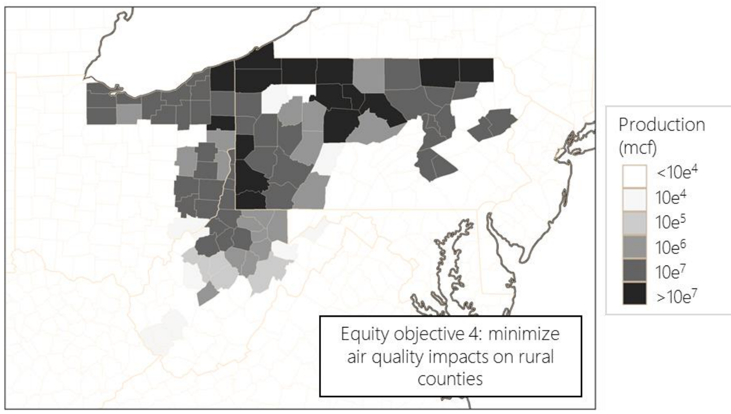

Figure 28. Spatial distribution of 2017 to 2030 production for equity objectives in which air quality impacts are minimized for rural populations. Production values are based on the employment optimal solution along the relevant employment-equity Pareto frontier.

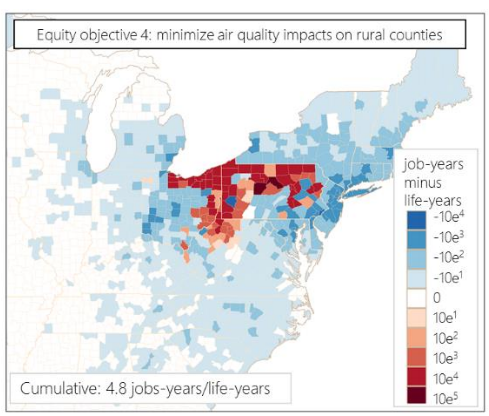

Figure 29. Spatial distribution of 2017 to 2030 impacts for equity objectives in which air quality impacts are minimized for rural populations. Impact values are based on the employment optimal solution along the relevant employment-equity Pareto frontier.

6.3.4 Minimize impacts on populations with existing environmental burden

Non-attainment counties are primarily more urban (refer to Figure 20). Thus, in minimizing this objective, the model selects a natural gas system pathway in which production and net positive impacts are more concentrated in rural areas of West Virginia, as depicted in Figure 31 and Figure 32. This is a systematically different pattern of development than equity objectives related to poor, minority, and rural populations. In climate and air quality optimization regimes, there are minor conflicts (if any) between impact and equity objectives, given that both types of objectives iterate to solutions which tend towards minimal natural gas activity; whereas, in an employment regime, there is a non-trivial tradeoff between employment and equity objectives, as shown in Figure 30.

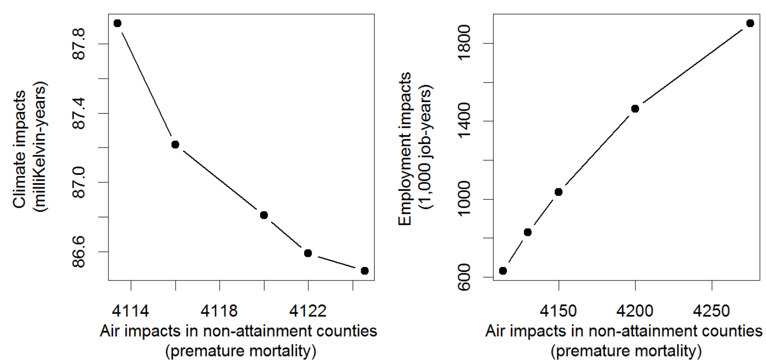

Figure 30. Pareto frontiers of pairwise equity and impact tradeoffs for equity objectives in which air quality impacts are minimized for non-attainment counties.

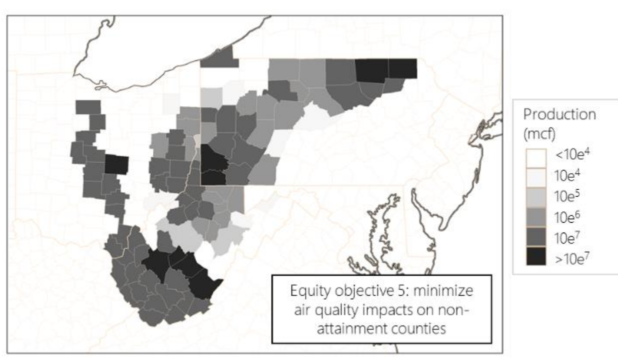

Figure 31. Spatial distribution of 2017 to 2030 production for equity objectives in which air quality impacts are minimized for non-attainment counties. Production values are based on the employment optimal solution along the relevant employment-equity Pareto frontier.

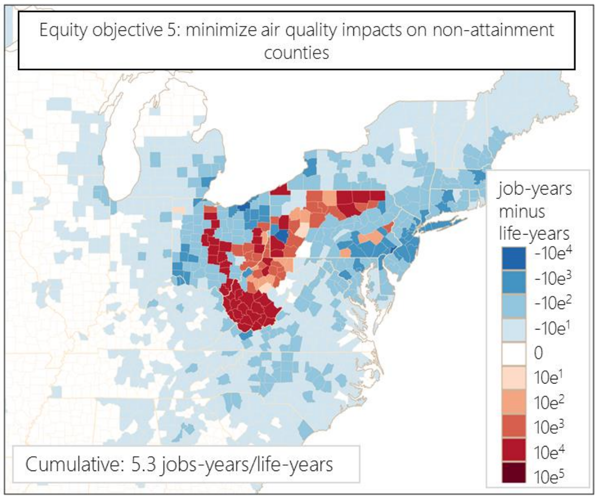

Figure 32. Spatial distribution of 2017 to 2030 impacts for equity objectives in which air quality impacts are minimized for non-attainment counties. Impact values are based on the employment optimal solution along the relevant employment-equity Pareto frontier.

6.4 The cost of equity

As described in previous sections, varying degrees of conflict and agreement may exist between equity and cumulative impact objectives. In air and climate minimization regimes, equity objectives limiting impacts on subpopulations based on demographics and existing environmental burdens are largely in agreement, given that both impact and equity objectives tend to limit production and other activity. Whereas, in an employment maximization regime, which tends to select a development pathway where there is greater natural gas activity, there are fundamental conflicts with all equity objectives. To further quantify these equity and impact conflicts, we estimate the average MRS by linearly regressing impacts on equity based on the Pareto optimal solutions; this approach is limited in that it does not capture increasing or decreasing MRS which are apparent in many of the Pareto frontiers. We then monetized the average cost of equity using the value of a statistical life or average income per capita, as shown in Figure 33. The cost of equity ranges from \$11 to \$580 million per avoided premature mortality for a given subpopulation.

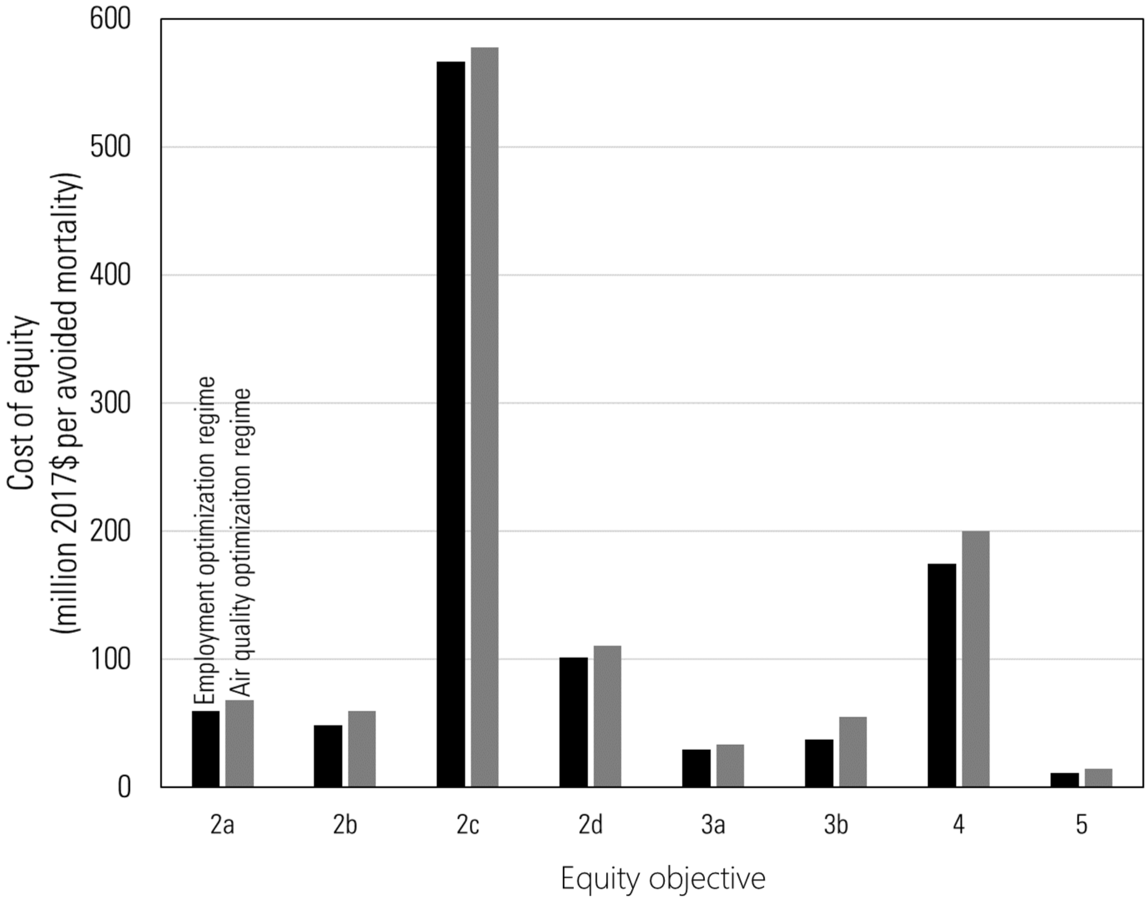

Figure 33. Cost of equity in units of million US$2017 per avoided mortality for given subpopulation. The cost is based on multiplying the mean MRS with either the value of a statistical life ($8.5M) or the annual, county-level income per capita.

6.5 Tradeoffs between equity objectives

There are inherent conflicts between equity objectives, given different spatial and temporal patterns of populations and impacts. Figure 34 presents a line graph, in which we normalize various equity metrics based on their best and worst values, to reveal the relative differences between equity objectives. We show the relationship between equity objectives, assuming metrics are derived from the employment optimal solution along the relevant employment-equity Pareto frontier. There are pronounced differences between the rural equity objective and other objectives, especially the non-attainment county objective which tends to select a development pathway that avoids urban impacts. We also note that minimizing impacts in poor and non-white communities differ. Even within equity categories, objectives do not follow the same or similar dominance structures (e.g., minimizing on air quality premature mortality on populations below the poverty line is not equivalent to minimizing on the 10% of counties with the highest poverty rates).

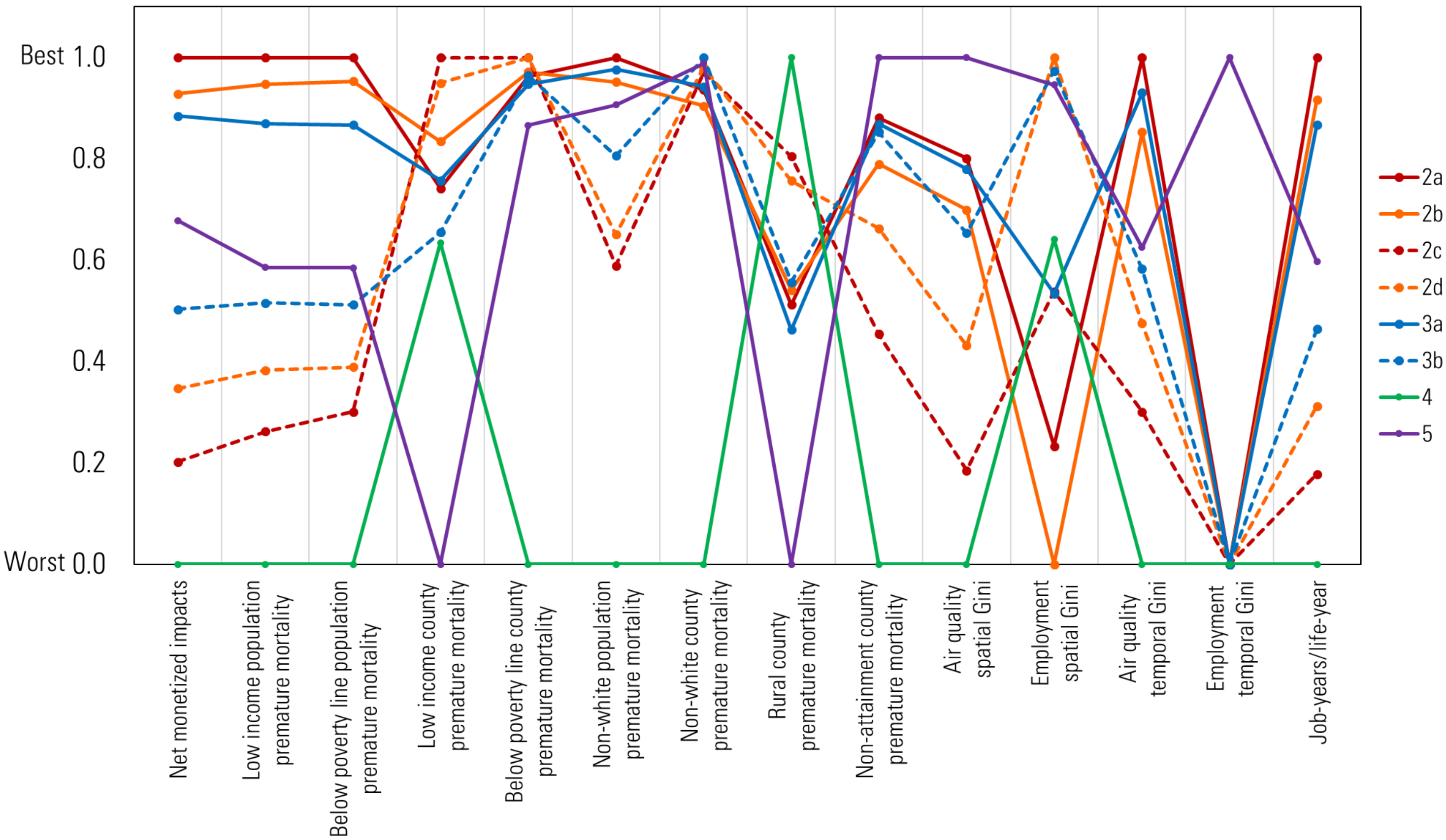

Figure 34. Comparison of equity objectives.

6.6 Bending equity rules through low-carbon energy technologies

In ongoing work, we explore and demonstrate how air pollutant and greenhouse gas emissions abatement, as well as integration of renewables to displace natural gas, can effectively bend the equity rules. Preliminary results suggest that despite inherent equity and impact conflicts in a fossil-fuel dominated energy system, low-carbon technologies have the potential to minimize inequities, while also reducing cumulative impacts.

6.7 Implications for policy and decision making

To compensate for, mitigate, or otherwise address societal inequities between populations or subpopulations delineated on the basis on geographical, temporal, or demographic attributes

This study provides an approach for developing future energy system pathways that optimize for different dimensions of equity. We operationalize a small subset of the seemingly limitless interpretations and combinations of impact and equity objectives as they relate to energy systems analysis, planning, and policy. We only generate pairwise comparisons to allow for comprehension of the most basic dynamics and tradeoffs, whereas the actual problem space has more than two dimensions and is much more complex. The framework presented is descriptive in nature, and there is no way to pick a particular Pareto optimal solution without appealing to ethical judgments. Within a decision-making process, the tradeoffs and energy system pathways generated using this approach can be presented to decision makers and stakeholders, and equity and impact objectives can be iteratively refined. This process may facilitate understanding that there are equity vs. impact and equity vs. equity tradeoffs, and may enable setting of goals and elucidation of preference hierarchies between different impact and equity criteria. Policies may then be designed which instill these impact and equity preference structures.

While it is possible to design public policy that explicitly addresses inequities, given preferences and technically feasible solutions, such policies are outside the norm of U.S. domestic environmental policy, which effectively adopts a Kaldor-Hicks efficiency criterion and treats equity as subsidiary or ignores it entirely. Moreover, the energy system is dynamic and all equity rules explored in this study have a spatial and/or temporal dimension, requiring adaptable policy that matches this heterogeneity, such as location-specific and temporally evolving siting, production, and consumption standards. This suggests that equity policies are likely more complex to implement and enforce than traditional policies. Furthermore, given that equity policies are non-standard, are derived based on potentially divergent ethical judgments, and have the potential to reverse the historically observed Not-In-My-Backyard phenomenon of siting decisions, such policies may be less politically tractable. Practical limitations abound, but the historical and current standards of public decision making and energy system planning have demonstratively contributed to observed and entrenched inequities.

6.8 Conclusions

The entangled strings of equity pull in many directions, which requires both delicately and pointedly addressing as a society, while also balancing cumulative societal impacts. Increasing real and perceived inequities on all sides of the political and cultural spectrum in the U.S. are seemingly driving a societal moment demanding reconciliation and potentially contributing to further ideological entrenchment that may have cascading effects in a transition to a low-carbon energy system. Many of the inequities within our society are rooted in slavery of the 19th and 20th centuries and the treatment of immigrant populations, which have percolated into racial and ethnic disparities in many facets of today's society including environmental injustices. Furthermore, our nation as a whole prospered from the exploitation of non-renewable energy resources, but not without placing a disparate burden on certain subpopulations, such as the rural poor of Appalachia that have experienced multiple boom-and-bust cycles since the Industrial Revolution. Piling onto the centuries-long inequities that persist in some form at present are the impending existential and intergenerational threats of climate change, which will be borne out on the order of millennia and temporally eclipse the historical inequities of the U.S.

7 Conclusions

Transforming the energy system entails balancing multiple and often conflicting societal objectives. This thesis presents new quantitative modeling approaches for energy and environmental systems planning and policy evaluation, with an emphasis on system heterogeneity, cumulative air, climate, and employment impacts, and temporal, spatial, and distributional equity. The following summarizes key objectives, findings, and policy implications of this thesis. In addition, we present a proposal for future work.

7.1 Key findings and policy implications

The chapters present a range of prescriptive and descriptive policy implications, applied and abstract insight derived through both retrospective and prospective lenses, and empirical and normative approaches. All chapters are heavily data-driven, adopt a systems-level quantitative modeling approach, and seek to address limitations of existing evaluative tools for environmental and energy policy and planning.

Chapter 2 adopts a traditional private and social economic efficiency maximization approach, coupled with methane emissions and abatement cost simulations reflecting system heterogeneity, to evaluate and design system-wide and superemitter policies related to methane abatement of the U.S. transmission and storage system. We demonstrate that there are high societal benefits from abatement policies, and minimal (if any) private costs given the existing suite of abatement options, which is counter to the belief held by some that methane abatement and detection are necessarily in conflict with private interests. We also find that superemitter policies, which target the highest emitting facilities, may reduce the private cost burden, while system-wide policies may be warranted given the non-trivial societal benefit of abating low-emitting sources. This work contributes to our understanding of the prospects and limitations of evaluating and designing policies that account for system heterogeneity. Our analysis also contributes to honest discourse regarding the merits of methane policies, especially in light of proposed rollbacks of federal methane rules.

Chapters 3 and 5 aim to develop and demonstrate data-driven metrics and approaches for characterizing systems-level cumulative impacts and equity of current energy systems. In Chapter 3, we assess the spatially-and temporally-resolved air, climate, and employment impacts from extraction to end use and over the life of natural gas plays in the Appalachian basin from 2004 to 2016. We show that short-lived air quality iand employment impacts track with the boom-and-bust cycle, while climate impacts persist for generations well beyond the period of natural gas activity. We also find that employment effects are spatially concentrated in rural areas with thin labor markets where development is occurring, and more than half of cumulative premature mortality is within source emissions states. We show that most premature mortality is associated with end uses, while upstream and midstream segments also account for a substantial portion of impacts. With respect to climate change impacts, the magnitude of methane emissions across the supply chain produces temperature impacts nearly equivalent to that of carbon dioxide over a 30-year time horizon, but over longer integration periods, the warming impact from carbon dioxide dominates. We estimate that a tax on production of $2 per thousand cubic foot (+172%/-76%) would compensate for cumulative climate and air quality externalities across the supply chain. While we do not present the production tax as a policy recommendation, it does contextualize impacts relative to natural gas prices and production costs, which stimulates questions regarding whether firm decisions would change if they internalized environmental damages.

In Chapter 5, we develop and demonstrate an approach for evaluating the equity state of an energy system. We apply variants of standard methods and present new methods and metrics to quantify spatial, temporal, and distributional equity, leveraging impact estimates of the shale gas boom in the Appalachian basin from Chapter 3. We find that there are high temporal and spatial inequities with respect to cumulative air and employment impacts, and that spatial inequities are constant over time reflecting largely fixed infrastructure and consumption patterns. We also present indicators of temporal climate inequities, estimating that long-term global temperature impacts are 100 times that of near-term impacts. With respect to distributional equity of air quality impacts, we do not observe a disparity in mortality rates across subpopulations on the basis of income and poverty; however, there is a trend of increasing income corresponding to decreasing damages, which demonstrates the higher health burden of lower income communities. With respect to distributional equity of labor markets, we find statistically significant declines in the income disparity and poverty rates in producing counties. Pairwise comparisons of impacts reveal that changes in air and climate impacts are sensitive to changes in employment impacts. Chapter 5 is intended to contribute to the policy evaluation toolset and to provide descriptive policy insight and salience, revealing the inequities that exist within the natural gas system. These system analytics can be used to further assess changes in the equity state of an energy system.

Chapters 4 and 6 develop and apply a multiobjective energy systems optimization model, which is informed and parameterized based on the data-driven quantitative modeling of Chapters 3 and 5. We formulate cumulative impact and equity objectives that reflect additional societal values, beyond the traditional economic efficiency maximization criterion. The modeling architecture is distinct from many existing energy systems optimization models because of the unique cumulative impact and equity objectives, spatial resolution, long time horizons, air and climate impact (rather than emission and monetary) measures leveraging reduced complexity models, and parameterization and structure reflecting a real (rather than illustrative) natural gas system. The generated modeling scenarios represent alternative energy system futures, based on our understanding of the historical evolution of the system. The scenarios also reflect the conflicts and agreements between cumulative air, climate, and employment impacts and spatial, temporal, and distributional equity objectives. Preliminary results also demonstrate that conflicts may be reduced or reversed as a collection of low-carbon technologies are introduced into a fossil fuel-dominated energy system. These chapters are intended to contribute to the policy process and evaluative toolset, rather than prescribe policy. Given that we are modeling complex and highly uncertain futures, the modeling scenarios provide more abstract policy insight. However, given that the modeling inputs and architecture represent details of the actual natural gas system in Appalachia, it is possible to generate prescriptive policy reflective of cumulative impact and equity preferences with iterative feedback from stakeholders and decision makers.

7.2 Future work

This section proposes refinements and extensions to the studies in this thesis, as well as, general areas of future work.

With respect to the multiobjective cumulative impact and equity optimization models presented in Chapters 4 and 6, additional scenarios are in progress. We are developing modules that allow for the abatement of air pollutant and GHG emissions and displacement of future natural gas electricity demand with renewable generation. In addition, we intend to perform sensitivity analyses to explore the uncertainty and variability associated with the value of a statistical life, emission factors, dose-response relationships, representative climate pathways, absolute global temperature potential, social cost of methane and carbon, marginal labor, well productivity, and end use demand.

There are several refinements and extensions to the multiobjective cumulative impact model which we do not intend to include in the studies presented in this thesis. Other low-carbon interventions, such as energy efficiency, demand reduction, and electrification of commercial and industrial end use, may be modeled. In addition, cumulative impact objectives related to ecological, land, and water impacts and private costs may be specified. The modeling structure can be broadened to include midstream natural gas activity and used for evaluating the tradeoffs of developing a petrochemical industry in southwest Pennsylvania. There are additional uncertainties and dynamics which may be incorporated, including changing atmospheric sensitivities, correlation between air pollutant and GHG emissions, spatial variability in emission factors, economies/diseconomies on marginal labor, spatial employment spillovers, correlation between segments of the supply chain, declining well productivity as the play ages, technological advancements increasing well productivity, and variability in well productivity. Some of these dynamics are structurally difficult to include in this optimization framework or not well understood, which may require auxiliary analyses. While we focus on natural gas activity within the Appalachian basin, the model is specified such that it is readily adaptable to the broader domestic energy system and can be used to explore environmental and labor market tradeoffs associated with policy proposals and large-scale green infrastructure investment, such as proposals for a Green New Deal.

The theoretical and applied field of incorporating systems-level equity into energy systems optimization and planning models is largely undeveloped, with the key exceptions of energy access for low-income households and the developing world. Overall, we hope that the foundations of an approach for energy equity systems analysis presented in Chapters 5 and 6, spur other researchers to consider, if not contribute to, energy equity systems analysis. There are seemingly limitless equity concepts that can be operationalized in quantitative terms and applied to other contexts. For example, in the case of large-scale green infrastructure investment and siting, variants of spatial equity objectives may be formulated which enforce a more equitable distribution of both benefits and costs across counties or Congressional districts. In the case of publicly-funded infrastructure, these types of equity rules may facilitate legislative and stakeholder support for funding and siting of certain types of infrastructure that are high cost or otherwise controversial, such as transmission lines.

Finally, much of the work in this thesis is currently framed for policy researchers and modelers. Some of the main findings of this thesis can be translated and distilled for policy makers, elucidating the historical impacts of natural gas development in the Appalachian basin and demonstrating how an energy system transition and public and private investment in low-carbon energy technologies may relieve conflicts between environmental and employment impacts and equity.

References

1. Field, C. B. *et al. 2014: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. (2009). doi:10.1017/CBO9781107415324.004

2. Stocker, T. F. *et al. IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. (2014).

3. Edenhofer, O. *et al. IPCC, 2014: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*.

4. Williams, J. H., Jones, R., Kwok, G. & Haley, B. Deep Decarbonization in the Northeastern United States and Expanded Coordination with Hydro-Québec. (2018).

5. Haley, B. *et al.* Deep Decarbonization. (2015).

6. The White House. United States Mid-Century Strategy for Deep Decarbonization. 1–111 (2016).

7. Socolow, R. & Pacala, S. Stabilization Wedges: Solving the Climate Problem for the Next Half-Century with Technologies Available Today. *Science (80-.).* **305,** 968–972 (2004).

8. Geels, F. W., Sovacool, B. K., Schwanen, T. & Sorrell, S. Accelerating innovation is as important as climate policy. *Science (80-.).* **357,** 1242–1244 (2017).

9. Bataille, C. *et al.* A review of technology and policy deep decarbonization pathway options for making energy-intensive industry production consistent with the Paris Agreement. *J. Clean. Prod.* **187,** 960–973 (2018).

10. Fann, N., Fulcher, C. M. & Baker, K. The recent and future health burden of air pollution apportioned across u.s. sectors. *Environ. Sci. Technol.* **47,** 3580–3589 (2013).

11. Jochem, E. & Madlener, R. Workshop on the Benefits of Climate Policy : Improving Information for Policy Makers The Forgotten Benefits of Climate Change Mitigation : Innovation , Technological Leapfrogging , Employment , and Sustainable Development. *Energy Policy* (2003).

12. Tol, R. S. J., Downing, T. E., Kuik, O. J. & Smith, J. B. Distributional aspects of climate change impacts. *Glob. Environ. Chang.* **14,** 259–272 (2004).

13. Shonkoff, S. B., Morello-Frosch, R., Pastor, M. & Sadd, J. The climate gap: Environmental health

and equity implications of climate change and mitigation policies in California-a review of the literature. *Clim. Change* **109,** 485–503 (2011).

14. Loulou, R. & Labriet, M. ETSAP-TIAM: The TIMES integrated assessment model Part I: Model structure. *Comput. Manag. Sci.* **5,** 7–40 (2008).

15. US Energy Information Administration. *The National Energy Modeling System: An Overview*. **0581,** (2009).

16. Moreno-Cruz, J. B., Ricke, K. L. & Keith, D. W. A simple model to account for regional inequalities in the effectiveness of solar radiation management. *Clim. Change* **110,** 649–668 (2012).

17. Cole, W. *et al.* Considering the Role of Natural Gas in the Deep Decarbonization of the U . S . Electricity Sector Considering the Role of Natural Gas in the Deep Decarbonization of the U . S . Electricity Sector Number 2. (2016).

18. Perkins, D. W. & Carey, M. P. Cost - Benefit Analysis and Financial Regulator Rulemaking. (2017).

19. Gowdy, J. M. The Revolution in Welfare Economics and Its Implications for Environmental Valuation and Policy. *Land Econ.* **80,** 239 (2004).

20. Arrow, K. *et al.* Are We Consuming Too Much ? Is humanity ' s use of Earth ' s resources endangering the economic possibilities. *J. Econ. Perspect.* **18,** 147–172 (2004).

21. Lave, L. Benefict-cost analysis: Do the benefits exceed the costs? in *Risks, Costs, and Lives Saved* (ed. Hahn, R. W.) (1996).

22. Arrow, K. J. *et al.* Is There a Role for Benefit-Cost Analysis in Environmental, Health & Safety Regulation 1996.pdf. *Science (80-.).* **272,** 221–222 (1996).

23. Zeckhauser, R. & Schaefer, E. Public Policy & Normative Economic Theory. in *The Study of Policy Formation* (eds. Bauer, R. & Gergen, K.) (1996).

24. Tol, R. S. J. Equitable cost-benefit analysis of climate change. *Nota di Lav.* **41,** (1999).

25. Farrow, S. Incorporating Equity in Regulatory and Benefit-Cost Analysis Using Risk Based Preferences. 1–21 (2009).

26. Arrow, K. *et al.* Determining Benefits and Costs for Future Generations. *Science (80-.).* **341,** 349–351 (2013).

27. Heal, G. & Millner, A. Uncertainty and decision making in climate change economics. *Rev. Environ. Econ. Policy* **8,** 120–137 (2014).

28. Kenneth J. Arrow *et al. How Should Benefits and Costs Be Discounted in an Intergenerational Context? The Views of an Expert Panel*. *Resources for the Future* (1992). doi:10.2139/ssrn.2199511

29. U.S. Environmental Protection Agency. Technical Guidance for Assessing Environmental Justice in Regulatory Analysis. (2016).

30. U.S. Environmental Protection Agency. Toolkit for Assessing Potential Allegations of Environmental Injustice. (2004). doi:EPA 300-R-04-002

31. BP. *BP Statistical Review of World Energy, 67th Edition*. (2018).

32. Schivley, G., Azevedo, I. M. L. & Samaras, C. Assessing the evolution of power sector carbon intensity in the United States. *Environ. Res. Lett.* (2018). doi:https://doi.org/10.1088/1478-3975/aa9768

33. Jacoby, H. D., Sullivan, F. M. O. & Paltsev, S. Environmental Policy. (2011).

34. U.S. Energy Information Adminstration. *U.S. Crude Oil and Natural Gas Proved Reserves, Year-end 2016*. (2018).

35. Alvarez, R. A., Pacala, S. W., Winebrake, J. J., Chameides, W. L. & Hamburg, S. P. Greater focus needed on methane leakage from natural gas infrastructure. *Proc. Natl. Acad. Sci.* **109,** 6435–6440 (2012).

36. Brandt, A. R., Heath, G. A. & Cooley, D. Methane Leaks from Natural Gas Systems Follow Extreme Distributions. *Environ. Sci. Technol.* **50,** 12512–12520 (2016).

37. U.S. Environmental Protection Agency (EPA). *Oil and Natural Gas Sector Compressors*. (2014).

38. U.S. Environmental Protection Agency (EPA). *Oil and Natural Gas Sector Leaks Report for Oil and Natural Gas Sector Leaks*. (2014).

39. U.S. Environmental Protection Agency (EPA). *Oil and Natural Gas Sector Pneumatic Devices*. *EPA White Pages* (2014).

40. U.S. Environmental Protection Agency. Regulatory Impact Analysis of the Proposed Emission Standards for New and Modified Sources in the Oil and Natural Gas Sector. (2015).

41. U.S. Environmental Protection Agency (EPA). *Oil and Natural Gas Sector: Emission Standards*

for New, Reconstructed, and Modified Sources. 35824–35941 (2016).

42. Fullerton, D. Distributional Effects of Environmental and Energy Policy : An Introduction. (2009).

43. Brandt, A. *et al.* Methane Leaks from North American Natural Gas Systems. *Science (80-.).* **343,** 733–735 (2014).

44. U.S. Environmental Protection Agency. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2016*. (2018).

45. The White House. *Climate Action Plan Strategy to Reduce Methane Emissions*. (2014).

46. Office of the Press Secretary. FACT SHEET: Administration Takes Steps Forward on Climate Action Plan by Announcing Actions to Cut Methane Emissions. *The White House* 2–5 (2015).

47. U.S. Environmental Protection Agency (EPA). *Natural Gas STAR Methane Challenge Program : Proposal for Feedback*. (2015).

48. Subramanian, R. *et al.* Methane emissions from natural gas compressor stations in the transmission and storage sector: Measurements and comparisons with the EPA greenhouse gas reporting program protocol. *Environ. Sci. Technol.* **49,** 3252–3261 (2015).

49. Zimmerle, D. J. *et al.* Methane emissions from the natural gas transmission and storage system in the United States. *Environ. Sci. Technol.* **49,** 1–50 (2015).

50. Allen, D. T. *et al.* Measurements of methane emissions at natural gas production sites in the United States. *Proc. Natl. Acad. Sci.* **110,** 17768–17773 (2013).

51. Lamb, B. K. *et al.* Direct measurements show decreasing methane emissions from natural gas local distribution systems in the United States. *Environ. Sci. Technol.* **49,** 5161–9 (2015).

52. Mitchell, A. L. *et al.* Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results_supplemental information. *Environ. Sci. Technol.* **49,** 3219–3227 (2015).

53. Zavala-Araiza, D. *et al.* Toward a Functional Definition of Methane Super-Emitters: Application to Natural Gas Production Sites. *Environ. Sci. Technol.* **49,** 8167–8174 (2015).

54. Saunier, S. Quantifying Cost-effectiveness of Systematic Leak Detection and Repair Programs Using Infrared Cameras. *Carbon Limits* (2014).

55. ICF International. Economic Analysis of Methane Emission Reduction Opportunities in the Canadian Oil and Natural Gas Industries. 118 (2015). doi:10.1016/S0065-2113(08)60504-0

56. Zavala-Araiza, D. *et al*. Super-emitters in natural gas infrastructure are caused by abnormal process conditions. *Nat. Commun.* **8,** 14012 (2017).

57. U.S. Environmental Protection Agency. Directed Inspection and Maintenance at Compressor Stations. 1–11 (2003).

58. U.S. Environmental Protection Agency (EPA). Installing Vapor Recovery Units on Storage Tanks - Lessons learned from Natural Gas STAR Partners. (2006).

59. U.S. Environmental Protection Agency. Replacing Wet Seals with Dry Seals in Centrifugal Compressors - Lessons learned from Natural Gas STARS Partners. 1–8 (2006).

60. U.S. Environmental Protection Agency. Oil and Natural Gas Sector: Standards of Performance for Crude Oil and Natural Gas Production , Transmission , and Distribution. (2011).

61. Office of Management and Budget. Circular-A94 Guidelines and Discount Rates for Benefit-Cost Analysis of Federal Programs. (1992). doi:10.1007/s10796-016-9632-3

62. National Gas Machinery Laboratory, Clearstone Engineering & Environmetnal Innovative Solutions. Cost Effective Directed Inspection and Maintenance Control Opportunities at Five Gas Processing Plants and Upstream Gathering Compressor Stations and Well Sites. 1–50 (2006).

63. U.S. Energy Information Adminstration. *Annual Energy Outlook 2015 DOE/EIA-0383*. (2015).

64. Bureau of Economic Analysis. *National Economic Accounts: Current-Dollar and 'Real' Gross Domestic Product*. (2015).

65. U.S. Environmental Protection Agency (EPA). Interactive Convertor Tool. Available at: http://www3.epa.gov/cmop/resources/converter.html. (Accessed: 20th October 2016)

66. Environmental Defense Fund. How do EDF and ARPA-E projects on methane sensing relate to each other. (2014).

67. Environmental Defense Fund. Methane Detectors Challenge Request for Proposal. (2014).

68. Marten, A. L. & Newbold, S. C. Estimating the social cost of non-CO2GHG emissions: Methane and nitrous oxide. *Energy Policy* **51,** 957–972 (2012).

69. Marten, A. L., Kopits, E. A., Griffiths, C. W., Newbold, S. C. & Wolverton, A. Corrigendum to: Incremental CH_4 and N_2O mitigation benefits consistent with the U.S. Government's SC-CO_2 estimates. *Clim. Policy* **15,** 678–679 (2015).

70. California Environmental Protection Agency Air Resources Board. Aliso Canyon Methane Leak

Climate Impacts Mitigation Program. (2016).

71. U.S. Environmental Protection Agency (EPA). Technical Support Document: Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866. 1–21 (2010).

72. Smith, C. US pipeline operators sink revenue growth into expansion. *Oil Gas J.* (2013).

73. Small, M. J. *et al.* Risks and Risk Governance in Unconventional Shale Gas Development. (2014).

74. Vidic, R. D., Brantley, S. L., Vandenbossche, J. M., Yoxtheimer, D. & Abad, J. D. Impact of shale gas development on regional water quality. *Science (80-.).* **340,** (2013).

75. Olmstead, S. M., Muehlenbachs, L. a, Shih, J., Chu, Z. & Krupnick, A. J. Shale gas development impacts on surface water quality in Pennsylvania. *Proc. Natl Acad. Sci. USA* **110,** 4962–4967 (2013).

76. Jiang, M., Hendrickson, C. T. & Vanbriesen, J. M. Life Cycle Water Consumption and Wastewater Generation Impacts of a Marcellus Shale Gas Well a . Direct and indirect water use across life cycle stages of a Marcellus shale gas well.

77. Vengosh, A., Jackson, R. B., Warner, N., Darrah, T. H. & Kondash, A. A Critical Review of the Risks to Water Resources from Unconventional Shale Gas Development and Hydraulic Fracturing in the United States. (2014).

78. Litovitz, A., Curtright, A., Abramzon, S., Burger, N. & Samaras, C. Estimation of regional air-quality damages from Marcellus Shale natural gas extraction in Pennsylvania. *Environ. Res. Lett.* **8,** (2013).

79. Roohani, Y. H., Roy, A. A., Heo, J., Robinson, A. L. & Adams, P. J. Impact of natural gas development in the Marcellus and Utica shales on regional ozone and fine particulate matter levels. *Atmos. Environ.* **155,** 11–20 (2017).

80. Swarthout, R. F. *et al.* Impact of marcellus shale natural gas development in southwest Pennsylvania on volatile organic compound emissions and regional air quality. *Environ. Sci. Technol.* **49,** 3175–3184 (2015).

81. Allred B. W.   D. Twidwell, J. H. Haggerty, S. W. Running, D. E. Naugle, S. D. Fuhlendorf, W. K. S. Ecosystem services lost to oil and gas in North America. Net primary production reduced in crop and rangelands. *Science (80-.).* **348,** 401–402 (2015).

82. Abrahams, L. S., Samaras, C., Griffin, W. M. & Matthews, H. S. Life cycle greenhouse gas emissions from U.S. liquefied natural gas exports: Implications for end uses. *Environ. Sci.*

Technol. **49,** 3237–3245 (2015).

83. Heath, G. A., O'Donoughue, P., Arent, D. J. & Bazilian, M. Harmonization of initial estimates of shale gas life cycle greenhouse gas emissions for electric power generation. *Proc. Natl. Acad. Sci.* **111,** E3167–E3176 (2014).

84. Weber, C. L. & Clavin, C. Life Cycle Carbon Footprint of Shale Gas: Review of Evidence and Implications. *Environ. Sci. Technol.* **46,** 5688–5695 (2012).

85. Jiang, M. *et al.* Life cycle greenhouse gas emissions of Marcellus shale gas. *Environ. Res. Lett.* **6,** (2011).

86. Paredes, D., Komarek, T. & Loveridge, S. Income and employment effects of shale gas extraction windfalls: Evidence from the Marcellus region. *Energy Econ.* **47,** 112–120 (2015).

87. Weber, J. G. A decade of natural gas development: The makings of a resource curse? *Resour. Energy Econ.* **37,** 168–183 (2014).

88. Weber, J. G. The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming. *Energy Econ.* **34,** 1580–1588 (2012).

89. Colborn, T., Kwiatkowski, C., Schultz, K. & Bachran, M. Natural Gas Operations from a Public Health Perspective. *Hum. Ecol. Risk Assess.* **17,** 1039–1056 (2011).

90. Hays, J. & Shonkoff, S. B. C. Toward an understanding of the environmental and public health impacts of shale gas development: an analysis of the peer-reviewed scientific literature, 2009-2015. 1–24 (2015).

91. Muller, N. Z. & Mendelsohn, R. *The Air Pollution Emission Experiments and Policy Analysis Model (APEEP) Technical Appendix*. (2006).

92. Muller, N. Z. Boosting GDP growth by accounting for the environment. *Science (80-.).* **345,** 873–874 (2014).

93. Heo, J., Adams, P. J. & Gao, H. O. Public health costs accounting of inorganic PM2.5pollution in metropolitan areas of the United States using a risk-based source-receptor model. *Environ. Int.* **106,** 119–126 (2017).

94. Tessum, C. W., Hill, J. D. & Marshall, J. D. InMAP: A model for air pollution interventions. *PLoS One* **12,** 9281–9321 (2017).

95. Pope III, C. A. *et al.* Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine

Particulate Air Pollution. **287,** (2002).

96. Lepeule, J., Laden, F., Dockery, D. & Schwartz, J. Chronic exposure to fine particles and mortality: An extended follow-up of the Harvard six cities study from 1974 to 2009. *Environ. Health Perspect.* **120,** 965–970 (2012).

97. U.S. EPA. *Guidelines for Preparing Economic Analyses*. **2010,** (2014).

98. Shine, K. P., Fuglestvedt, J. S., Hailemariam, K. & Stuber, N. Alternatives to the Global Warming Potential for comparing climate impacts of emissions of greenhouse gases. *Clim. Change* **68,** 281–302 (2005).

99. Boucher, O. & Reddy, M. S. Climate trade-off between black carbon and carbon dioxide emissions. *Energy Policy* **36,** 193–200 (2008).

100. Fuglestvedt, J. S. *et al.* Transport impacts on atmosphere and climate: Metrics. *Atmos. Environ.* **44,** 4648–4677 (2010).

101. Berntsen, T. & Fuglestvedt, J. Global temperature responses to current emissions from the transport sectors. *Proc. Natl. Acad. Sci.* **105,** 19154–19159 (2008).

102. Peters, G. P., Aamaas, B., Berntsen, T. & Fuglestvedt, J. S. The integrated global temperature change potential (iGTP) and relationships between emission metrics. *Environ. Res. Lett.* **6,** 044021 (2011).

103. Shindell, D. *et al.* Climate, health, agricultural and economic impacts of tighter vehicle-emission standards. *Nat. Clim. Chang.* **1,** 59–66 (2011).

104. Chang, C. Y. *et al.* Investigating ambient ozone formation regimes in neighboring cities of shale plays in the Northeast United States using photochemical modeling and satellite retrievals. *Atmos. Environ.* **142,** 152–170 (2016).

105. Alvarez, R. A. *et al.* Assessment of methane emissions from the U . S . oil and gas supply chain. **7204,** 1–9 (2018).

106. O'Neill, B. C. The jury is still out on Global Warming Potentials. *Clim. Change* **44,** 427–443 (2000).

107. Fuglestvedt, J. S. *et al.* Metrics of climate change: Assessing radiative forcing and emission indices. *Clim. Change* **58,** 267–331 (2003).

108. Myhre, G. *et al.* Anthropogenic and Natural Radiative Forcing. *Clim. Chang. 2013 Phys. Sci.*

Basis. Contrib. Work. Gr. I to Fifth Assess. Rep. Intergov. Panel Clim. Chang. 659–740 (2013). doi:10.1017/ CBO9781107415324.018

109. Boucher, O. Comparison of physically- and economically-based CO2- equivalences for methane. *Earth Syst. Dyn.* **3,** 49–61 (2012).

110. Ricke, K., Drouet, L., Caldeira, K. & Tavoni, M. Country-level social cost of carbon. *Nat. Clim. Chang.* doi:10.1038/s41558-018-0282-y

111. Reilly, J. M. & Richards, K. R. Climate change damage and the trace gas index issue. *Environ. Resour. Econ.* **3,** 41–61 (1993).

112. Schmalensee, R. Comparing Greenhouse Gases for Policy Purposes. *Energy J.* **14,** 245–255 (2018).

113. Hammitt, J. K., Jain, A. K., Adams, J. L. & Wuebb. A welfare-based index for assessing environmental effects of GHG emissions. *Nature* **381,** 301–303 (1996).

114. Archer, D. Fate of fossil fuel CO_2 in geologic time. *J. Geophys. Res.* **110,** C09S05 (2005).

115. Shoemaker, J. K., Schrag, D. P., Molina, M. J. & Ramanathan, V. What Role for Short-Lived Climate Pollutants in Mitigation Policy ? **342,** 1323–1325 (2013).

116. Shindell, D. & Faluvegi, G. The net climate impact of coal-fired power plant emissions. *Atmos. Chem. Phys.* **10,** 3247–3260 (2010).

117. Marchand, J. The distributional impacts of an energy boom in Western Canada. *Can. J. Econ.* **48,** 714–735 (2015).

118. Marchand, J. Local labor market impacts of energy boom-bust-boom in Western Canada. *J. Urban Econ.* **71,** 165–174 (2012).

119. Marchand, J. & Weber, J. The Local Effects of the Texas Shale Boom on Schools, Students, and Teachers. *Ssrn* (2018). doi:10.2139/ssrn.3096293

120. Brown, J. P. Production of natural gas from shale in local economies: A resource blessing or curse? *Econ. Rev.* **99,** 119–147 (2014).

121. Feyrer, J., Mansur, E. T. & Sacerdote, B. Geographic dispersion of economic shocks: Evidence from the fracking revolution. *Am. Econ. Rev.* **107,** 1313–1334 (2017).

122. Weber, J. G. A decade of natural gas development: The makings of a resource curse? *Resour. Energy Econ.* **37,** 168–183 (2014).

123. Wei, M., Patadia, S. & Kammen, D. M. Putting renewables and energy efficiency to work. *Energy Policy* 1–36

124. Mayfield, E. N. *et al.* Cumulative air, climate, and employment impacts of natural gas systems.

125. U.S. Environmental Protection Agency. National Emissions Inventory. (2018). Available at: https://www.epa.gov/air-emissions-inventories/2014-national-emissions-inventory-nei-data.

126. U.S. Environmental Protection Agency. Continuous Emissions Monitoring System. (2018). Available at: https://ampd.epa.gov/ampd/.

127. US Census Bureau. American Community Survey. (2018). Available at: factfinder.census.gov.

128. Centers for Disease Control and Prevention. Population Projections, United States, 2004 - 2030, by state, age and sex, on CDC WONDER Online Database. (2005). Available at: http://wonder.cdc.gov/population-projections.html.

129. Trump, D. J. *Presidential Executive Order on Promoting Energy Independence and Economic Growth*. (2017).

130. Ehrgott, M. & Ruzika, S. Improved ε-constraint method for multiobjective programming. *J. Optim. Theory Appl.* **138,** 375–396 (2008).

131. Pennsylvania Department of Environmental Protection. PA DEP Oil & Gas Reporting Website. Available at: https://www.paoilandgasreporting.state.pa.us/publicreports/Modules/Welcome/Welcome.aspx.

132. Ohio Department of Natural Resources. Oil & Gas Resources. Available at: http://oilandgas.ohiodnr.gov/production.

133. West Virginia Department of Environmental Protection. Oil and Gas Production Data. Available at: https://dep.wv.gov/oil-and-gas/databaseinfo/Pages/default.aspx.

134. U.S. Geological Survey. National Assessment of Oil and Gas Project Devonian Marcellus Shale of the Appalachian Basin Province (067) Assessment Units. (2011).

135. U.S. Geological Survey. National Assessment of Oil and Gas Project - Utica Shale Unconventional Assessment. (2015).

136. Bell, E. & Clough, D. Just fracking: a distributive environmental justice analysis of unconventional gas development in Pennsylvania, USA. *Environ. Res. Lett.* **11,** 25001 (2016).

137. Emanuel, R. E. Flawed environmental justice analyses Mexico ' s basic science funding falls short

Fringe on the brink : Intertidal reefs at risk. **2018,** 2–5 (2018).

138. Healy, N., Stephens, J. C. & Malin, S. A. Embodied energy injustices: Unveiling and politicizing the transboundary harms of fossil fuel extractivism and fossil fuel supply chains. *Energy Res. Soc. Sci.* **48,** 219–234 (2019).

139. Klinsky, S. & Dowlatabadi, H. Conceptualizations of justice in climate policy. **9,** 88–108 (2009).

140. Centers for Disease Control and Prevention, N. C. for H. S. Underlying Cause of Death 1999-2016 on CDC WONDER Online Database, released December, 2017. Data are from the Multiple Cause of Death Files, compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Prog. (2018). Available at: http://wonder.cdc.gov/.

141. Mohai, P. & Saha, R. Racial Inequality in the Distribution of Hazardous Waste: A National-Level Reassessment. *Source Soc. Probl. UTC Soc. Probl.* **5443,** 343–370 (2007).

142. U.S. Census Bureau. Small Area Income & Poverty Estimates. Available at: https://www.census.gov/did/www/saipe/.

143. Organisation for Economic Co-operation and Deveopment. Income inequality (indicator). (2018). Available at: https://data.oecd.org/inequality/income-inequality.htm. (Accessed: 2nd August 2018)

144. Hsiang, S. *et al.* Estimating economic damage from climate change in the United States. *Science* **356,** 1362–1369 (2017).

145. Mohai, P. & Saha, R. Reassessing Racial and Socioeconomic Disparities in Environmental Justice Research. *Demography* **43,** 383–399 (2006).

146. Black, D., Mckinnish, T. & Sanders, S. The Economic Impact of the Coal Boom and Bust Author (s): Dan Black , Terra McKinnish and Seth Sanders Published by : Wiley on behalf of the Royal Economic Society Stable URL : http://www.jstor.org/stable/3590402 . THE ECONOMIC IMPACT OF THE COAL BOOM A. **115,** 449–476 (2005).

147. Marchand, J. & Weber, J. Local labor markets and natural resources: A synthesis of the literature. *J. Econ. Surv.* **00,** 1–22 (2017).

Appendix A: Supplemental Information for Chapter 2

Table of Contents

List of Figures

List of Tables

A1 Introduction

This supporting information document provides additional details regarding model formulations, assumptions, data, sources of uncertainty, and results. **Error! Reference source not found.** is a conceptual mode depicting six interacting sub-modules used in the assessment of policy options.

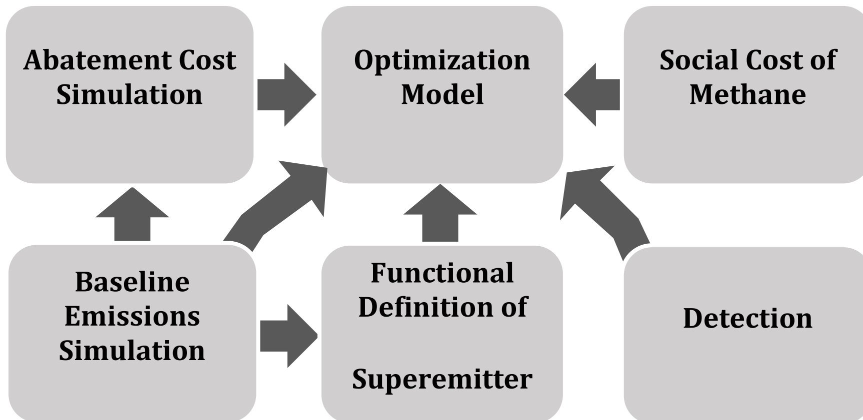

Figure A1. Conceptual model.

This study focuses on existing infrastructure in the transmissions and storage (T&S) sector. We define the T&S system as the infrastructure between the receipt meter with the gathering and processing sector and the delivery meter with the distribution system. The T&S sector contains both interstate and intrastate infrastructure, including compressor stations, pressurized pipeline networks, metering and regulation stations, and supporting equipment. This analysis excludes pipelines and concentrates on compressor stations and storage facilities.

We further focus this analysis on existing infrastructure, rather than new and modified natural gas facilities that are regulated by New Source Performance Standards (NSPS) under the Clean Air Act that were finalized by the Environmental Protection Agency (EPA) in June 2016.

A2 Input Parameter and Assumption Tables

A2.1 Baseline emissions simulation input parameters and distributions

The following table presents a list of additional key parameters and assumptions for the baseline emissions simulation. Table A1 presents the distributions of emission factors and operating hours for each component type in each operating mode.

Table A1. Emission factor and operating hour inputs and distributions.

Emission Category[1]		**Emissions Factor Distribution [Standard cubic feet per minute]**[2,3]	**Annual Operating Hours Distribution [Hours]**
Centrifugal Compressor	Blowdown Vent - NOP	Exponential (1.422)	TruncExponential (720,8784)
	Blowdown Vent - OP	Exponential (1.422)	TruncExponential (2636,8784)
	Isolation Valve - NOD	Exponential (2.10)	Uniform (0,8784)
	Dry Seal Vent - OP	Exponential (2.63)	TruncExponential (2869,8784)
	Wet Seal Vent - OP	Exponential (7.772)	TruncExponential (2313,8784)
Compressor Component Leaks	Connector	Lognormal (0.18,0.71)	8784
	Open Ended Line	Lognormal (3.65,48.2)	8784
	Pressure Release Valve	Lognormal (0.43,5.53)	8784
	Valve	Lognormal (0.20,0.88)	8784
	Meter	Lognormal (0.43,5.53)	8784
Non-Compressor Component Leaks	Connector	Lognormal (0.14,0.45)	8784
	Open Ended Line	Lognormal (1.41,19.9)	8784
	Pressure Release Valve	Lognormal (0.43,5.53)	8784
	Valve	Lognormal (0.27,1.51)	8784
	Meter	Lognormal (0.43,5.53)	8784
Pneumatics	High Bleed	Lognormal (0.35,1.03)	8784
	Intermittent Bleed	Not Modeled	Not Modeled
	Low Bleed	Lognormal (0.35,1.03)	8784
Reciprocal Compressor	Blowdown Vent - NOP	Exponential (1.42)	TruncExponential (2816,8784)
	Blowdown Vent - OP	Exponential (1.42)	TruncExponential (2992,8784)
	Isolation Valve - NOD	Exponential (3.52)	TruncExponential (2975,8784)
	RodPacking - NOP	Exponential (6.34)	TruncExponential (2816,8784)
	RodPacking - OP	Exponential (3.49)	TruncExponential (2992,8784)
Transmission Tank	Tank Vent	Exponential (3.27)	8784
	Tank Flare	Not Modeled	Not Modeled
Wellhead Components		Not Modeled	Not Modeled
Blowdowns		Not Modeled	Not Modeled
Combustion	Lean 2 Stroke - OP	Triangular (0,0.94,2.63)	TruncExponential (2977,8784)
	Lean 4 Stroke - OP	Triangular (0.09,0.75,3.22)	TruncExponential (2940,8784)
	Rich 4 Stroke - OP	Triangular (0,0.02,0.67)	TruncExponential (2984,8784)
	Turbine - OP	Exponential (0.0089)	TruncExponential (2667,8784)

1. Operating modes for components of reciprocating and centrifugal compressors are as follows: not operating pressurized (NOP), operating pressurized (OP), and not operating depressurized (NOD). For some components, operating modes are not differentiated.
2. Distribution designations are as follows: Exponential (lamba); Lognormal (mean, standard deviation); Triangular (minimum, median, maximum); TruncExponential (lambda, maximum); Uniform (minimum, maximum).
3. Although we leverage datasets compiled by and some assumptions from the Zimmerle et al. (2015) study, we do not employ their emissions model. Zimmerle et al. (2015) modeled emission factor and hour distributions differently, directly drawing from the empirical distributions. The purpose of the Zimmerle et al. (2015) study was to derive aggregate estimates for the T&S system, whereas the purpose of the modeling in this study is to develop the details of emissions from individual components and facilities to assess policy options. Thus, the functional form, treatment of uncertainty and variability, and other model details necessarily differ between the two models. We explicitly make note of these differences where appropriate.

A2.2 Abatement cost simulation input parameters and distributions

The following tables present a list of additional key parameters and assumptions for the abatement cost simulation. Table A2 presents brief descriptions of abatement measures for each component type in each operating mode. Table A3 includes the distributions of incremental abatement costs for each component type; note that these are not presented as annual costs, but rather total costs. Table A4 and Table A5 present the abatement interval and abatement efficacy for each component type, respectively.

Table A2. Abatement measure description.

Emission Source		**Abatement Measure**
Centrifugal Compressor	Blowdown Vent - NOP	Rebuild or replace valve.
	Blowdown Vent - OP	Rebuild or replace valve.
	Isolation Valve - NOD	Clean and inject sealant or replace valve.
	Wet Seal Vent - OP	Replace wet seal vents with dry seal vents.
Compressor Component Leaks	Connector	Replace or repair connector.
	Open Ended Line	Replace or repair open ended line.
	Pressure Release Valve	Replace or repair pressure release valve.
	Valve	Replace or repair valve.
Non-Compressor Component Leaks	Connector	Replace or repair connector.
	Open Ended Line	Replace or repair open ended line.
	Pressure Release Valve	Replace or repair pressure release valve.
	Valve	Replace or repair valve.
Pneumatics	High Bleed	Replace high bleed with low bleed.
Reciprocal Compressor	Blowdown Vent - NOP	Rebuild or replace valve.
	Blowdown Vent - OP	Rebuild or replace valve.
	Isolation Valve - NOD	Clean and inject sealant or replace valve.
	Rodpacking - NOP	Replace rodpacking rods and rings.
	Rodpacking - OP	Replace rodpacking rods and rings.
Transmission Tank	Tank Vent	Reroute to combustion component.

Table A3. Total cost of abatement.

Emission Source		Parameter Distribution[1,2] [2014 USD]	Source
Centrifugal Compressor	Blowdown Vent - NOP	Uniform (150, 1,200)	(Carbon Limits 2014; Clearstone 2006; EPA 2003; EPA 2014b)
	Blowdown Vent - OP	Uniform (150, 1,200)	(Carbon Limits 2014; Clearstone 2006; EPA 2003; EPA 2014b)
	Isolation Valve - NOD	Uniform (95, 4,000)	(EPA 2003; EPA 2014b)
	Wet Seal Vent - OP	Uniform (55,000, 110,000)	(EPA 2006b; EPA 2011; EPA 2014a; ICF 2014)
Compressor Component Leaks	Connector	Triangular (11, 66, 5,000)	(Carbon Limits 2006; Carbon Limits 2014; EPA 2003; EPA 2014b)
	Open Ended Line	Triangular (69, 200, 1,900)	(Clearstone 2006; EPA 2003; EPA 2014b)
	Pressure Release Valve	Triangular (90, 480, 830)	(Clearstone 2006; EPA 2003; EPA 2014b)
	Valve	Triangular (20, 50, 5,500)	(Carbon Limits 2014; Clearstone 2006; EPA 2014b)
Non-Compressor Component Leaks	Connector	Triangular (11, 66, 5,000)	(Carbon Limits 2006; Carbon Limits 2014; EPA 2003; EPA 2014b)
	Open Ended Line	Triangular (69, 200, 1,900)	(Clearstone 2006; EPA 2003; EPA 2014b)
	Pressure Release Valve	Triangular (90, 480, 830)	(Clearstone 2006; EPA 2003; EPA 2014b)
	Valve	Triangular (20, 50, 5,500)	(Carbon Limits 2014; Clearstone 2006; EPA 2014b)
Pneumatics	High Bleed	Triangular (580, 2,300, 9,800)	(EPA 2011; EPA 2014c; ICF 2014)
Reciprocal Compressor	Blowdown Vent - NOP	Uniform (150, 1,200)	(Carbon Limits 2014; Clearstone 2006; EPA 2003; EPA 2014b)
	Blowdown Vent - OP	Uniform (150, 1,200)	(Carbon Limits 2014; Clearstone 2006; EPA 2003; EPA 2014b)
	Isolation Valve - NOD	Uniform (95, 4,000)	(EPA 2003; EPA 2014b)
	Rodpacking - NOP	Uniform (5,900, 8,000)	(Carbon Limits 2014; EPA 2011; EPA 2014a; EPA 2015c; ICF 2014)
	Rodpacking - OP	Uniform (5,900, 8,000)	(Carbon Limits 2014; EPA 2011; EPA 2014a; EPA 2015c; ICF 2014)
Transmission Tank	Tank Vent	Uniform (91,000, 110,000)	(EPA 2006a; ICF 2014)

1. Estimates rounded to two significant figures.
2. Distribution designations are as follows: Triangular (low, median, high); Uniform (low, high).

Table A4. Abatement interval.

Emission Source		Parameter Value [Years]	Source
Centrifugal Compressor	Blowdown Vent – NOP	2	Average based on (Clearstone 2006; EPA 2014b)
	Blowdown Vent - OP	2	Average based on (Clearstone 2006; EPA 2014b)
	Isolation Valve - NOD	3	Average based on (EPA 2014b)
	Wet Seal Vent - OP	10	(EPA 2011; EPA 2014b)
Compressor Component Leaks	Connector	3	Average based on (Carbon Limits 2006; EPA 2014b)
	Open Ended Line	2	(Clearstone 2006; EPA 2014b)
	Pressure Release Valve	2	(Clearstone 2006; EPA 2014b)
	Valve	3	Average based on (EPA 2014b)
Non-Compressor Component Leaks	Connector	3	Average based on (Carbon Limits 2006; EPA 2014b)
	Open Ended Line	2	(Clearstone 2006; EPA 2014b)
	Pressure Release Valve	2	(Clearstone 2006; EPA 2014b)
	Valve	3	Average based on (EPA 2014b)
Pneumatics	High Bleed	10	(EPA 2011)
Reciprocal Compressor	Blowdown Vent - NOP	2	Average based on (Clearstone 2006; EPA 2014b)
	Blowdown Vent - OP	2	Average based on (Clearstone 2006; EPA 2014b)
	Isolation Valve - NOD	3	Average based on (EPA 2014b)
	Rodpacking - NOP	3	Average based on (Carbon Limits 2014; EPA 2015c)
	Rodpacking - OP	3	Average based on (Carbon Limits 2014; EPA 2015c)
Transmission Tank	Tank Vent	10	(Richards)

Table A5. Abatement efficacy.

Emission Source		Parameter Value	Source
Centrifugal Compressor	Blowdown Vent - NOP	0.95	(Carbon Limits 2014; EPA 2003; EPA 2014b)
	Blowdown Vent - OP	0.95	(Carbon Limits 2014; EPA 2003; EPA 2014b)
	Isolation Valve - NOD	0.93	Average based on (Carbon Limits 2014; EPA 2006; ICF 2014)
	Wet Seal Vent - OP	0.92	Average based on (EPA 2011; ICF 2014)
Compressor Component Leaks	Connector	0.95	(Carbon Limits 2014; ICF 2014)
	Open Ended Line	0.95	(Carbon Limits 2014; ICF 2014)
	Pressure Release Valve	0.95	(Carbon Limits 2014; ICF 2014)
	Valve	0.93	Average based on (Carbon Limits 2014; EPA 2006; ICF 2014)
Non-Compressor Component Leaks	Connector	0.95	(Carbon Limits 2014; ICF 2014)
	Open Ended Line	0.95	(Carbon Limits 2014; ICF 2014)
	Pressure Release Valve	0.95	(Carbon Limits 2014; ICF 2014)
	Valve	0.93	Average based on (Carbon Limits 2014; EPA 2006; ICF 2014)
Pneumatics	High Bleed	0.93	Average based on (EPA 2011; ICF 2014)
Reciprocal Compressor	Blowdown Vent - NOP	0.95	(Carbon Limits 2014; EPA 2003; EPA 2014b)
	Blowdown Vent - OP	0.95	(Carbon Limits 2014; EPA 2003; EPA 2014b)
	Isolation Valve - NOD	0.93	Average based on (Carbon Limits 2014; EPA 2006; ICF 2014)
	Rodpacking - NOP	0.79	Average based on (EPA 2014c; EPA 2015c)
	Rodpacking - OP	0.79	Average based on (EPA 2014c; EPA 2015c)
Transmission Tank	Tank Vent	0.95	(EPA 2006; EPA 2015c; ICF 2014)

A2.3 Detection input parameters

The following table presents a list of additional key parameters and assumptions for the detection sub-module. Table A6 presents the detection limits and costs.

Table A6. Detection input parameters.

Detection Method	Annual Detection Cost		Detection Limit	
	Parameter Value [2014 USD per facility]	**Source**	**Parameter Value [SCFM]**	**Source**
EPA Method 21	13,400[1]	(Clearstone 2006; EPA 2014; ICF 2014)	0.00001	(EPA 2014)
Optical Gas Imaging (Base Assumption)	5,700[2]	(Clearstone 2006; ICF 2013; ICF 2014)	0.024	(Allen 2013; EC/R Inc 2001)
Emerging Technology	469[3]	(EDF 2014b)	2.27	(EDF 2014b)

1. The detection costs are the sum of the annualized capital costs and labor costs. We assume detection surveys occur on a quarterly basis. The annualized capital costs were calculated assuming a discount rate of 3%, a 5-year capitalization period, and capital costs of detection equipment and other costs of $52,200. The labor costs are $3,100.
2. The detection costs are the sum of the annualized capital costs and labor costs. We assume detection surveys occur on a quarterly basis. The annualized capital costs were calculated assuming a discount rate of 3%, 5-year capitalization period, and capital costs of detection equipment and other costs of $153,000. The labor costs are $1,200.
3. The detection costs are the annualized costs. We assume detection occurs continuously. The annualized capital costs were calculated assuming a discount rate of 3%, 10-year equipment life, and capital costs of equipment of $1,000.

In other analyses (ICF 2014; Clearstone 2006), it was assumed that more frequent detection is proportional to increased abatement (e.g., an increase in detection frequency from annually to quarterly results in a four-fold increase in abated emissions). Although more frequent or continuous measurement conceivably would allow for more rapid detection and abatement, lacking continuous methane measurement data, we do not explicitly model this. Note that in addition to varying the detection method, to account for uncertainty in detection and subsequent timing of an abatement action, we perform sensitivity analysis of abatement efficacy and calculate detection costs at the break-even point (i.e., when system-wide net benefits are zero).

A2.4 Throughput input parameters

The following table presents a list of additional key parameters and assumptions for estimating throughput and proportional loss rates. Table A7 presents the distributions of horsepower multiplied by efficiency for reciprocating and centrifugal compressors. Distributions are based on datasets compiled by Zimmerle et al. (2015).

Compressor Type	Horsepower x Efficiency Distribution[1]
Reciprocating	Invgaussian (651,890)
Centrifugal	Lognormal (3897, 60,315)

1. Distribution designations are as follows: Lognormal (mean, standard deviation); Invgaussian (mean,lambda).

Table A7. Throughput input parameters.

A3 Baseline Methane Emissions Simulation Formulation

Figure A2 is a depiction of the types of components (e.g., rod-packing, isolation valves) or emission sources at transmission compressor stations and storage facilities.

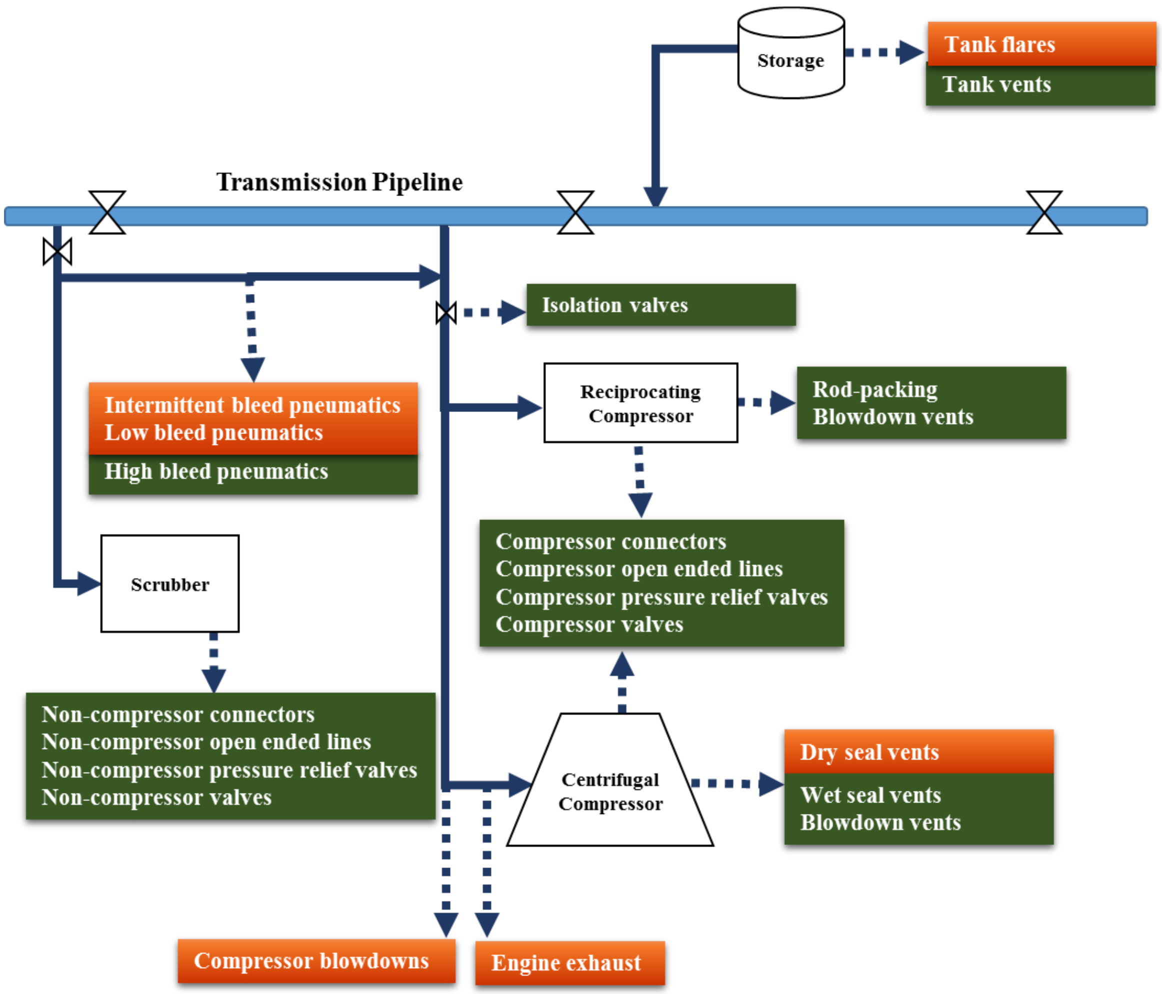

Figure A2. Transmission and storage system diagram. Orange boxes represent emission sources that we assume are not abatable. Green boxes represent emission sources that we assume are abatable.

Annual baseline emissions ($\tilde{e}_{i,j}^{\text{ANNUAL}}$) are simulated for each emission category $i = 1,\ldots, m$ (assuming there are m emission categories) and component number $j = 1,\ldots, n$ in the population (assuming there are n components). We account for uncertainty in emission factors ($\tilde{e}_i^{\text{EF}}$) and annual operating hours ($\tilde{h}_i$). Annual baseline emissions are simulated as follows:

0. Set $j = 0$
1. Draw random values from the emissions factor distribution ($\tilde{e}_i^{\text{EF}}$) and annual operating hour distribution ($\tilde{h}_i$) for each $j \in D_i = \{\text{component } j \text{ is of type } i\}$
2. Estimate annual emissions as follows:

$$\tilde{\mathbf{e}}_{i,j}^{\textbf{ANNUAL}} = \mathbf{60} \cdot \tilde{\mathbf{e}}_i^{\textbf{EF}} \cdot \tilde{\mathbf{h}}_i \tag{1}$$

3. Repeat steps 1-2, 100 times
4. $j = j+1$
5. Repeat steps 1-4 until $j=n$

Estimates of baseline annual emissions for each facility l in the U.S. natural gas transmission and storage sector are also simulated by assigning annual emissions for each component in the population to a facility. Each facility is described by a facility profile or inventory (i.e., number of components of each type at a facility), which are known, at least in part, for only some facilities in the population. However, for a subset of facilities in which facility profiles are unknown, we simulate profiles based on known facility profiles, accounting for correlations between numbers and types of components. Based on the facility profiles, we then assign component-level emissions to each facility, and develop facility-level emission estimates:

$$\tilde{\mathbf{e}}_l^{\textbf{FACILITY}} = \sum_{j \in \mathbf{F}_l = \{\textbf{component } j \textbf{ is of facility } l\}} \tilde{\mathbf{e}}_j^{\textbf{ANNUAL}} \tag{2}$$

This model simulates 100 realizations of each component and facility in the sector.

A4 Abatement Cost Simulation Mathematical Formulation

For each component type i, we estimate annualized costs of abatement ($\tilde{c}_i^{\mathrm{ABATEMENT}}$), representing the present value of the capital cost of an abatement measure over an assumed financial life, converted to equal annual payments. Annualized costs are a function of the capital recovery factor (CRF) and capital expenditures for the abatement measure for component type i ($\widetilde{\mathrm{CAPEX}}_i^{\mathrm{ABATEMENT}}$). The capital costs are incremental costs beyond the status quo. The CRF is a function of the discount factor r and the abatement interval t (e.g., replacement interval). Abatement costs are calculated as follows:

$$\tilde{c}_i^{\mathrm{ABATEMENT}} = \mathrm{CRF} \cdot \widetilde{\mathrm{CAPEX}}_i^{\mathrm{ABATEMENT}} \tag{3}$$

where the CRF is given by the following equation:

$$\mathrm{CRF} = \frac{r \cdot (1+r)^t}{(1+r)^t - 1} \tag{4}$$

Abatement costs ($\tilde{c}_{i,j}^{\mathrm{ABATEMENT}}$) are simulated for each component type $i = 1,\ldots, m$ (assuming there are m component types) and component number $j = 1,\ldots, n$ in the population (assuming there are n components total of all types). Abatement costs are simulated as follows:

0. Set $j = 0$
1. Draw a random value from the levelized annual cost of abatement distribution ($\tilde{c}_i^{\mathrm{ABATEMENT}}$) for each $j \in \mathrm{D}_i = \{\text{component } j \text{ is of type } i\}$
2. Draw a random value from the baseline annual emissions ($\tilde{e}_{i,j}^{\mathrm{ANNUAL}}$)
3. Estimate abatement costs as follows:

$$\boldsymbol{\tilde{c}_{i,j}^{\mathrm{ABATEMENT}} = \tilde{c}_i^{\mathrm{ABATEMENT}} - 1.1 \cdot \mathrm{A}_i \cdot \tilde{e}_{i,j}^{\mathrm{ANNUAL}} \cdot \mathrm{C}^{\mathrm{NG}}} \tag{5}$$

4. Repeat steps 1-3, 100 times
5. $j = j+1$
6. Repeat steps 1-5 until $j=n$

Note that in Equation 5, to estimate NG emissions, we multiply methane emissions by a factor of 1.1, which assumes that the methane content of NG is 90.4%.

A5 Social cost of methane comparison

The social cost of methane emissions can be estimated indirectly or directly. Indirect estimation, suggested by Price et al. (2007), converts non-CO_2 emissions to carbon dioxide equivalents (CO_2eq) using global warming potentials (GWPs), then applies estimates of the social cost of carbon (SCC). GWP – which is a simplified index relating the contribution of non-CO_2 greenhouse gas emissions to long-run measures of atmospheric radiative forcing to that of CO_2 – was developed for practical purposes of transparency and ease of use (Marten and Newbold 2012). The metric, GWP, has been critiqued from economic and scientific perspectives for reasons such as, among other things, arbitrary time horizons and constant concentration assumptions. A few studies evaluated the magnitude of error in the context of cost-effectiveness analysis, finding that the increased cost of abatement required to achieve a temperature stabilization target due to utilizing proxies based on GWPs may be small in relative terms, but may not be insubstantial in absolute (Johansson et al. 2006).

While there are limitations of indirect estimation, minimal direct estimates of the marginal costs of methane exist. Marten and Newbold (2015) estimated the social costs of methane for the years 2010 to 2050 using an integrated assessment model that combines the DICE, PAGE, and FUND integrated assessment models with a climate model, Model for the Assessment of Greenhouse-gas Induced Climate Change (MAGICC), employing assumptions similar to those used by the United States Government Interagency Working Group on the Social Cost of Carbon (IWG SCC). This study suggests that for single-gas policies, such as emissions abatement in natural gas systems, the estimation error when using GWPs can be large; the global damage potential for methane, which is analogous to GWP but based on direct estimation, is 4-84% higher in 2020 than the 100-year GWP reported in the IPCC's Fourth Assessment Report (Marten et al. 2014; Marten et al. 2015).

For comparative purposes, Table A8 includes direct estimates for abatement conducted in 2020 and 2050, as well as indirect estimates, derived using 100-year GWPs of 28 to 34, as reported in the IPCC's Fifth Assessment Report, and the SCC estimates reported by the IWG SCC (IWG SCC 2015; Myhre et al. 2013). This comparison suggests that direct and indirect estimates are similar. In this analysis, we use direct estimates ranging from 601 to 3560 USD per metric ton of methane.

Table A8. Social Cost of Methane [Units: 2014 USD per ton of methane].

Year	Direct Estimates				Indirect Estimates			
	r=5% (mean)	*r*=3% (mean)	*r*=2.5% (mean)	*r*=3% (95^{th} percentile)	*r*=5% (mean)	*r*=3% (mean)	*r*=2.5% (mean)	*r*=3% (95^{th} percentile)
2020	601	1330	1780	3560	374 - 454	1310 - 1590	1930 - 2350	3830 - 4650
2050	1450	2780	3450	7450	810 - 983	2150 - 2610	2960 - 3590	6600 - 8020

A6 Optimization model formulation

Section A6.1 and A6.2 are the optimization formulations for the system-wide and superemitter policies, respectively. Table A9 includes parameter and decision variable descriptions.

A6.1 System-wide formulation

For the system-wide policies, we assume that all components in the T&S system are regulated, and detection and abatement activities occur at all facilities. Two system-wide formulations, aligning with different policy options, are derived: 1) an unconstrained formulation in which net social benefits are maximized (Section A6.1.1), and 2) a uniform percent emissions reduction target for each facility (10/50/75%) in which private costs are minimized, while meeting target (Section A6.1.2). The models select the subset of components in the population to abate.

A6.2 Unconstrained formulation

The objective function of the unconstrained formulation, in which net social benefits are maximized, is as follows:

$$\{\mathbf{Min}_{\{x_{i,j,k,l}\}} \sum_{i,j}\{\boldsymbol{x_{i,j,k,l}} \cdot [\tilde{\mathbf{c}}_{i,j,k,l}^{\mathbf{ABATEMENT}} - \mathbf{A}_{i} \cdot \tilde{\mathbf{e}}_{i,j,k,l}^{\mathbf{ANNUAL}} \cdot (\mathbf{1.1} \cdot \mathbf{C}^{\mathbf{NG}} + \mathbf{C}^{\mathbf{SCCH4}})]\} \tag{6}$$

$$+\mathbf{Y} \cdot \mathbf{C}^{\mathbf{DETECTION}}\}_{\boldsymbol{k}}$$

$x_{i,j,k,l} \in \{0,1\}\ \forall i,j,k,l$ is a binary decision variable, where a value of one indicates abatement of emissions for component type *i*, component number *j* in the population, and facility *l*, and a value of zero indicates emissions are not abated. Each realization *k* aligns with a single iteration of the Monte Carlo simulation of baseline emissions and abatement costs for the entire system; in other words, each realization is a random configuration of each component and facility in the system. Essentially, we are performing 100 sub-optimization problems.

We assume that only emissions that are detected may be abated. Thus, we perform a pre-processing step prior to the optimization in which we eliminate components in the system in which the emissions factor in units of standard cubic feet per minute ($\tilde{e}_{i,j,k,l}^{EF}$) is less than the detection limit ($S^{DETECTION}$) for a given detection technology.

In alternative formulations of the base model, we assume a system-wide emissions reduction target ($\mathrm{R}_{\mathrm{SYSTEMWIDE}}^{\mathrm{PERCENTAGE}}$) that cannot be exceeded:

$$\frac{\sum_{i,j,l} \tilde{\mathrm{e}}_{i,j,k,l}^{\mathrm{ANNUAL}} \cdot \mathrm{A}_i \cdot x_{i,j,k,l}}{\sum_l \tilde{\mathrm{e}}_{k,l}^{\mathrm{UNABATABLE}} + \sum_{i,j,l} \tilde{\mathrm{e}}_{i,j,k,l}^{\mathrm{ANNUAL}}} \leq \mathrm{R}_{\mathrm{SYSTEMWIDE}}^{\mathrm{PERCENTAGE}} \;\; \forall k \tag{7}$$

From a policy implementation perspective, we perform *ex-post* calculations of component-level performance standards or taxes to achieve the optimal policy.

Facility-level percentage reduction target

The objective function of the unconstrained formulation is as follows:

$$\{\mathrm{Min}_{\{x_{i,j,k,l}\}} \sum_{i,j} \{x_{i,j,k,l} \cdot [\tilde{\mathrm{c}}_{i,j,k}^{\mathrm{ABATEMENT}} - \mathrm{A}_i \cdot \tilde{\mathrm{e}}_{i,j,k,l}^{\mathrm{ANNUAL}} \cdot (1.1 \cdot \mathrm{C}^{\mathrm{NG}})]\} \tag{8}$$

$$+\mathrm{Y} \cdot \mathrm{C}^{\mathrm{DETECTION}}\}_k$$

It is the same formulation as the net benefit maximization objective function (Equation F-1), except the social cost of methane term is removed. We minimize private costs, rather than maximize net social benefits because we assume (from a policy implementation perspective) that the policy will be a facility-level percentage reduction target, whereby the facility can choose the optimal (least cost) abatement strategy.

We similarly account for detection limits in this formulation. In addition, we instill the logic that each facility l must reduce emissions enough to just achieve the facility-level emissions reduction target ($\mathrm{R}_l^{\mathrm{PERCENTAGE}}$), which we assume is 10/50/75% in alternative scenarios:

$$\frac{\sum_{i,j} \tilde{\mathrm{e}}_{i,j,k,l}^{\mathrm{ANNUAL}} \cdot \mathrm{A}_i \cdot x_{i,j,k,l}}{\tilde{\mathrm{e}}_{k,l}^{\mathrm{UNABATABLE}} + \sum_{i,j} \tilde{\mathrm{e}}_{i,j,k,l}^{\mathrm{ANNUAL}}} \leq \mathrm{R}_l^{\mathrm{PERCENTAGE}} \;\; \forall k, l \tag{9}$$

A facility abates components in the order of least to most marginal abatement costs, until it conducts just enough abatement to achieve the target. If a facility cannot achieve an emissions reduction target, given that some emissions are not abatable and abatement is not completely effective, then the facility must perform the maximum level of abatement possible.

A6.3 Superemitter formulation

For the superemitter policies, we assume that bottom-up detection is conducted at all facilities, and abatement is only conducted at a subset of facilities. Two superemitter formulations, aligning with different policy options, are derived: 1) net social benefits are maximized for a subset of facilities with absolute annual emissions exceeding an ex-ante threshold (50th and 90th percentile absolute annual emissions across all facilities) (Section A6.3.1), and 2) net social benefits are maximized for a subset of facilities with proportional loss rates exceeding an ex-ante threshold (50th and 90th percentile proportional loss rate across all facilities) (Section A6.3.2). Refer to Figure A3 for a depiction of the ex-ante absolute annual emissions and proportional loss rate thresholds. The models select the components to abate.

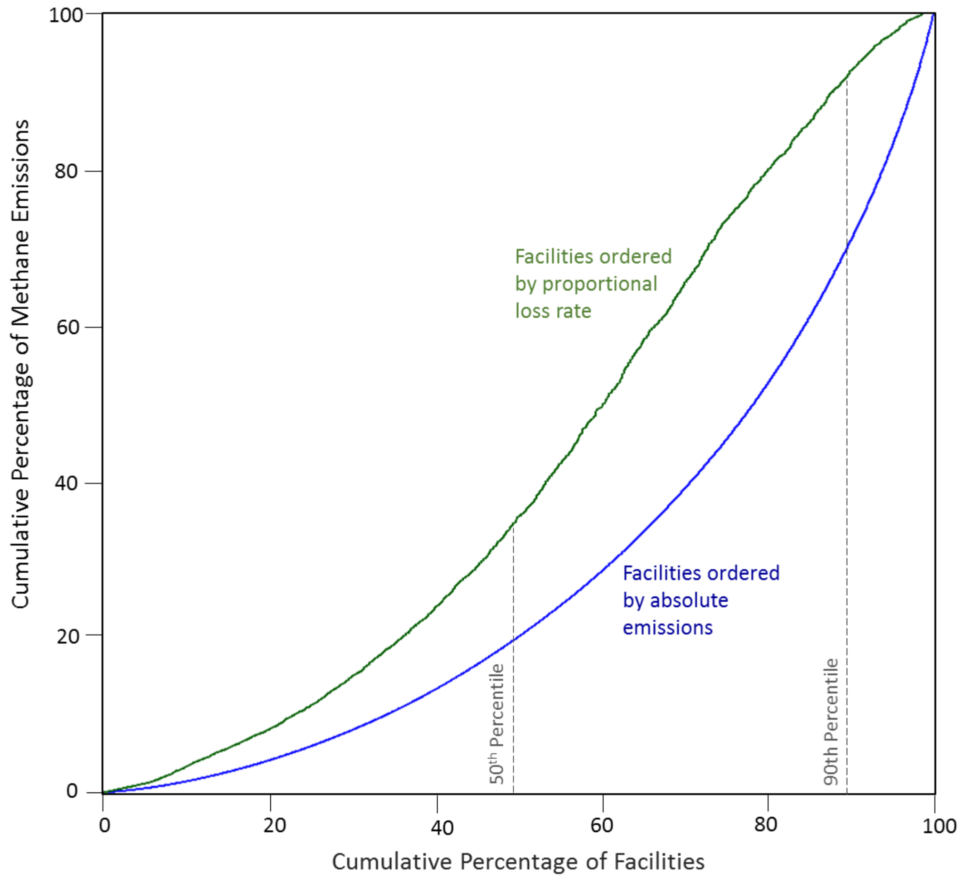

Figure A3. Cumulative percentage of facilities versus the cumulative percentage of methane emissions. The blue line represents simulated emissions (for a single model iteration), whereby facilities are ordered from least to greatest absolute annual emission. The green line represents simulated emissions (for a single model iteration), whereby facilities are ordered from least to greatest proportional loss rate. The vertical gray lines depict 50th and 90th percentiles of facilities.

A6.3.1 Absolute annual emissions threshold

The formulation is the same as for the unconstrained system-wide policy, except that abatement is conducted at a subset of facilities with absolute annual emissions ($\tilde{e}_{k,l}^{\mathrm{UNABATABLE}} + \sum_{i,j} \tilde{e}_{i,j,k,l}^{\mathrm{ANNUAL}}$) that are less than an *ex-ante* absolute annual emissions threshold ($F^{\mathrm{STANDARD,ABSOLUTE}}$).

A6.3.2 Proportional loss rate threshold

The formulation is the same as for the unconstrained system-wide policy, except that abatement is conducted at a subset of facilities with proportional loss rates ($\mathrm{L}_{k,l}$) that are less than an *ex-ante* proportional loss rate threshold ($\mathrm{F}^{\mathrm{STANDARD,RATE}}$).

Table A9. Parameter and Decisions Variable Descriptions.

Parameter	Description
$x_{i,j,k,l}$	{0 – do not abate emissions from component type *i*, population component number *j*, and for realization *k*; 1 – abate emissions}
A_i	Abatement efficacy (percentage emissions reduction) [%]
Y	Number of facilities in population
$\mathrm{S}^{\mathrm{DETECTION}}$	Detection method threshold [SCFM]
$\mathrm{R}^{\mathrm{PERCENTAGE}}_{\mathrm{SYSTEMWIDE}}$	Percentage emissions reduction target for the system [%]
$\mathrm{R}^{\mathrm{PERCENTAGE}}_{l}$	Percentage emissions reduction target for facility *l* [%]
$\mathrm{F}^{\mathrm{STANDARD,RATE}}$	Ex-ante proportional loss rate threshold for facility [%]
$\mathrm{F}^{\mathrm{STANDARD,ABSOLUTE}}$	Ex-ante annual emissions threshold for facility [SCF]
$\mathrm{L}_{k,l}$	Proportional loss rate of facility *l* for realization *k* []
$\mathrm{C}^{\mathrm{DETECTION}}$	Annual cost of onsite detection for each facility [2014 USD]
$\tilde{\mathrm{c}}^{\mathrm{ABATEMENT}}_{i,j,k,l}$	Annualized abatement cost for component type *i*, population component number *j*, and for realization *k*; [2014 USD]
C^{NG}	Market price of natural gas [2014 USD per SCF]
$\mathrm{C}^{\mathrm{SCCH4}}$	Social cost of methane [2014 USD per SCF]
$\tilde{\mathrm{e}}^{\mathrm{ANNUAL}}_{i,j,k,l}$	Annual emissions from component type *i*, component number *j*, and for realization *k* [SCF]
$\tilde{\mathrm{e}}^{\mathrm{EF}}_{i,j,k,l}$	Minutely emissions from component type *i*, component number *j*, and for realization *k* [SCFM]
$\tilde{\mathrm{e}}^{\mathrm{UNABATABLE}}_{k,l}$	Annual emission from sources that are not considered abatable for facility *l* and realization *k* [SCF]

A7 Uncertainty

Table A10 summarizes the sources of uncertainty that we currently incorporate into the model and those we exclude. We model uncertain inputs in two way, as either probabilistic inputs in the simulation or parametric inputs in a sensitivity analysis. In the paper, we explicitly qualify uncertainty reflected in our estimates and describe potential impacts of omitted uncertainties on policy recommendations.

We parametrically vary uncertainties in detection technologies, discount rates, social costs of methane, abatement efficacy, and natural gas prices. Other sources of uncertainty that we do not model (e.g., correlation between emissions and abatement costs, uncertainty in activity factors, spatial heterogeneity in abatement costs, changes in T&S infrastructure over time, and different years of abatement) will likely augment the variation between realizations of the system. Of the sources of uncertainty that we do not model, we believe that incorporating spatial heterogeneity may provide insight into potential differences in policy recommendations by region; however, we do not believe incorporating the other sources of uncertainty that we list will alter policy recommendations in a material way.

We acknowledge that the measurements for each emission category may not be representative of the entire population; however, the Subramanian et al. (2014) study was a substantial effort and a vast improvement on previous emission measurements. Site selection for the Subramanian et al. (2014) study was not random, but rather constricted by location, survey team schedules, and site suitability. Sites were reported as being broadly representative of the fleets of participating companies, which comprise ∼56% of interstate transmission facilities reporting to FERC (Zimmerle et al. 2015). We do not aggregate across studies because the dataset used is relatively robust. In addition, Brandt et al. (2016) outline the difficulties of cross-study aggregation; they show that measurements from similarly named sources across different studies rarely pass the Kolmogorov−Smirnov test, which indicates the studies' samples seem to be drawn from different underlying populations.

In an effort to reflect uncertainty and skewness in emission factors, we employ parametric distributions. Brandt et al. (2016) find that fitted parametric distribution may generate a narrow uncertainty range around the mean estimate and poorly represent the upper tail because body observations may outweigh the small number of tail observations. However, nonparametric approaches, such as resampling, require an adequate sample size and a limited dataset may not include higher emissions (Brandt et al. 2016; Zavala-Araiza et al. 2017).

In addition, the exact number of T&S facilities and components is unknown because only a subset of transmission and storage facilities annually report to the GHGRP and FERC, and the level of information collected by these reporting programs does not necessarily align with modeling needs. Zimmerle et al.

(2015) explicitly modeled uncertainty in activity counts, which was appropriate for that modeling effort, given the purpose was to estimate aggregate emissions across the T&S system. However, we do not incorporate uncertainty in the number of facilities and assume the mean estimate reported in Zimmerle et al. (2015).

While we model different detection methods and perform a bounding analysis, there are several detection methods with associated costs, sensitivity, and error rates, which we do not model and that are out of the scope of this study. Performing a more robust analysis of detection is a useful future extension of this model, potentially requiring structural changes to the model that capture error rates.

Table A10. Sources of uncertainty and variability.

Uncertainty and variability accounted for in modeling	**Uncertainty and variability not accounted for in modeling**
• Emission factors • Operating hours • Correlation between emission factors and operating hours • Abatement capital costs • Correlation between types and counts of components at facilities • Discount rate • Natural gas prices • Abatement efficacy • Detection technology • Social cost of methane	• Emissions measurement methods • Number of components • Number of facilities • Correlation between emissions and location • Correlation between costs and emissions • Correlation between emissions and efficacy • Spatial heterogeneity in abatement labor costs • Detection limit • Replacement interval • Year of analysis (i.e., year of abatement) • Change in transmission and storage system infrastructure over time

A8 Additional Results

Section A8.1 includes results from the baseline methane emissions model. Section A8.2 includes additional marginal abatement cost curves for emissions category. Section A8.3 includes an additional aggregate marginal abatement cost curve step function. A8.4 includes additional sensitivity analyses. A8.5 includes the estimated performance standards.

A8.1 Baseline methane emissions results

Figure A4 depicts the simulated mean methane emissions by emissions category for the T&S system. The estimated aggregate methane emissions for the T&S sector (440 MT) is within the range of published values: 2014 Greenhouse Gas Reporting Program (210 MT), 2014 Greenhouse Gas Initiative (1.9 MMT), and Zimmerle et al. (2015) (545 MT +128/-80%; 1.4 MMT +36/-23%). These figures vary because of differing modeling approaches, data, incorporation of uncertainty and variability, and source inclusion. Some emission categories (including intermittent bleed valves, wellhead components, transmission tank flares, blowdowns) were excluded from this analysis based on the availability of data.

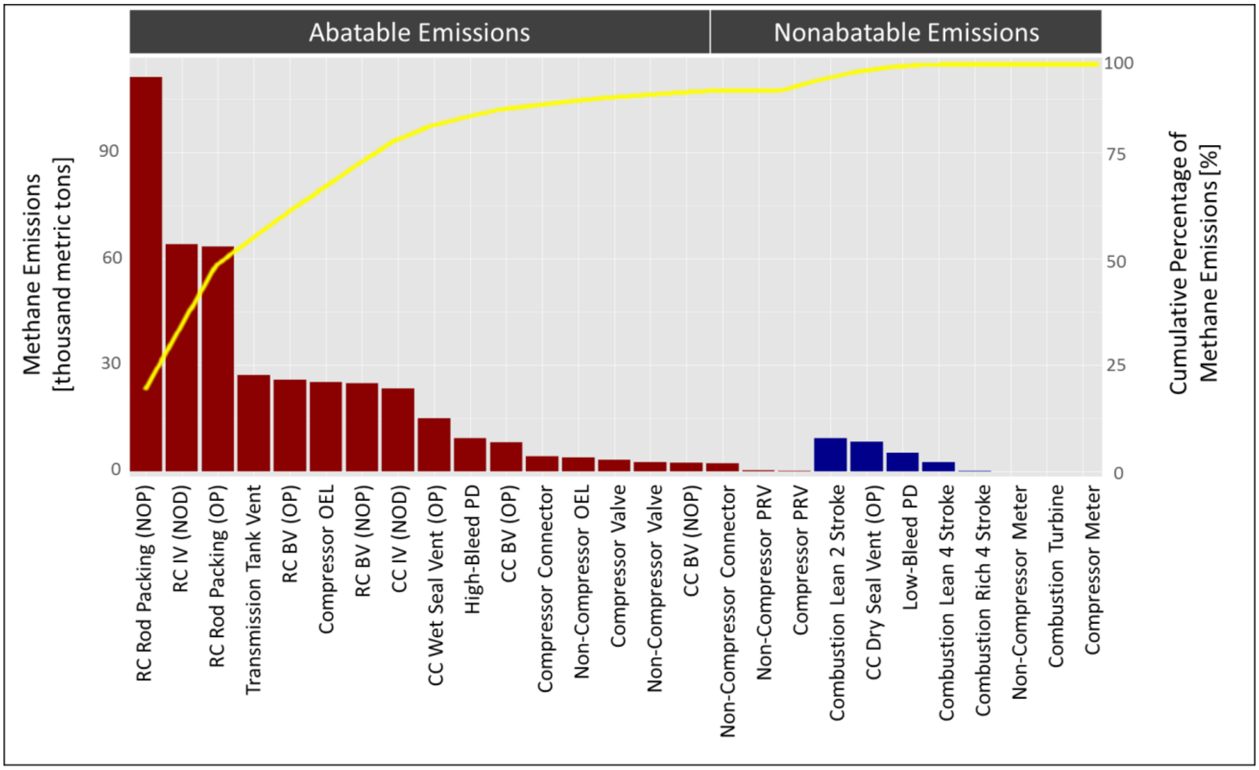

Figure A4. Mean methane emissions by emission category for the entire U.S. T&S system, accounting for variability in emissions factors and operating hours. Emission categories for which abatement is and is not possible, given the current slate of abatement technologies, are indicated in maroon and blue, respectively. The cumulative percentage of emissions attributable to each emission category is depicted by the yellow line. Approximately 94% of emissions are from abatable emission categories, with a few emission categories (i.e., rod-packing, reciprocating compressor isolation valves) accounting for approximately 57% of emissions from abatable emission categories. Component abbreviations are as follows: CC – centrifugal compressor, RC – reciprocating compressors, BV – blowdown valve, IV – isolation valve, OEL – open ended line, PRV – pressure release valve. Operating modes for components of reciprocating and centrifugal compressors are as follows: not operating pressurized (NOP), operating pressurized (OP), and not operating depressurized (NOD).

A8.2 Component-level marginal abatement cost curves

Figure A5 presents marginal abatement cost curves for each type of component, under conditions without and with natural gas savings, and employing base case assumptions. Base case model assumptions are as follows: discount rate = 3%, detection technology – optical gas imaging, social cost of methane – 3% discount rate mean scenario, and no savings from capturing marketable natural gas. Each marginal abatement cost curve is based on a single realization of the U.S. transmission and storage system. The marginal abatement cost curves for each component include one or more abatement measures. These curves account for uncertainty in both abatement costs and emissions.

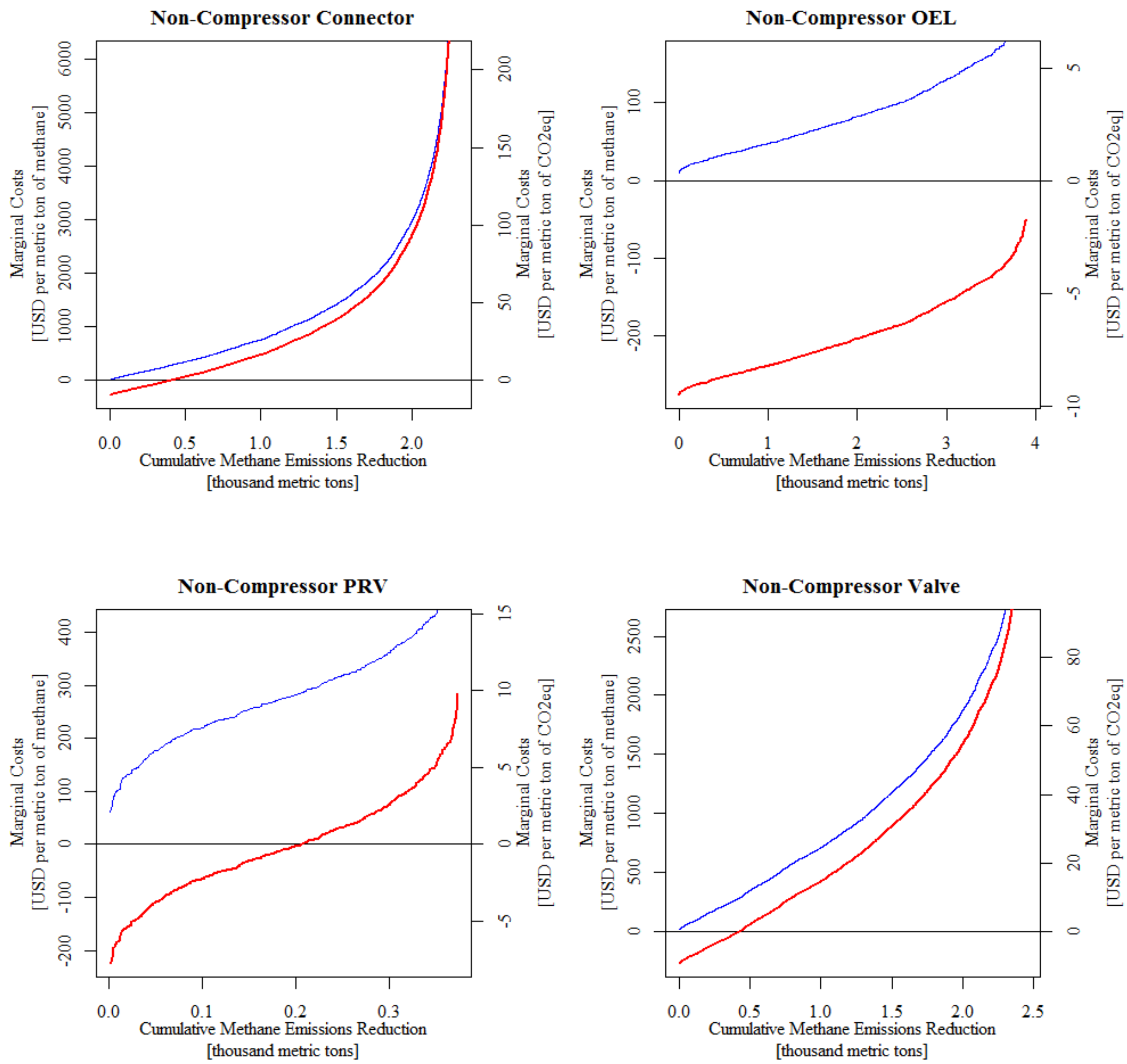

Figure A5. Marginal abatement cost curve for each emissions category. Blue lines assume no savings. Red lines assume savings, where the natural gas prices is the EIA projection for 2020 for the reference case ($4.88/MMBTU). Component abbreviations are as follows: CC – centrifugal compressor, RC – reciprocating compressors, BV – blowdown valve, IV – isolation valve, OEL – open ended line, PRV – pressure release valve. Operating modes for components of reciprocating and centrifugal compressors are as follows: not operating pressurized (NOP), operating pressurized (OP), and not operating depressurized (NOD).

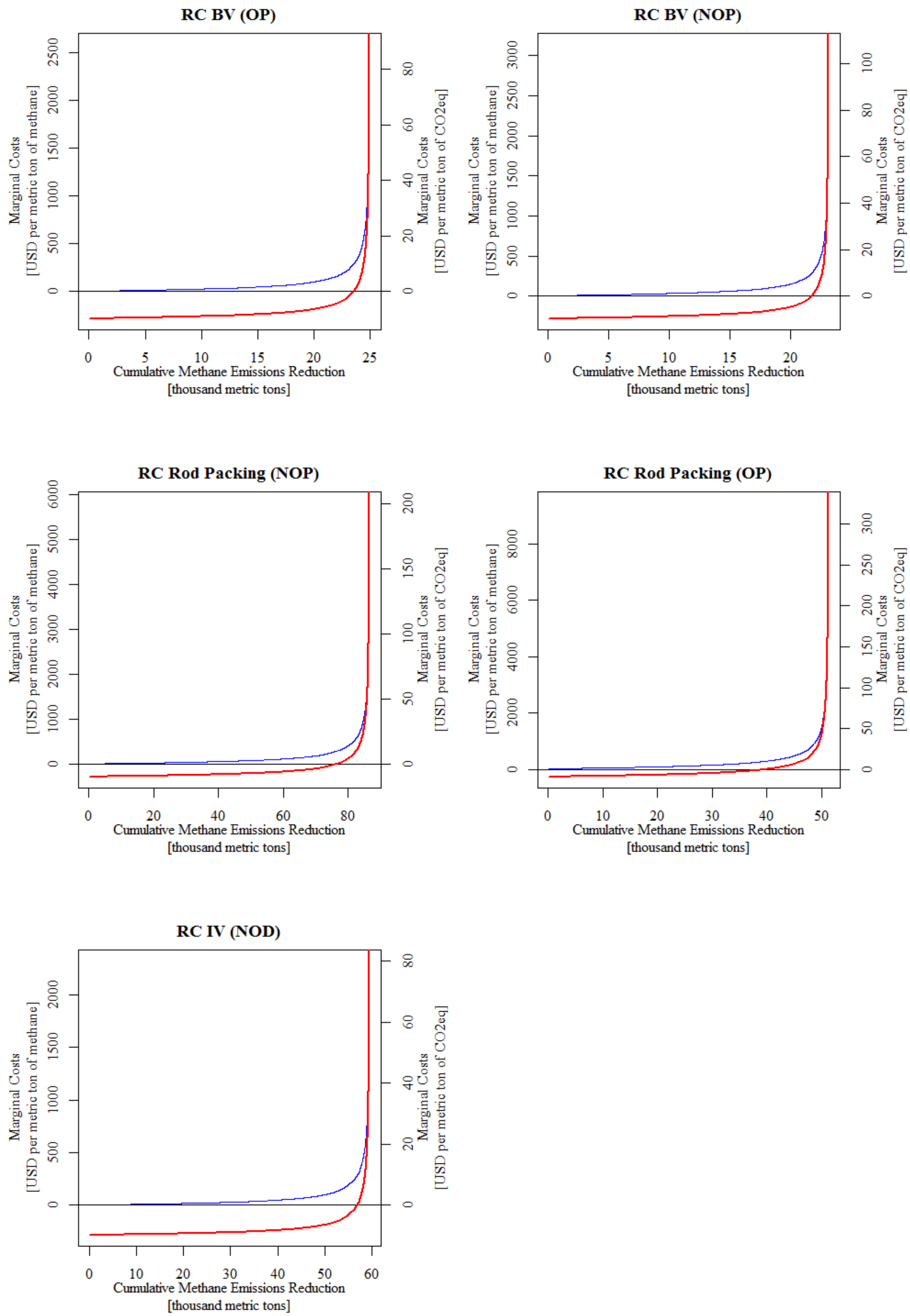

Figure A5 (continued)

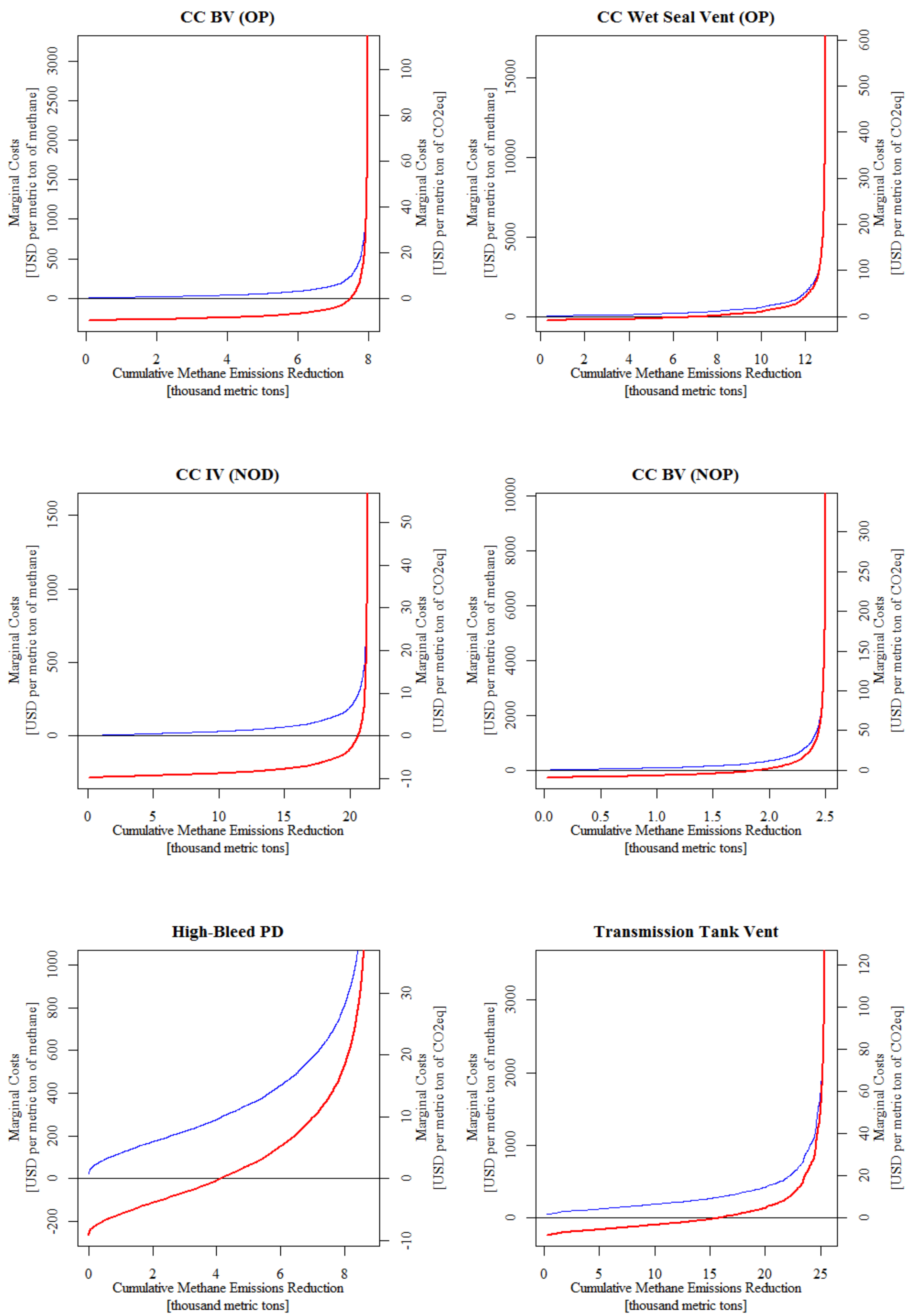

Figure A5 (continued)

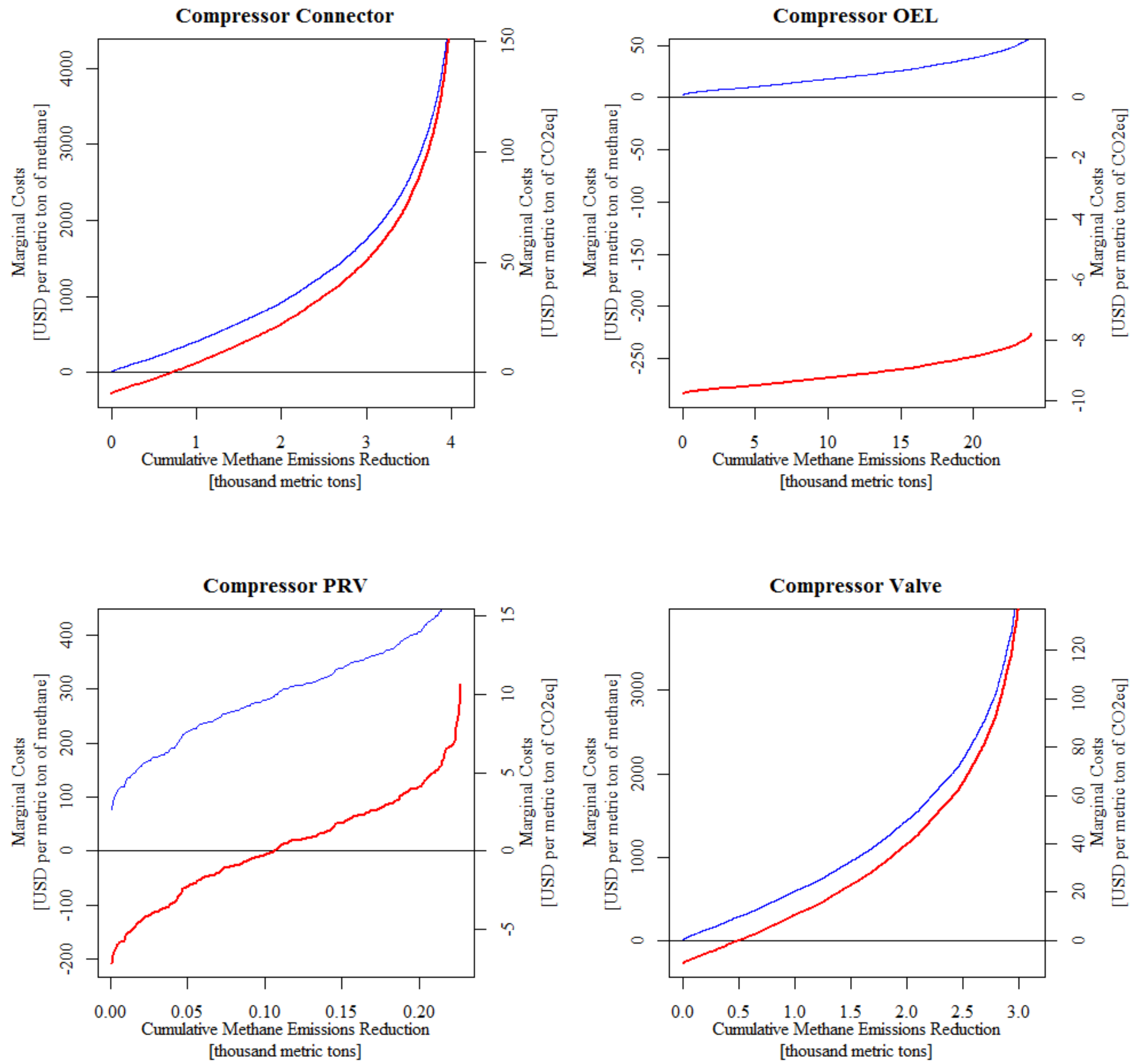

Figure A5 (continued)

A8.3 System-level marginal abatement cost curves

Figure A6 is the marginal abatement cost curve step function, assuming savings and the EIA reference case natural gas price projection.

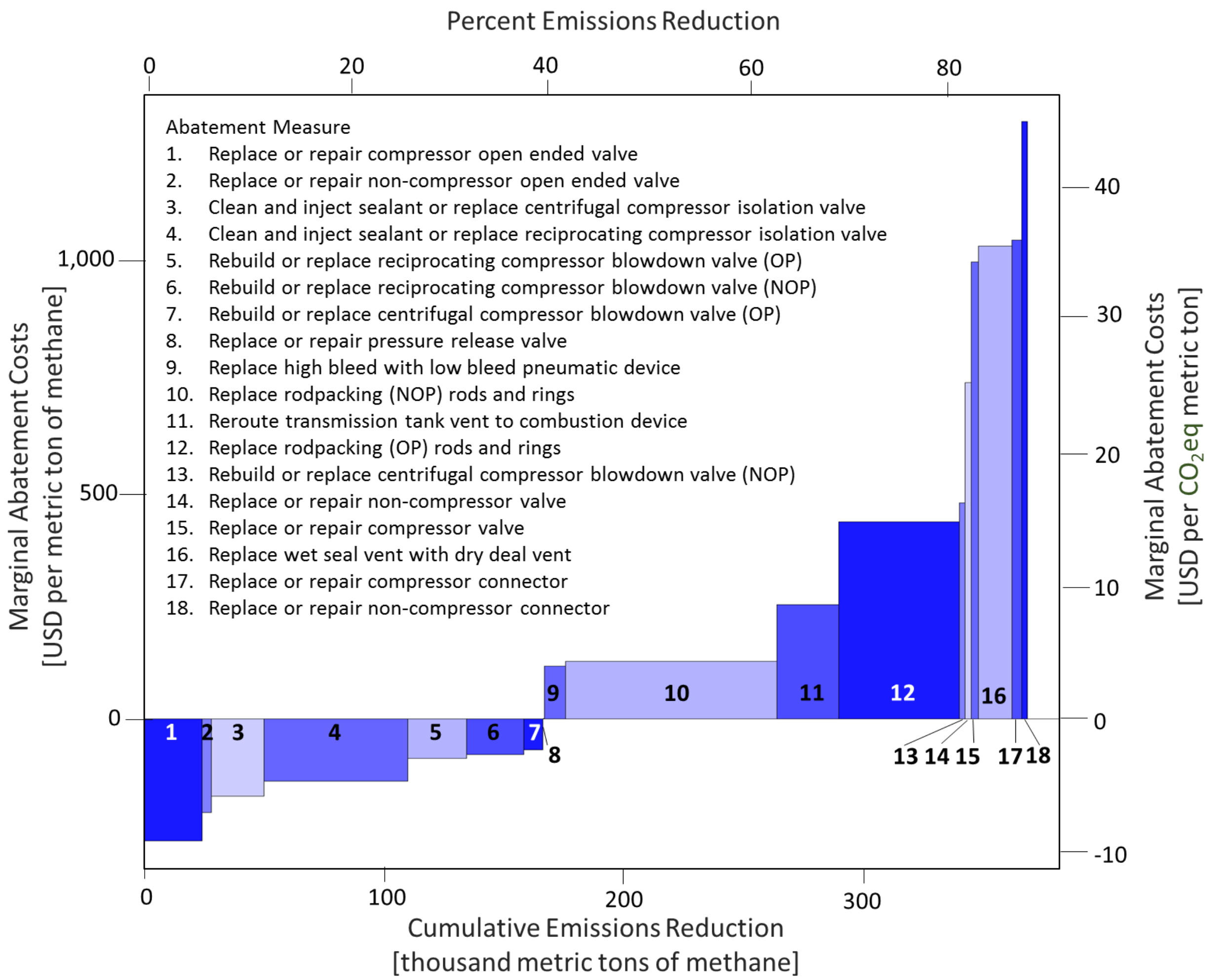

Figure A6. Marginal abatement cost curve based on simulated median emissions and marginal costs, assuming savings based on the EIA reference case natural gas price projection for 2020 ($4.88/MMBTU).

A8.4 Sensitivity analysis

This section describes sensitivity analyses performed for the unconstrained system-wide policy. We parametrically vary key input parameters to the optimization, including social cost of methane, natural gas price, detection technology, discount rate, and abatement efficacy. Figures H7 to H10 depict net benefits, private costs (under performance standard and tax policies), and optimal emission reductions for the single factor sensitivity analyses.

Social Cost of Methane

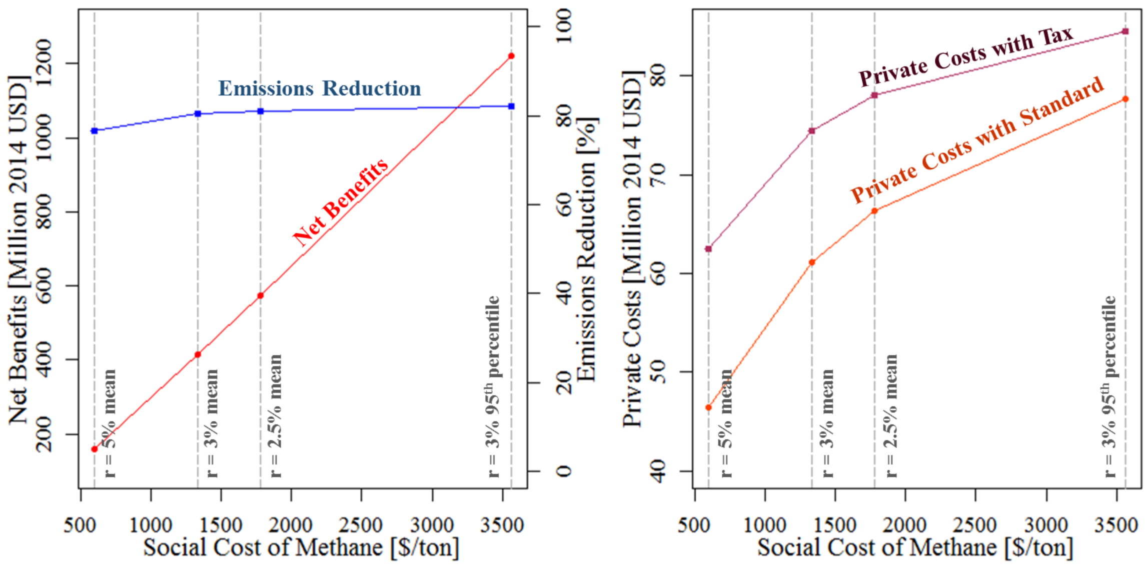

Figure A7. Sensitivity of net benefits, private costs, and optimal emission reductions to social cost of carbon. All estimates are mean values over 100 realizations of the U.S. T&S system.

Natural Gas Price

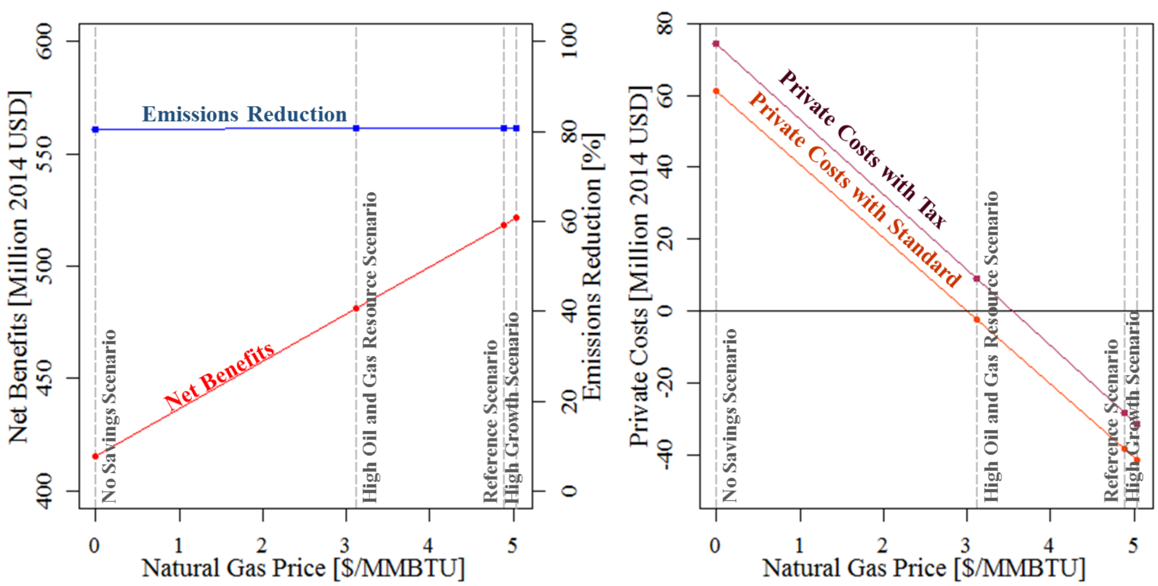

Figure A8. Sensitivity of net benefits, private costs, and optimal emission reductions to natural gas price. All estimates are mean values over 100 realizations of the U.S. T&S system.

Detection Technology

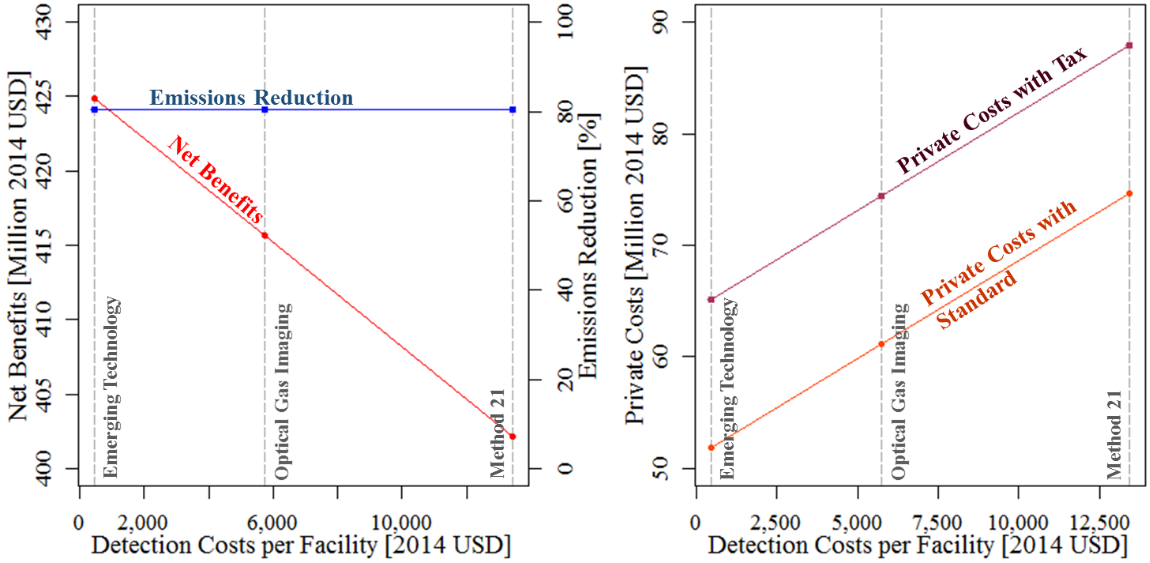

Figure A9. Sensitivity of net benefits, private costs, and optimal emission reductions to detection technology. All estimates are mean values over 100 realizations of the U.S. T&S system.

Discount Rate

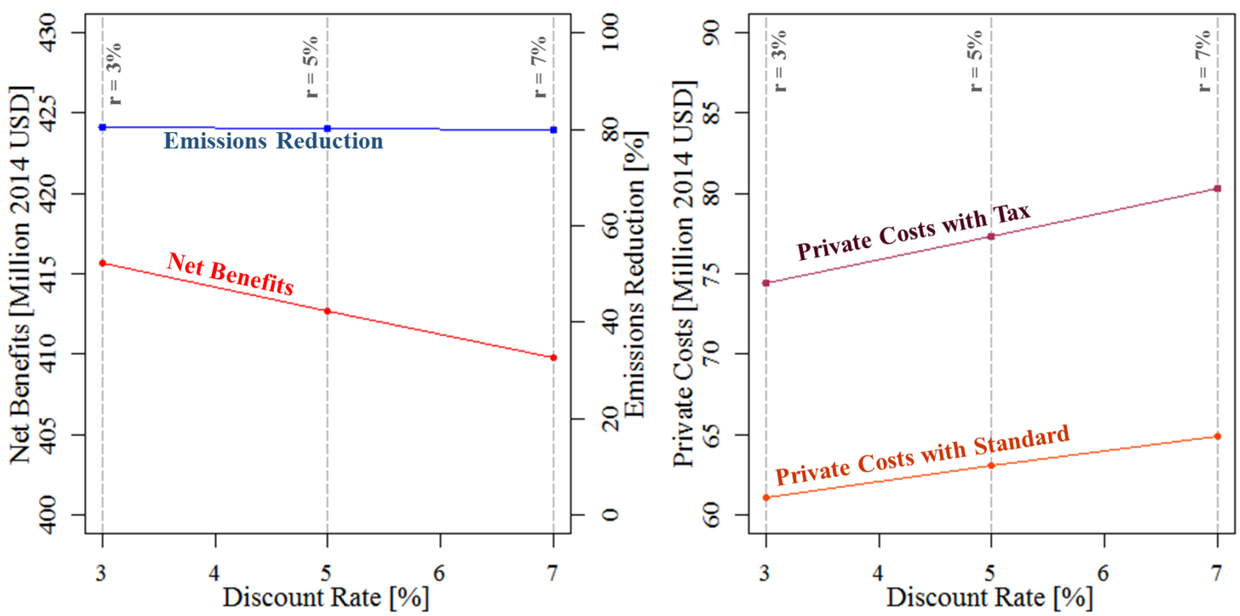

Figure A10. Sensitivity of net benefits, private costs, and optimal emission reductions to discount rate. All estimates are mean values over 100 realizations of the U.S. T&S system.

A8.5 Facility-level comparisons

In this section, we provide of comparison of the policies at the facility-level. Figures A11 to H15 depict differences in annual emissions reductions per facility, percent emissions reduction per facility, private costs per facility, net benefits per facility, and marginal abatement costs per facility.

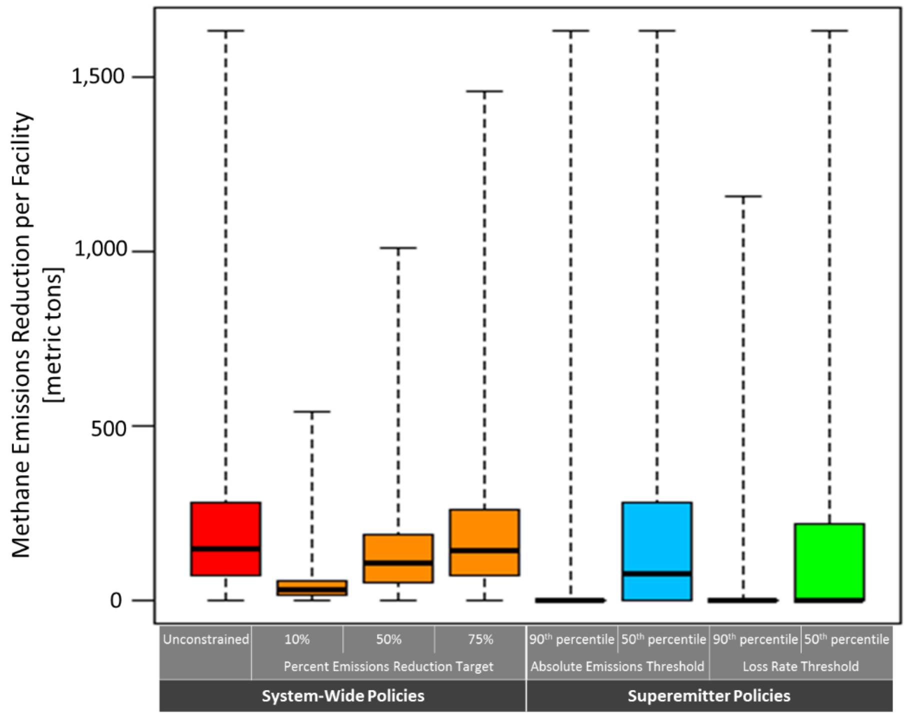

Figure A11. Simulated annual emissions reductions per facility across different policy options.

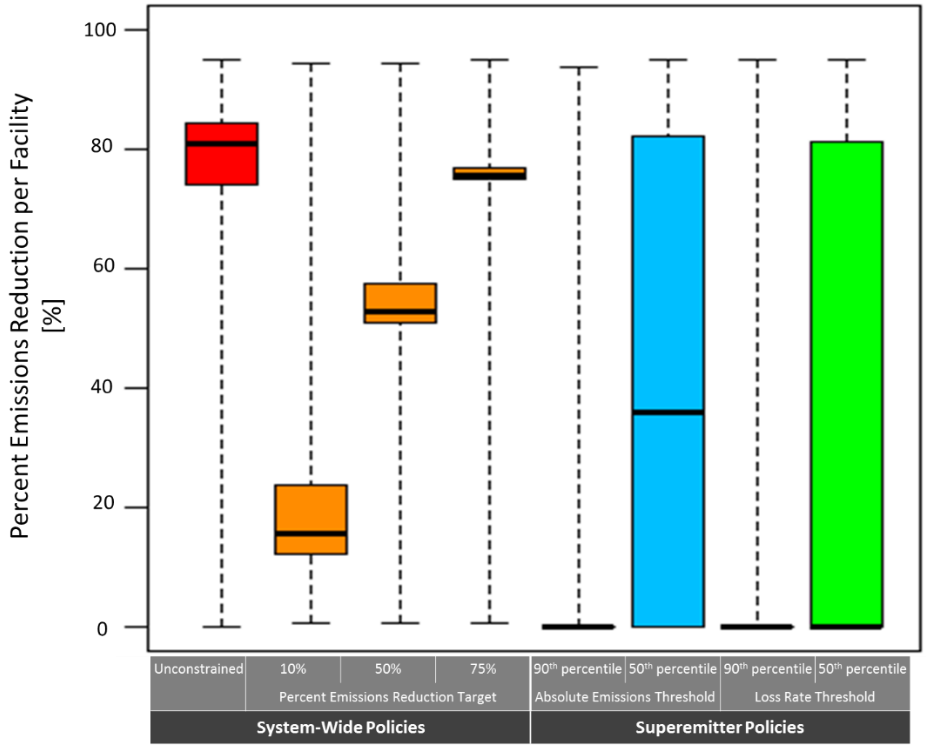

Figure A12. Simulated percent emissions reductions per facility across different policy options.

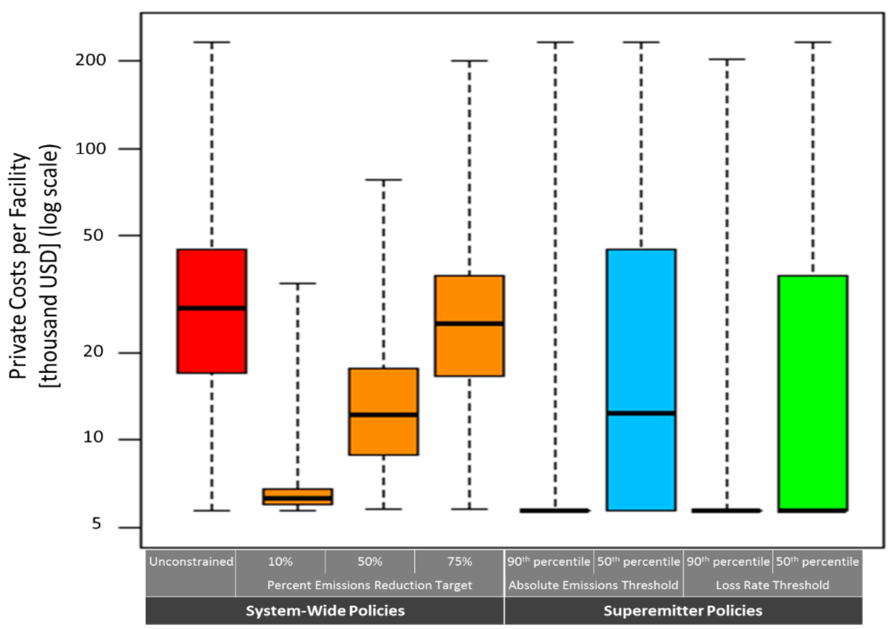

Figure A13. Simulated private costs per facility across different policy options. Private costs account for only detection and abatement costs and do not account for savings.

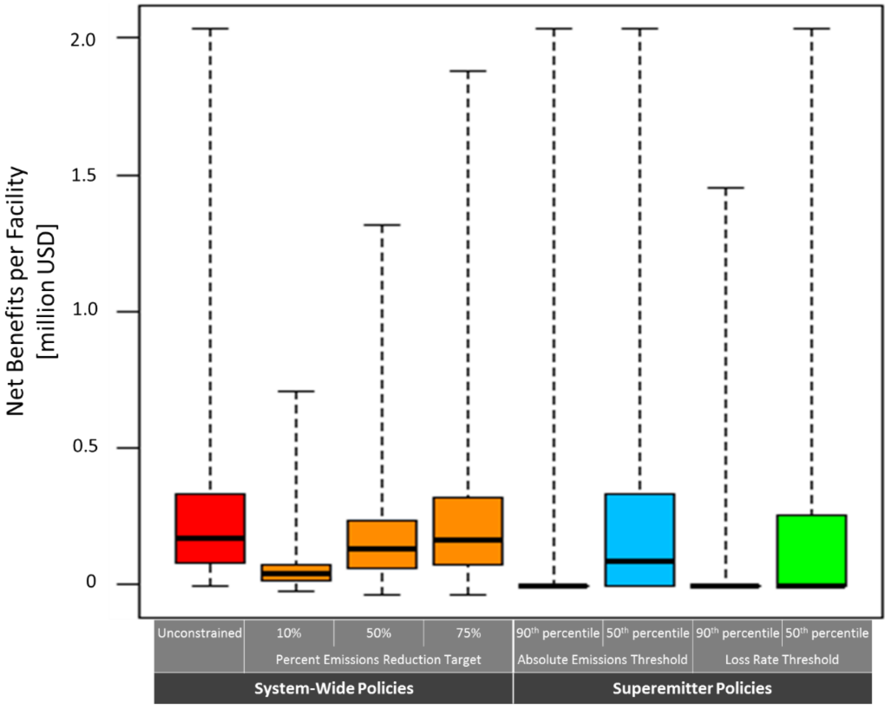

Figure A14. Simulated net benefits per facility across different policy options. Net benefits include the social cost of methane and detection and abatement costs.

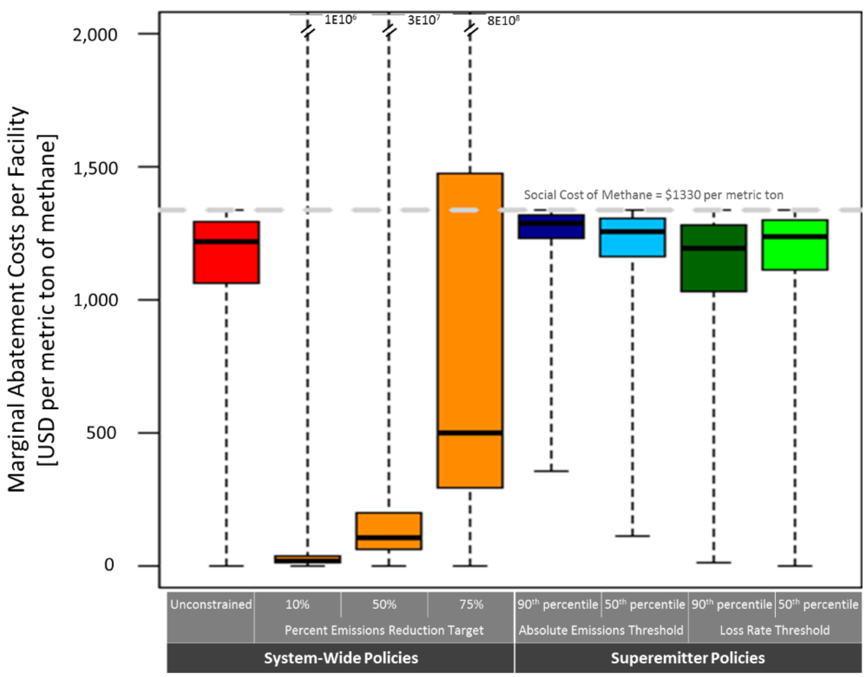

Figure A15. Simulated marginal abatement costs per facility across different policy options. These include only marginal abatement costs for facilities that perform abatement under each policy.

A8.6 Performance standards

Box 1 describes the methodology derived to estimate component-level performance standards that would be needed to achieve the optimal level of abatement.

Box 1: Estimation of performance standards needed to achieve optimal level of abatement

We provide an example for one component type – rod-packing. The plot on the left depicts the marginal net benefits for each component as a function of annual emissions for a single realization of the U.S. T&S system. Blue dots represent the optimal subset of components in the system selected by the model to abate, and black open dots represent components not selected. The performance standard is set at the minimum annual emissions of the optimal subset (as shown in the inset). We then map this annual standard to an emissions factor (EF) standard, as shown in the plot on the right. We repeat this for all realizations of the T&S system. For rod-packing, the mean (and 95th percentile confidence interval) annual standard would need to be 20.0 (0.5 to 69.7) metric tons, or alternatively, the EF would need to be 11.9 (0.0 to 19.8) scfh to achieve the optimal level of abatement for this component type.

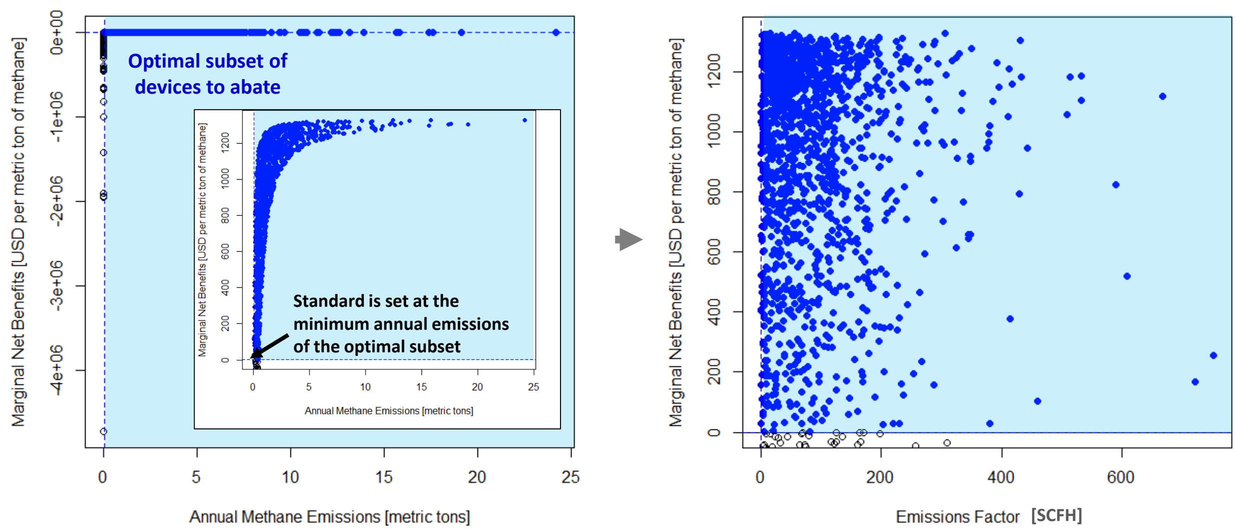

Table A11 presents the performance standards necessary to achieve the optimal level of abatement. We provide results for the unconstrained system-wide policy under base case assumptions and without no savings of marketable natural gas. The performance standards would vary with changing assumptions, given that the optimal subset of devices and marginal net benefits would change.

Table A11. Performance standards needed to achieve optimal level of abatement for each component

Emission Source		**Emission Factor Standard [SCFH]**	**Annual Emissions Standard [SCF]**
Centrifugal Compressor	Blowdown Vent - NOP	2.3 (0.0 to 5.5)	5.2 (0.0 to 22.6)
	Blowdown Vent - OP	1.0 (0.0 to 2.5)	12.7 (0.3 to 46.8)
	Isolation Valve - NOD	1.0 (0.0 to 2.30	73.6 (0.8 to 312)
	Wet Seal Vent - OP	61.0 (39.6 to 83.3)	337 (7.6 to 1.1e3)
Compressor Component Leaks	Connector	0.4 (0.2 to 0.5)	2.4e3 (1.6e3 to 3.0e3)
	Open Ended Line	89.4 (87.1 to 90.9)	7.9e5 (7.7e5 to 8.0e5)
	Pressure Release Valve	3.9 (3.5 to 4.1)	3.42e4 (3.1e4 to 3.6e4)
	Valve	0.6 (0.4 to 0.7)	3.9e3 (2.6e3 to 5.0e3)
Non-Compressor Component Leaks	Connector	0.2 (0.1 to 0.3)	1.0e3 (503 to 1.5e3)
	Open Ended Line	21.3 (20.5 to 22.1)	1.9e5 (1.8e5 to 1.9e5)
	Pressure Release Valve	3.8 (3.5 to 4.1)	3.4e4 (3.1e4 to 3.6e4)
	Valve	1.2 (0.9 to 1.4)	9.4e3 (7.1e3 to 1.1e4)
Pneumatics	High Bleed	1.2 (0.8 to 1.5)	7.6e3 (5.2e3 to 1.0e4)
Reciprocal Compressor	Blowdown Vent - NOP	0.3 (0.0 to 1.5)	4.4 (0.1 to 19.7)
	Blowdown Vent - OP	0.3 (0.0 to 1.6)	4.3 (0.1 to 16.9)
	Isolation Valve - NOD	0.8 (0.0 to 2.5)	10.1 (0.4 to 38.3)
	RodPacking - NOP	11.9 (0.0 to 19.8)	20.0 (0.5 to 69.7)
	RodPacking - OP	9.5 (0.0 to 17.3)	14.4 (0.2 to 50.4)
Transmission Tank	Tank Vent	46.2 (0.0 to 54.1)	2.9e3 (45.9 to 1.3e4)

type. Mean estimates and 95% confidence intervals for 100 realization of the U.S. T&S system are included.

Appendix B: Supplemental Information for Chapter 3

Table of Contents

List of Figures

List of Tables

B1 Scope

This study considers all segments of the natural gas supply chain from preproduction to end use. However, we focus solely on impacts associated with shale gas, a type of unconventional gas in low permeability shale. Unconventional refers to both the resource (i.e., shale gas, methane hydrates, tight sands) and extraction technologies (e.g., hydraulic fracturing, horizontal drilling) used to facilitate production.

We further focus this analysis on impacts associated with natural gas activity within Pennsylvania, Ohio, and West Virginia in the Appalachian basin. The geologically-defined Appalachian basin consists of the Marcellus and Utica, vast plays with combined 2016 proved reserves of 100 trillion cubic feet (tcf) (Figure B1).[1] As of 2016, the Appalachian Basin was the largest natural gas basin in the U.S. with respect to both reserves and production. Shale gas production in the Appalachian basin began in 2004, and the plays were rapidly developed, with production reaching 7.7 tcf in 2016 (Figure B2).

We model impacts from 2004, the year in which the first well was drilled in the Marcellus play, to 2016. Only impacts which directly stem from shale activity within Pennsylvania, Ohio, and West Virginia are modeled; for example, we model emissions from activity within the source states and the corresponding premature mortalities which extend beyond these source state boundaries, but exclude emissions from shale gas exported to interstate and international markets and end use outside of the region.

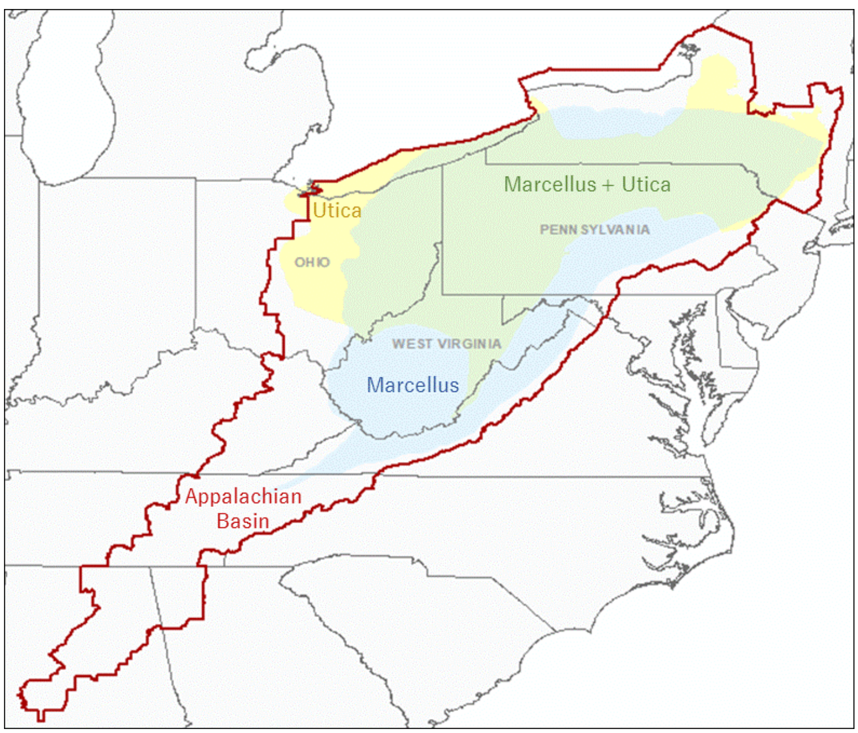

Figure B1. Map of Appalachian basin. The red outline is the extent of the Appalachian basin, as delineated by the U.S. Geological Survey.[2] The shaded blue region is the Marcellus play, the shaded yellow region is the Utica play, and the shaded green region is the intersection of the Marcellus and Utica plays.[3,4]

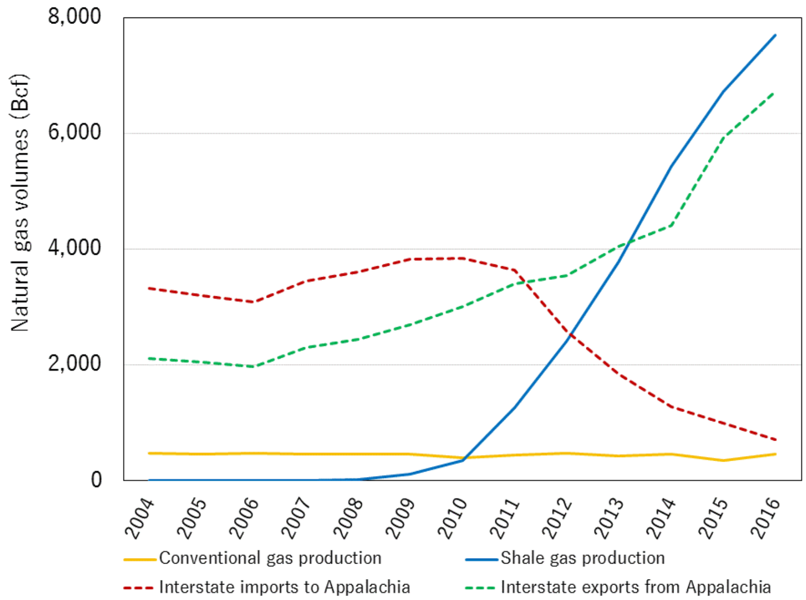

Figure B2. Shale gas withdrawals[5–8], conventional natural gas withdrawals, and export and import volumes[9] from 2004 to 2016 for Appalachia. Import volumes include net interstate receipts to the tristate, and export volumes include net deliveries to outside of the tristate. We exclude import and export volumes for transactions between Pennsylvania, Ohio, and West Virginia.

B2 Natural gas activity

The magnitude, timing, and geographic location of natural gas activity, including producing and spud well counts, production, transmission, distribution, and processed volumes, and upstream, midstream, and end use fuel consumption volumes, are fundamental inputs cross-cutting each impact model. The following gives an overview of natural gas activity data sources, data cleaning, and treatment of missing or misreported data.

B2.1 Production

We compile operator-reported data from the PA Department of Environmental Protection (PA DEP), Ohio Department of Natural Resources (ODNR), and West Virginia Department of Environmental Protection (WV DEP) of production, coordinate location, and other attributes (e.g., producing formation, well direction, spud date, operator, owner, etc.) for each shale well from 2004 to 2016.[5–7] There was notable variation across states and reporting years with respect to production reporting requirements and data quality. Significant data cleaning to ensure internal consistency across states and years, as well as, comparisons to national datasets was performed.

For Pennsylvania, we include natural gas wells designated by the PA DEP as unconventional, which are horizontally or vertically drilled in unconventional formations and require stimulation through hydraulic fracturing. We exclude conventional wells, which produce from conventional formations and are vertically drilled; although conventional wells typically require hydraulic fracturing, they do not require the volume of fluids typically required for unconventional wells. Due to changing reporting requirements, annual production for 2010 cannot be extracted from the reported datasets; thus, we simulate 2010 production by interpolating production per well between 2009 and 2011, and account for the year in which the well was spud. For West Virginia, we include all wells designated in well permits by the WV DEP as horizontally configured and/or as having the Marcellus or Utica formations as the target producing formation. For Ohio, we include all wells designated by the ODNR as horizontal shale and/or Utica wells.

Figure B3 depicts the cumulative production aggregated by county, and Figure B4 depicts the time series of production aggregated by county. Most unconventional natural gas wells drilled in Pennsylvania between 2004 and 2016 produce from the Marcellus shale formation, with production from the Utica shale beginning in 2012 and increasingly thereafter, although many wells produce from or penetrate multiple formations. Production has continued to rise from 2004 to 2016, as shown in Figure B5, and there is high correspondence between the cleaned well-level production data and state-level estimates reported by the Energy Information Agency (EIA).

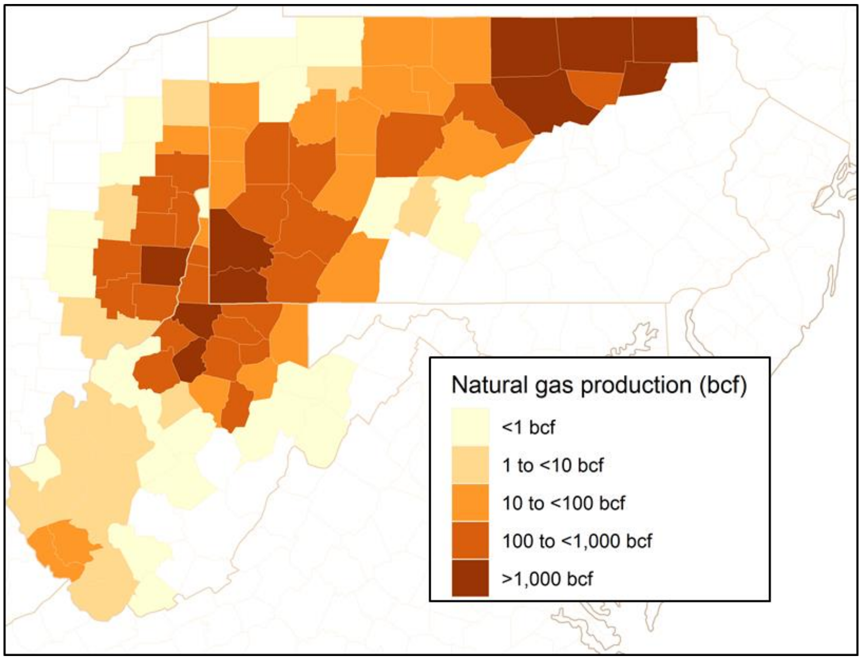

Figure B3. Cumulative production from 2004 to 2016 for each shale gas well. The brown shades indicate the cumulative production aggregated at the county level.

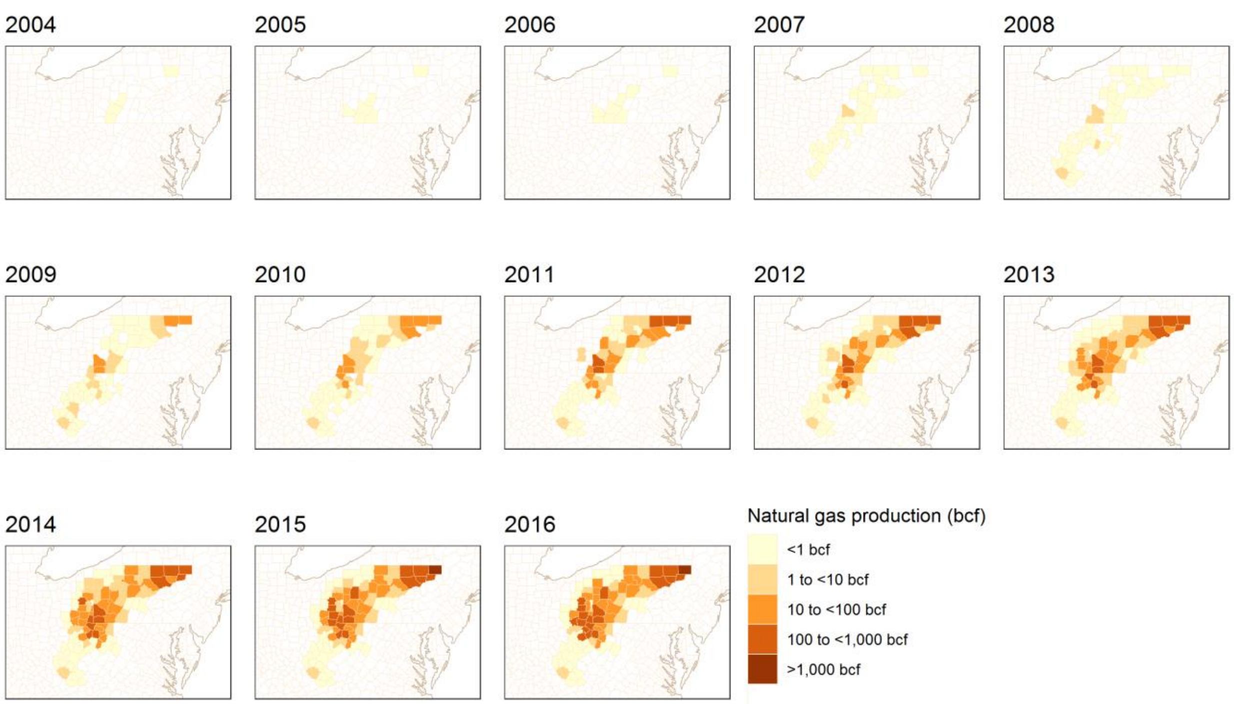

Figure B4. Maps of annual production from 2004 to 2016. The brown shades indicate the production aggregated at the county level.

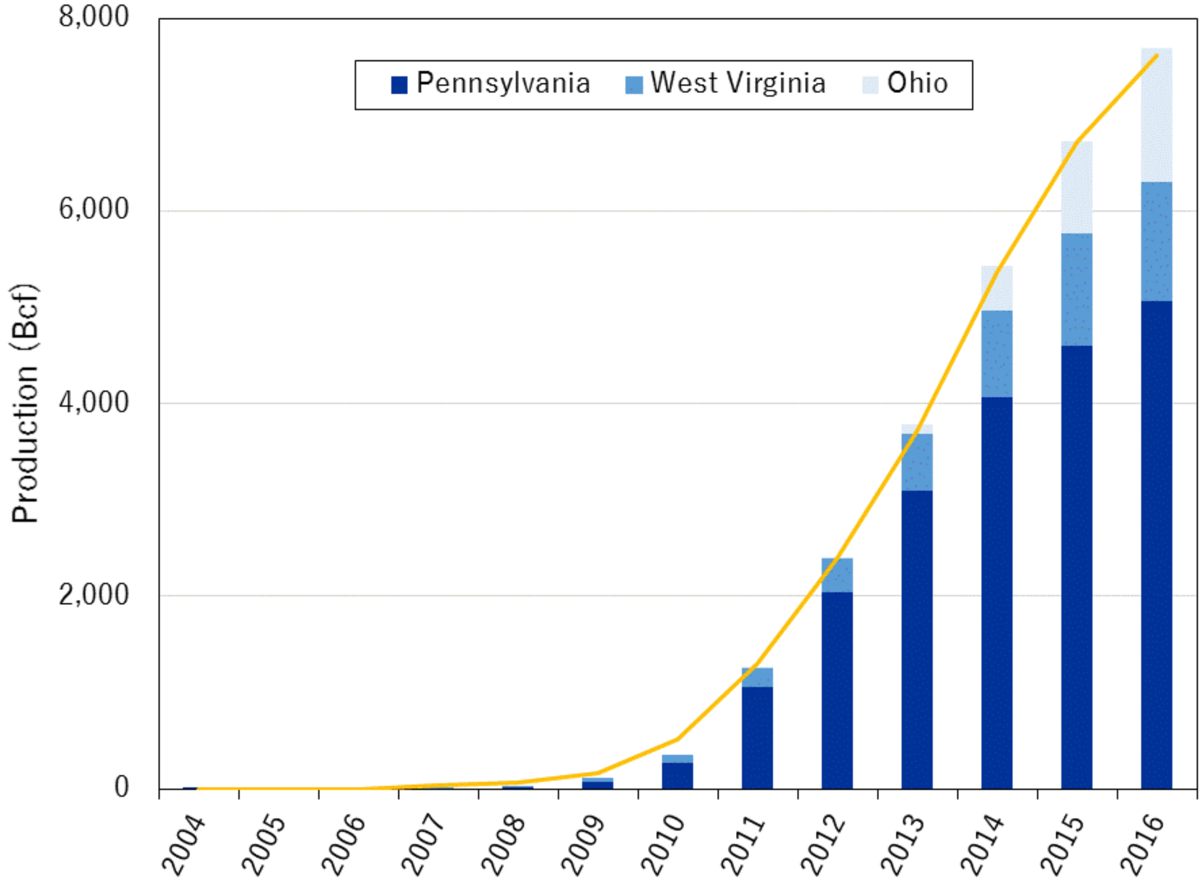

Figure B5. Annual shale gas production from 2004 to 2016 for each state. Bars represent cleaned, well-level production data aggregated by state.[5–7] The solid yellow line depicts shale production aggregated across Pennsylvania, Ohio, and West Virginia and reported at the state level by the EIA.[8]

B2.2 Producing wells

We derive spatially-resolved producing well counts based on the cleaned datasets described in section B2.1. We include all wells that reported production in a given year. Producing well counts have continued to rise from 2004 to 2016, as shown in Figure B6.

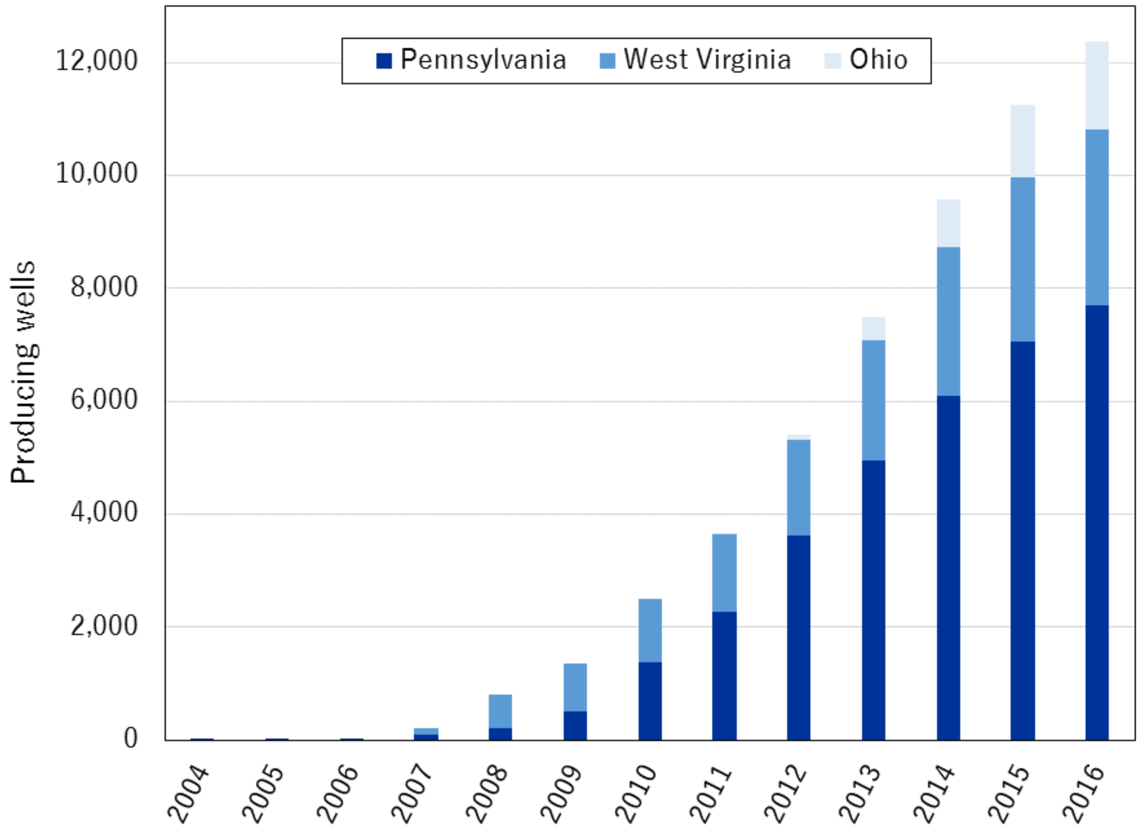

Figure B6. Annual shale gas producing wells from 2004 to 2016 for each state.

B2.3 Spud wells

We derive spatially-resolved spud well counts based on the cleaned datasets described in section B2.1. The number of wells spud in a given year is estimated based on the reported or simulated year in which drilling commences. PA DEP and ODNR publish the spud date for most wells, while WV DEP does not. For wells without a reported spud year, we assume that the spud year is the year prior to when production is first reported. To test the validity of this assumption, we compare the reported spud year to the year of first production for the subset of wells where both data exist, finding that there is an average lag of a year between spudding and start of production, as shown in the Table B1 and Figure B7. We also note that some reported spud years for Pennsylvania are several years prior to 2004 when shale gas production began; this is possibly a reporting anomaly in which an unconventional well was re-spud in the same location as an existing conventional well.

Figure B8 depicts the annual number of spud wells. The first well was spud in Pennsylvania in 2004, while the first reported spud wells in West Virginia and Ohio were a few years later. The aggregate number of spud wells peaked in 2013, after initial rapid purchases and development of leases.

Table B1. Comparison of first production year and reported spud year.

State	# of producing wells	# of wells with reported spud date	Difference between first production year and reported spud year
Pennsylvania	7962	7919	mean=1.15 sd=1.02 median=1
Ohio	1917	1623	mean=0.92 sd=0.71 median=1

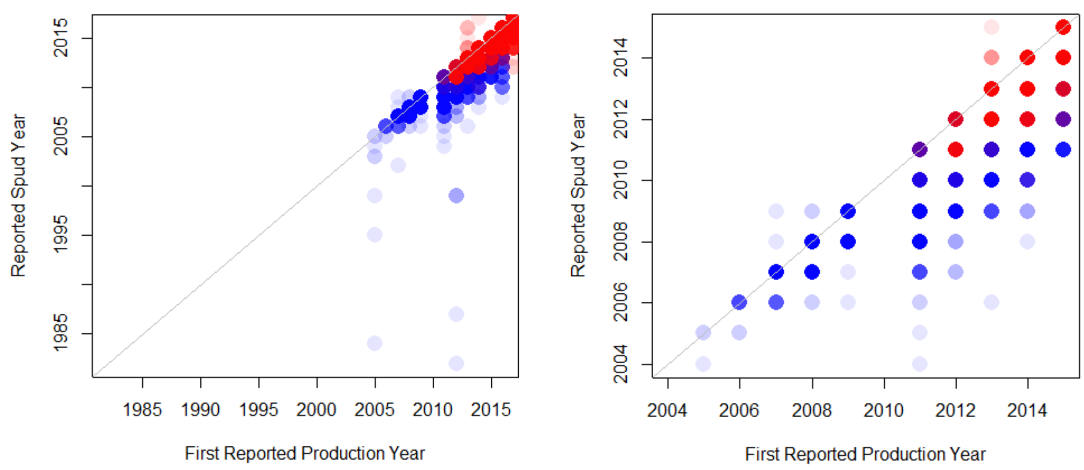

Figure B7. First reported production year versus reported spud year. Blue dots are observations for Pennsylvania and red dots are observations for Ohio. The opacity of the dots is indicative of the number of observations. The grey line represents exact correspondence between first reported production year and reported spud year.

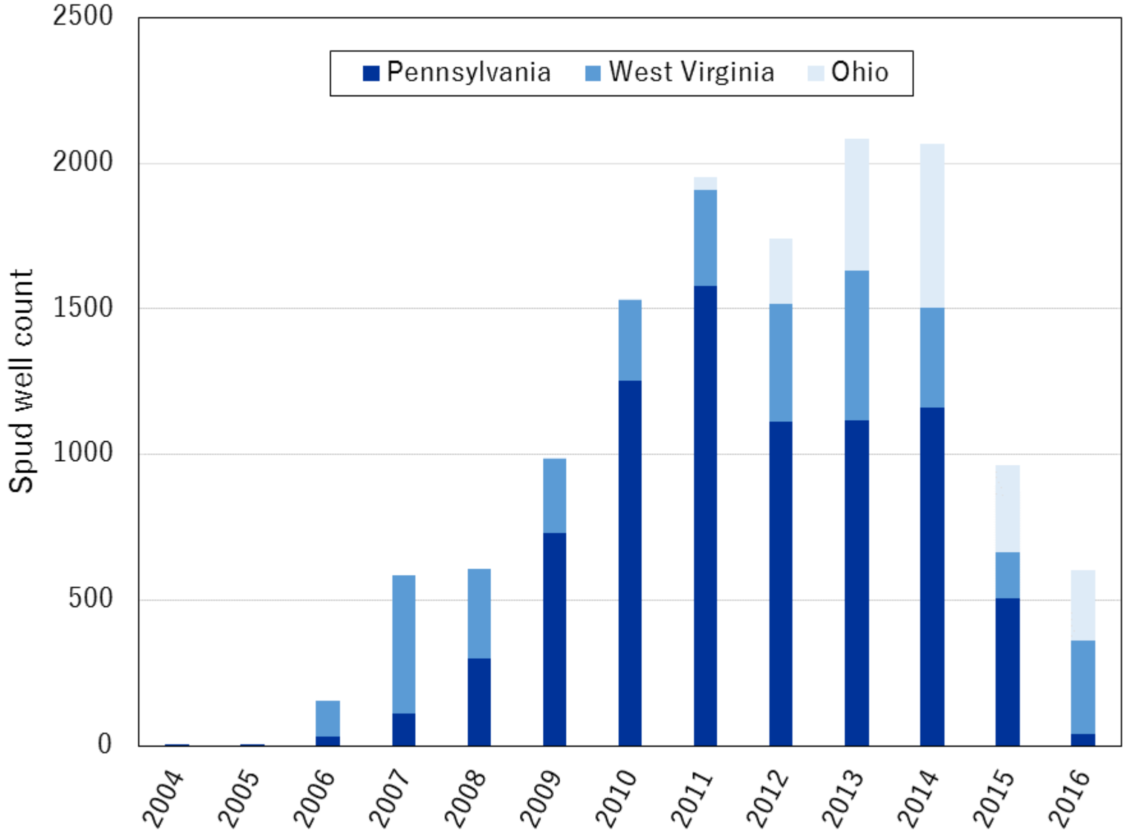

Figure B8. Annual shale gas spud wells from 2004 to 2016 for each state.

B2.4 Midstream and end use volumes

Fuel consumption volumes from midstream and end use activity, a key input for modeling climate change impacts, are depicted in Figure B9. Fuel consumption, transmission, and distribution volumes have increased slightly from 2004 to 2016, while processing volumes have increased 18-fold.

As noted in section B1, we model impacts from shale gas only, excluding conventional natural gas. However, attributing air quality and climate change impacts from midstream and end use segments to shale production is challenging. For example, the volume of natural gas that enters a processing plant is indistinguishably derived from conventional or unconventional production sources. Moreover, flows of natural gas between segments and across state boundaries are inconsistent Figure B2. As a proxy, we prorate reported natural gas volumes (or emissions), based on the percentage of shale gas out of total natural gas production, as provided in Figure B10.

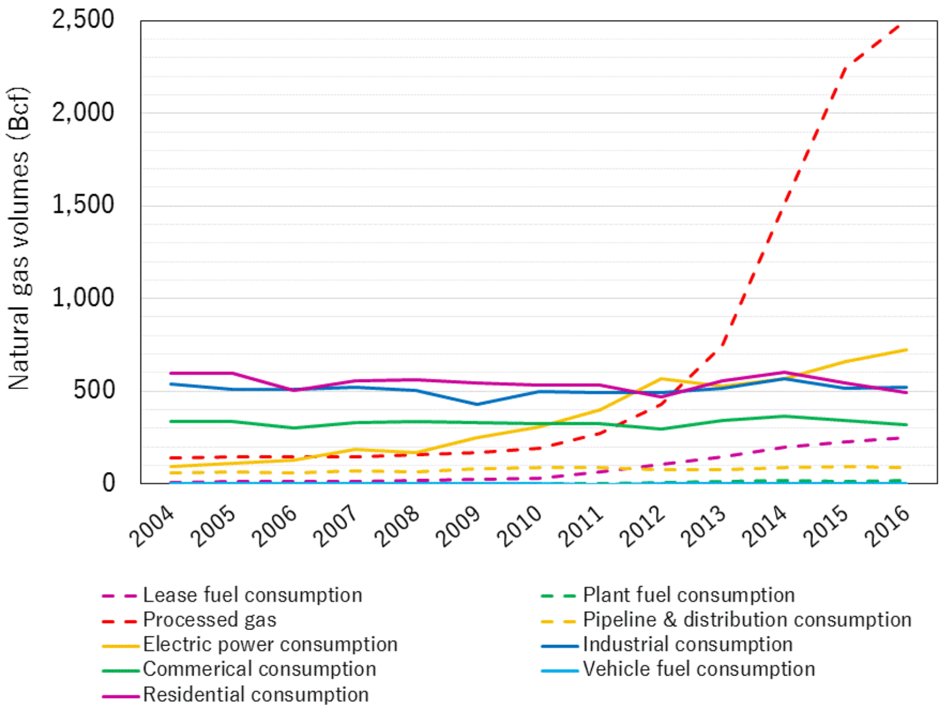

Figure B9. Midstream and end use natural gas volumes from 2004 to 2016 for Pennsylvania, Ohio, and West Virginia. Processed gas[10], lease fuel consumption[11], plant fuel consumption[12], pipeline and distribution consumption[13], and electric power, industrial, commercial, residential, and vehicle consumption[14–18] within Pennsylvania, Ohio, and West Virginia.

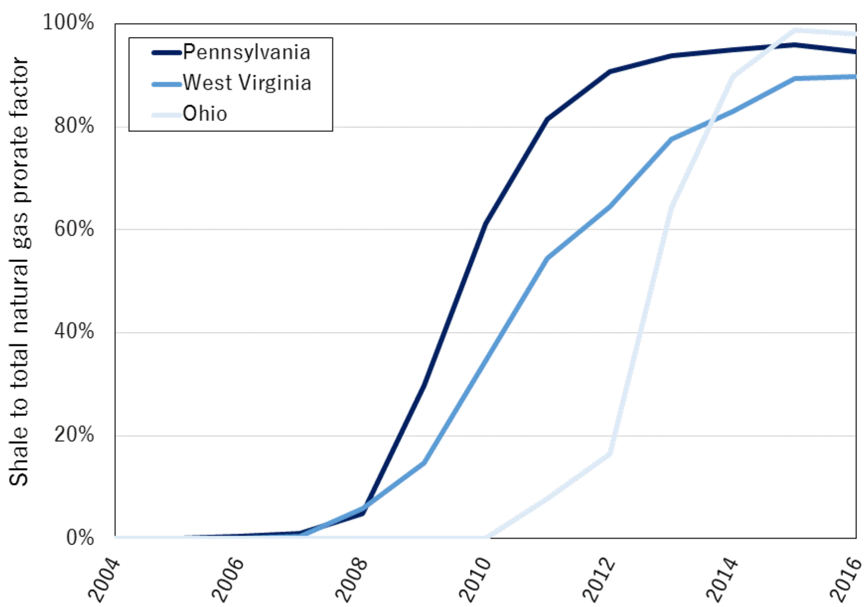

Figure B10. Prorate factor of shale to total natural gas production for Pennsylvania, Ohio, and West Virginia from 2004 to 2016[5–8].

B3 Air quality model

B3.1 Methods

We model spatially-resolved emissions of nitrogen oxides (NOX), volatile organic compounds (VOCs), and fine particulate matter (PM2.5) and health impacts associated with shale gas activities across the supply chain from preproduction to end use from 2004 to 2016. We estimate health impacts in terms of premature mortalities and monetized mortalities associated with NOx, PM2.5, and VOCs using two source-receptor reduced complexity models (RCMs), Air Pollution Emission Experiments and Policy (Version 3) (AP3) and Air Pollution Social Cost Accounting (APSCA) model. We also use another RCM, Intervention Model for Air Pollution (InMAP); although we do not use a receptor-resolved version of the model. To account for major sources of uncertainty, we perform sensitivity analyses of key inputs, such as the dose-response relationship and the value of statistical life, and develop process-level emission ranges for upstream and midstream segments of the supply chain.

B3.1.1 Emissions model

In the following sections, we describe regional and unit-level emissions model formulations and assumptions for each segment. Table B2 is a summary of emissions modeling input parameters, including definitions, values, and data sources.

The inclusion of and detail in which we model each segment of and process within the natural gas sector varies based on the availability and quality of data, and the relative emissions contribution of each source. While we include many significant emission sources, there are several sources and species which we do not model. We limit our analysis to emissions within Pennsylvania, Ohio, and West Virginia, and exclude sources from unconventional natural gas export to interstate and international markets and end use outside of the tristate. In addition, we do not model emissions from the following sources: production pneumatic pumps, transmission and storage pneumatics, gathering station fugitives, transmission pipelines, and distribution service lines and mains.

We use a combination of parametric and probabilistic methods to characterize uncertainty in upstream emissions, focusing on representing uncertainty of systems-level emissions rather than unit-level variation, We derive deterministic estimates of midstream and end use emissions.

We account for time-varying parameters (where practicable), such as changing regulation, upstream practices, and activity factors. With respect to regulation, we account for the phase-in of Tier 2 and Tier 4 standards for non-road diesel engines, New Source Performance Standards (NSPS) Subpart JJJJ Standards of Performance for Stationary Spark Ignition Internal Combustion Engines, and NSPS Subpart OOOO/OOOOa Oil and Natural Gas Sector Emission Standards for New, Reconstructed, and Modified Sources. With respect to midstream and end use segments, we use reported facility-, plant-, or county-level emissions, which we assume inherently incorporate time-varying regulatory, efficiency, and activity factors.

Coordinate locations for stationary sources, including wells, compressor stations, and electric generating units, and county-level location information for residential, commercial, and industrial sources, are based on publicly reported location data. We assume that gathering and trucking are co-located within the county and grid cell that wells are located. We aggregate emissions at county and 36 x 36 km grid cell resolutions, which are the resolutions of the marginal damage factors from the AP3 and APSCA, respectively.

For modeling the preproduction, production, and gathering segments, we explicitly use shale gas activity factors. To attribute emissions from midstream and end use segments to shale production, we use a prorate factor relative to total natural gas production (refer to section B2.4).

Table B2. Emissions modeling parameter values, definitions, and data sources.

Parameter	Parameter Definition	Parameter Value	Units	Source
$CF_{completion}$	emission control factor	VOC: 95	%	19
$CF_{condensate}$	emission control factor	VOC: 95	%	20
$CF_{drilling,i}$	emission control factor[a]	NO_x: Triangular (10,95,30) $PM_{2.5}$: Triangular (60,81,97) VOC: Triangular (60,81,97)	%	21,22
$CF_{fracking,i}$	emission control factors[a]	NO_x: Triangular (10,95,30) $PM_{2.5}$: Triangular (60,81,97) VOC: Triangular (60,81,97)	%	21,22
$CF_{wellheadcompressor,i}$	emission control factors[a]	NO_x: Triangular (15,50,95) VOC: Triangular (30,60,95)	%	21
$DT_{drilling,t}$	time to drill a single well	Triangular (14,30,35)	days	21
$DPW_{trucking}$	distance between public water source and well site	Uniform (4.9, 27.3)	mi	23
$DSW_{trucking}$	distance between surface water source and well site	Uniform (0, 9.9)	mi	23
$DWW_{trucking,t}$	distance between well site and wastewater site	Time-variant (see Figure B12)	mi	24
$EF_{completion,uncontrolled}$	VOC emissions factor for well completion	Empirical (see Figure B11)	metric ton/well	21
$ET_{drilling}$	percentage of time the drilling equipment operates	Triangular (20,50,100)	%	21,25
$F_{condensate}$	VOC emissions per volume of condensate production without emission controls	Empirical (see Figure B11)	g/bbl	26
$F_{drilling,i}$	emissions factor for all drilling rig engines (in base year of 2009)[b]	Empirical (see Figure B11)	g/hp-hr	21
$F_{fracking.i}$	emissions factor for pump engine for fracturing	Empirical (see Figure B11)	g/hp-hr	21
$F_{fugitives,c}$	fugitive TOC emission factor for each component types c	valves: 4.5 connectors: 0.2 OEL: 2 flanges: 0.39	g/hr	21
$F_{heater,i}$	emissions factor per heater	NO_X: Triangular (0.022,0.045,0.091) VOC: Triangular (0.0013,0.005,0.003)	g/scf	21
$F_{trucking,i}$	emission factor per mile of heavy duty diesel trucks	Empirical (see Figure B11)	g/mi	21
$F_{pneumatics}$	VOC emissions per pneumatic device per year[b]	Empirical (see Figure B11)	g/device/yr	27
$F_{wellheadcompressor,i}$	emissions factor for wellhead compressor engines	Empirical (see Figure B11)	g/hp-hr	21
$FC_{completion,t}$	fraction of well completions in which emission controls are used	Time-variant points	-/-	28,29
$FC_{wellheadcompressor,t}$	fraction of wells with compressors	3	%	

Table B2 (continued). Emissions modeling parameter values, definitions, and data sources.

Parameter	Parameter Definition	Parameter Value	Units	Source
$FT_{drilling,t}$	cumulative fleet turnover fraction (relative to base year of 2009)[c]	Time-variant points (see Figure B12)	%	30–32
$FT_{fracking,t}$	cumulative fleet turnover fraction (relative to base year of 2009)[d]	Time-variant points (see Figure B12)	%	30–32
HHV_{heater}	heating value of natural gas	1000	BTU/cf	21
$HP_{compressorstation}$	horsepower per unit of production	Uniform(0.125,0.15)	hp/BCF	21
$HP_{drilling}$	total horsepower of all engines on the drilling rig	Triangular (2000,7000,4260)	hp	21
$HP_{fracking}$	total horsepower of pump engine for fracturing	Triangular (35,000, 40,000, 45,000)	hp-hr	21
$HP_{wellheadcompressor}$	total horsepower of wellhead compressor engine	Triangular (30, 101, 242)	hp-hr	21,25
$LF_{compressorstation}$	average load factor of compressor station engine	Uniform(40,80)	%	21
$LF_{drilling}$	average load factor for all engines on drilling rig	Triangular (26,56,90)	%	21,33
$LF_{fracking}$	average load factor of pump engine for fracturing	0.5	%	21,34
$LF_{wellheadcompressor}$	average load factor of wellhead compressor engine	Empirical	%	21
$N_{fugitives,c}$	number of components of type c per well	valves: 15 connectors: 43 OEL: 5 flanges: 25	components/well	25,35
N_{heater}	number of heaters per well	Triangular (0,0.63,1.1)	heaters/well	25
$N_{pneumatics}$	total number of pneumatic device per well	Empirical	devices/well	27
$P_{condensate,site,s,j,t}$	condensate production per well per year	Empirical (see Figure B11)	bbl/well/yr	
$PPW_{trucking}$	percentage of water sourced (excluding reused/recycled water) from public water supplies	0.2	%	23
$PSW_{trucking}$	percentage of water sourced (excluding reused/recycled water) from surface water supplies	0.8	%	23
Q_{heater}	heater throughput	10^6	BTU/hr	21
Q_{heater}	heater throughput	10^6	BTU/hr	21
$R_{trucking,t}$	rate of reuse/recycling of wastewater	Time-variant points (see Figure B12)	%	24
$RE_{commercial,state,i,j,t}$	reported (or simulated) county level emissions from commercial end use	Time-varying	metric tons	36
$RE_{distribution,state\ i,j,t}$	reported (or simulated) emissions from distribution compressor stations	Time-varying	metric tons	36
$RE_{electric,state\ i,j,t}$	reported (or simulated) emissions from electric utility facilities	Time-varying	metric tons	36–38

Table B2 (continued). Emissions modeling parameter values, definitions, and data sources.

Parameter	Parameter Definition	Parameter Value	Units	Source
$RE_{insutrial,state\ i,j,t}$	reported (or simulated) county level emissions from industrial end use	Time-varying	metric tons	[36]
$RE_{processing,state\ i,j,t}$	reported (or simulated) emissions from processing facilities	Time-varying	metric tons	[36]
$RE_{residential,state\ i,j,t}$	reported (or simulated) emissions from residential end use	Time-varying	metric tons	[36]
$RE_{transmission,state\ i,j,t}$	reported (or simulated) emissions from transmission compressor stations	Time-varying	metric tons	[36]
$S_{fracking.t}$	number stages to fracture well	Triangular (4,15,33)	stages	[21]
$T_{compressorstation}$	number of hours engine operates per day	24	hrs	[21]
$T_{fugitives}$	annual operating hours per well	8760	hrs/year	[35]
T_{heater}	number of hours heater operates per year	Triangular (126, 2982,4601)	hrs/yr	[25]
$T_{wellheadcompressor}$	number of hours engine operates per year	8760	hrs/yr	[21]
$U_{state,t}$	fraction of total unconventional production out of total production	Time-varying (ranges from 0 to 99%)	-/-	[5–8]
$TC_{trucking}$	truck fluid capacity	4620	gal	[39]
$UW_{trucking}$	water usage per well	Triangular (1.00, 4.05, 6.97)	million gal/well	[23]
$UWW_{trucking}$	wastewater produced per well	Triangular (1.22, 1.38, 1.64)	million gal/well	[40]
$VF_{fugitives}$	VOC fraction of natural gas	0.02	-/-	[21]
$W_{producing,j,t}$	number of spud wells	Time-varying (refer to section B2)	wells	[5–7]
$W_{spud,j,t}$	number of spud wells	Time-varying (refer to section B2)	wells	[5–7]

a We assume ignition timing retard and selective catalytic reduction for NOx, diesel particulate filters for PM2.5, diesel oxidation catalysts for VOCs.

b To develop an empirical VOC emission factor distribution, we use whole gas emission factors and mole fractions of propane, butane, and higher hydrocarbons for a subset of 53 wells in the Appalachian region, as reported in Allen et al. (2014). While the Allen et al. study measures emissions in multiple basins, we subset the reported data for those devices measured within the Appalachian region because of that finding that there is great variability in regional emissions resulting from differences in controller type (e.g., continuous and intermittent venting), frequency of actuation (e.g., more actuation in wet areas), and different applications (e.g., separators, process heaters). We convert reported emission factors in terms of scf per hour to grams per year, under standard conditions (14.7 psia and 70°F) and assuming the pneumatic devices are operating 8760 hours per year.

c We assume the following to derive the fleet turnover curve: a normally distributed scrappage curve[32], average load factor of 0.43 for diesel bore / drill rigs[30], annual activity of 466 hours per year for diesel bore / drill rigs[30], an average life of large diesel engines of 7000 hours[30], and a 3.7% growth rate for diesel industrial engines[31].

d We assume the following to derive the fleet turnover curve: a normally distributed scrappage curve[32], average load factor of 0.43 for diesel other oil equipment[30], annual activity of 1231 hours per year for other oil equipment[30], an average life of large diesel engines of 7000 hours[30], and a 3.7% growth rate for diesel industrial engines[31].

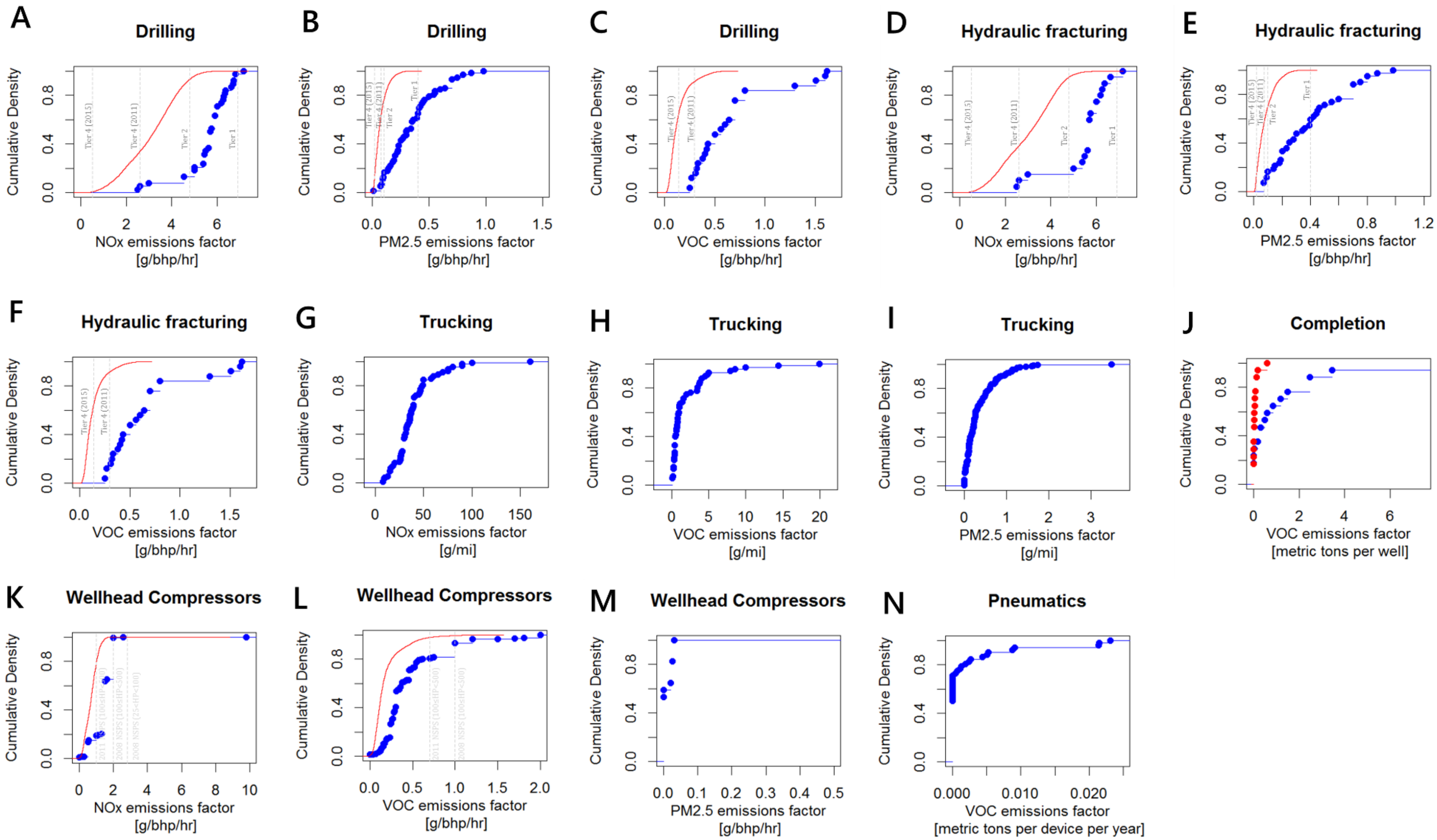

Figure B11. Empirical distributions used in air quality emissions modeling. Drilling emission factor distributions for (A) NO_x, (B) VOC, and (C) $PM_{2.5}$. Hydraulic fracturing emission factor distributions for (D) NO_x, (E) VOC, and (F) $PM_{2.5}$. Trucking emission factor distributions for (G) NO_x, (H) VOC, and (I) $PM_{2.5}$. Completion emission factor distribution for (J) VOC. Wellhead compressor emission factor distributions for (K) NO_x, (L) VOC, and (M) $PM_{2.5}$. Pneumatic devices emission factor distribution for (N) VOC. Blue lines and dots are empirical emission factor distributions (without emission controls). Red lines are emission factor distributions, accounting for emission controls. Grey dashed lines are regulatory standards, including Tier 1, 2, and 4 non-road (hp>750) diesel engine standards and NSPS for spark-ignition natural gas engines.

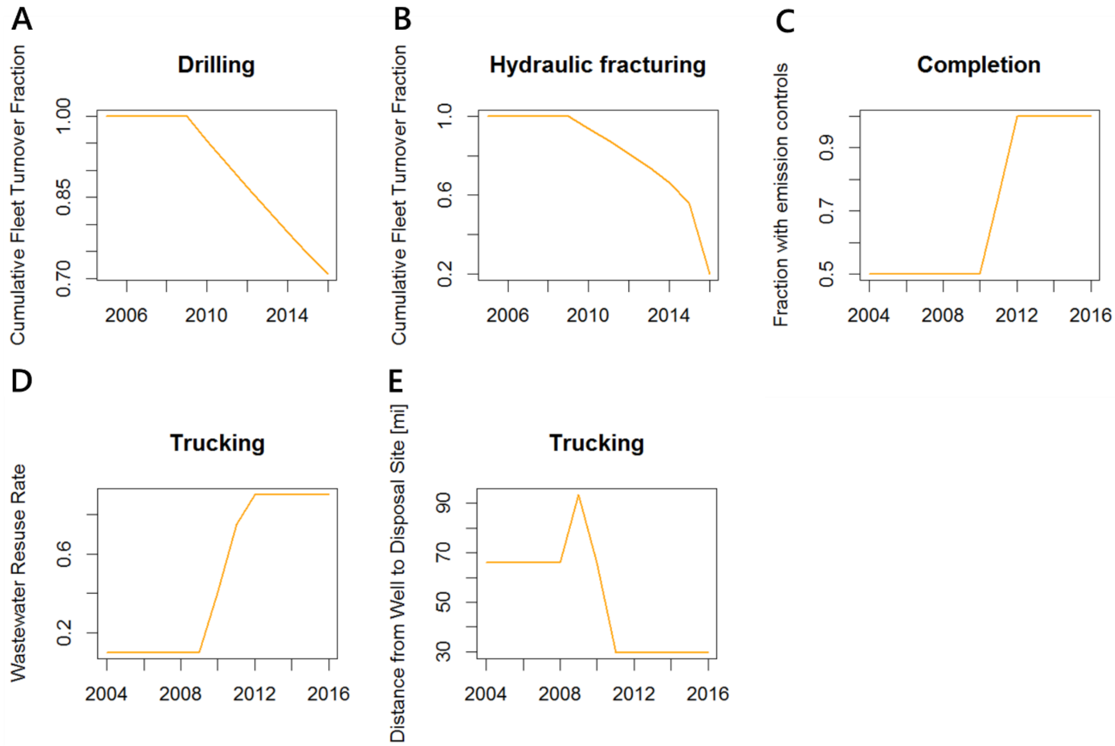

Figure B12. Time-varying parameters used in air quality emissions modeling. (A) Drilling rig fleet turnover based on EPA's nonroad growth and scrappage methodology.[32] (B) Hydraulic fracturing fleet turnover for fracturing pumps based on EPA's nonroad growth and scrappage methodology.[32] (C) Fraction of well completions with emission controls.[28,29] (D) Wastewater reuse rate. (E) Distance from well to disposal site.

B3.1.1.1 Preproduction

We estimate drilling emissions ($E_{drilling,i,j,t}$) for each species i, source location j, and year t, using a modification of the approach used in Bar-Ilan et al. (2008), as follows[25]:

$$E_{drilling,i,j,t} = W_{spud,j,t} \cdot \left[FT_{drilling,t} \cdot EF_{drilling,controlled,i,t} + (1 - FT_{drilling,t}) \cdot EF_{drilling,uncontrolled,i,t}\right] \quad (1)$$

where $W_{spud,j,t}$ is the number of spud wells and $FT_{drilling,t}$ is the fleet turnover fraction. $EF_{drilling,controlled,i,t}$ and $EF_{drilling,uncontrolled,i,t}$ are the emission factors per well for each year when emission controls are or are not employed, respectively, and are given by:

$$EF_{drilling,controlled,i,t} = F_{drilling,i} \cdot (1 - CF_{drilling,i}) \cdot HP_{drilling} \cdot LF_{drilling} \cdot DT_{drilling,t} \cdot ET_{drilling} \quad (2)$$

$$EF_{drilling,uncontrolled,i,t} = F_{drilling,i} \cdot HP_{drilling} \cdot LF_{drilling} \cdot DT_{drilling,t} \cdot ET_{drilling} \quad (3)$$

where $HP_{drilling}$ is the horsepower of drilling rig engines, $LF_{drilling}$ is the load factor of the drilling rig engines, $T_{drilling,t}$ is the drilling time, and $ET_{drilling}$ is the engine-on-time percentage. We use empirical distributions of engine emissions factors ($F_{drilling,i}$) for a base year of 2009, which were compiled in Roy et al. (2014). We assume a fleet turnover fraction ($FT_{drilling,t}$) and apply control factors ($CF_{drilling,i}$), which together reflect the phase-in of Tier 2 and Tier 4 standards for non-road diesel engines. We develop emission factors ($EF_{drilling,controlled,i,t}$ and $EF_{drilling,uncontrolled,i,t}$) representing the uncertainty in the systems-level mean rather than between-unit variability; we use a Monte Carlo simulation approach and find the mean and the upper and lower 95% confidence interval.

We estimate hydraulic fracturing emissions ($E_{fracking,i,j,t}$), using a modification of the approach used in Bar-Ilan et al. (2008), as follows[25]:

$$E_{fracking,i,j,t} = W_{spud,j,t} \cdot \left[FT_{fracking,t} \cdot EF_{fracking,controlled,i,t} + (1 - FT_{fracking,t}) \cdot EF_{fracking,uncontrolled,i,t}\right] \quad (4)$$

where $W_{spud,j,t}$ is the number of spud wells and $FT_{fracking,t}$ is the pump engine fleet turnover fraction. $EF_{fracking,controlled,i,t}$ and $EF_{fracking,uncontrolled,i,t}$ are the emission factors per spud well for each year when emission controls are or are not employed, respectively, and are given by:

$$EF_{fracking,controlled,i,t} = F_{fracking,i} \cdot (1 - CF_{fracking,i}) \cdot HP_{fracking} \cdot LF_{fracking} \cdot S_{fracking.t} \quad (5)$$

$$EF_{fracking,uncontrolled,i,t} = F_{fracking,i} \cdot HP_{fracking} \cdot LF_{fracking} \cdot S_{fracking.t} \quad (6)$$

where $HP_{drilling}$ is the horsepower of drilling rig engines, $LF_{drilling}$ is the load factor of the drilling rig engines, $T_{drilling,t}$ is the drilling time, and $ET_{drilling}$ is the engine-on-time percentage. We use empirical distributions of pump engine emissions factors ($F_{fracking,i}$) for a base year of 2009, which were compiled in Roy et al. (2014). We incorporate time variant elements, including emission regulations. We assume a fleet turnover fraction ($FT_{fracking,t}$) and apply control factors ($CF_{fracking,i}$), which together reflect the phase-in of Tier 2 and Tier 4 standards for non-road diesel engines. We develop emission factors ($EF_{fracking,controlled,i,t}$ and $EF_{fracking,uncontrolled,i,t}$) representing the uncertainty in the systems-level mean rather than between-unit variability; we use a Monte Carlo simulation approach and find the mean and the upper and lower 95% confidence interval.

Emissions from well completions are highly uncertain.[41–43] We estimate VOC emissions from well completion ($E_{completion,i,j,t}$) as follows:

$$E_{completion,i,j,t} = W_{spud,j,t} \cdot \left[FC_{completion,t} \cdot EF_{completion,controlled} + (1 - FC_{completion,t}) \cdot EF_{completion,uncontrolled,i,t}\right] \tag{7}$$

where $W_{spud,j,t}$ is the number of spud wells and $FC_{completion,t}$ is the fraction wells with emission controls. VOC emission factors for wells with ($EF_{completion,controlled}$) and without emission controls ($EF_{completion,uncontrolled}$), where the controlled emission factor is given by:

$$EF_{completion,controlled} = EF_{completion,uncontrolled} \cdot (1 - CF_{completion}) \tag{8}$$

We use an empirical distribution of emissions factors ($EF_{completion,uncontrolled,i,t}$) for a base year of 2009, which were compiled in Roy et al. (2014). We attempt to reflect voluntary adoption of emission controls and rapidly evolving regulation, including the implementation of the NSPS subpart OOOO and National Emissions Standards for Hazardous Air Pollutants (NESHAP) standards, that require reduced emission completions (RECs) for hydraulically fractured wells. We assume a changing fraction of well completions with emission controls ($FC_{completion,t}$); the penetration of emission controls over time is uncertain, although recent emission measurement studies suggest that most wells in Appalachia (for which measurements were taken) have emission controls; in a study conducted in 2012 by Allen et al. (2013), all five measured well completions in Appalachia had emission controls, and in a study conducted in 2014 by Omara et al. (2016), all four measured well completions had emission controls, with three of having installed RECs.[28,29] The emission factors represent the uncertainty in the systems-level mean rather than between-unit variability; we use a Monte Carlo simulation approach and find the mean and the upper and lower 95% confidence interval.

Emissions from trucking, including transport of drilling and fracturing water and wastewater ($E_{trucking,i,j,t}$) are estimated as follows:

$$E_{trucking,i,j,t} = W_{spud,j,t} \cdot EF_{trucking,i,t} \quad (9)$$

where $W_{spud,j,t}$ is the number of spud wells, and $EF_{trucking,i,t}$ is the emission factor, as given by:

$$EF_{trucking,i,t} = \left[UWW_{trucking} \cdot DWW_{trucking,t} + UW_{trucking} \cdot \left(DSW_{trucking} \cdot PSW_{trucking} + DPW_{trucking} \cdot PPW_{trucking}\right) \cdot \left(1 - R_{trucking,t}\right)\right] / TC_{trucking} \cdot 2 \cdot F_{trucking,i} \quad (10)$$

where $WWU_{trucking}$ and $WU_{trucking}$ are the water and wastewater use per well, respectively. $DWW_{trucking,t}$, $DSW_{trucking}$, and $DPW_{trucking}$ are the distance per trip for transporting between the well site and the wastewater disposal, surface water source, and public water source, respectively. $PSW_{trucking}$ and $PPW_{trucking}$ are the percentage of water from surface and public water supplies, respectively. $R_{trucking,t}$ is the percentage of wastewater that is reused and $TC_{trucking}$ is the capacity per truck. We use empirical distributions of emissions factors per mile for heavy duty diesel trucks ($F_{trucking,i}$) compiled in Roy et al. (2014). We exclude trucking from other site operations, given that fluid transport is the most significant source of truck traffic. We incorporate time variant elements reflective of changing wastewater management practices; specifically, we reflect the declining distance traveled to wastewater disposal sites and the increasing portion of wastewater that is reused, resulting from changing regulation and increasing wastewater infrastructure.

B3.1.1.2 Production

Emissions from wellhead compressors ($E_{wellheadcompressor,i,t}$) are estimated as follows:

$$E_{wellheadcompressor,i,t} = W_{producing,j,t} \cdot FC_{wellheadcompressor} \cdot EF_{wellheadcompressor,i,t} \quad (11)$$

where $W_{producing,j,t}$ is the number of producing wells, $FC_{wellheadcompressor}$ is the fraction of wells with compressors, and $EF_{wellheadcompressor,i,t}$ is the emissions factor per well, given by the following set of equations:

$$EF_{wellheadcompressor,i,t\leq 2011} = F_{wellheadcompressor,i} \cdot HP_{wellheadcompressor} \cdot LF_{wellheadcompressor} \cdot T_{wellheadcompressor} \quad (12)$$

$$EF_{wellheadcompressor,i,t\leq 2011} = (1 - CF_{wellheadcompressor,i}) \cdot F_{wellheadcompressor,i} \cdot HP_{wellheadcompressor} \cdot LF_{wellheadcompressor} \cdot T_{wellheadcompressor} \quad (13)$$

where $HP_{wellheadcompressor}$ is the total horsepower of wellhead compressor engines per well, $LF_{wellheadcompressor}$ is the load factor of wellhead compressor engines, $T_{wellheadcompressor}$ is the number of operating hours per year, and $CF_{wellheadcompressor,i}$ is the emission control factor. We modified the approach used in Roy et al. (2014) to incorporate time variant elements, including the fraction of wellheads that have compressors and emission regulations. We use empirical distributions of emissions factors for wellhead compressor engines ($F_{wellheadcompressor,i}$) (for a base year of 2009) compiled in Roy et al. (2014). To estimate emissions in other years, we assume an increasing fraction of wellheads with compressor and apply control factors, reflecting the phase-in of NSPS subpart JJJJ Standards of Performance for Stationary Spark Ignition Internal Combustion Engines for NO_x and VOC. Although wellhead compressors are currently uncommon in the Appalachian Basin, production pressure declines as the field ages, necessitating the use of wellhead compressors. We assume all wellhead compressors after 2011 have emission controls, given that wellhead compressor engines have short lifespans and there is low penetration of wellhead compressors (<1%) prior to 2011.

Emissions from condensate volatilization ($E_{condensate,j,t}$) from tanks are estimated as follows:

$$E_{condensate,j,t} = \sum_{site} P_{condensate,site,s,j,t} \cdot EF_{condensate,s} \quad (14)$$

where $P_{condensate,site,s,j,t}$ is the well site-specific condensate production for a given spud year s, and $EF_{condensate,s,t}$ is the emissions factor per unit condensate production, given by the following set of equations:

$$EF_{condensate,s<2011} = F_{condensate} \quad (15)$$

$$EF_{condensate,s\geq 2011} = F_{condensate} \cdot (1 - CF_{condensate}) \quad (16)$$

where $F_{condensate}$ is the VOC emission factor per unit condensate production without controls, and $CF_{condensate}$ is the emission control factor. We incorporate time variant elements, namely NSPS subpart OOOO, promulgated in 2012, for VOC emissions from storage vessels at production sites; the rule requires 95% control for all storage vessels emitting at least 6 tpy of VOC, which were constructed, modified, or reconstructed after August 2011. We use empirical distributions of emissions factors, representing uncontrolled condensate tank emissions compiled in Roy et al. (2014). We assume that all wells spud prior to 2011 are uncontrolled, and those spud in 2011 or later are controlled and comply with the NSPS 95% control factor.

Emissions from pneumatic devices ($E_{pneumatics,i,j,t}$) are estimated as follows:

$$E_{pneumatics,i,j,t} = W_{producing,j,t} \cdot EF_{pneumatics,i} \quad (17)$$

where $W_{producing,j,t}$ is the number of producing wells, and $EF_{pneumatics,i}$ is the VOC emissions per well, as given by:

$$EF_{pneumatics,i,j,t} = F_{pneumatics,i} \cdot N_{pneumatics} \quad (18)$$

We base the emission factor per device ($F_{pneumatics,i}$) and number of devices per well ($N_{pneumatics}$) distributions on measurements from a study by Allen et al. (2014). The NSPS subpart OOOOa for VOC emissions from pneumatic controllers at production sites were promulgated in 2012; the rule sets a whole gas bleed rate limit of 6 scf/h for continuous bleed, natural gas-driven pneumatic controllers, which commenced construction after August 2011, and are located between the wellhead and the point at which the gas enters the transmission and storage segment. We do not account for any time-varying parameters, given field data for the Appalachian region reported in Allen et al. (2014) indicate most whole gas emission factors below the 6 scf/h standard and most pneumatic controllers are intermittent rather than continuous bleed; in addition, most wells in the Appalachian basin were constructed after the date put forth in the 2012 NSPS.

Heater emissions for each producing well ($E_{heater,i,j,t}$) are estimated, using the approach described in Roy et al. (2014), as follows[21]:

$$E_{heater,i,j,t} = W_{producing,j,t} \cdot EF_{heater,i} \tag{19}$$

where $W_{producing,j,t}$ is the number of producing wells, and $EF_{heater,i}$ is the emissions per well, as given by:

$$EF_{heater,i,j} = F_{heater,i} \cdot Q_{heater} \cdot T_{heater} \cdot N_{heater}/HHV_{heater} \tag{20}$$

where Q_{heater} is the heater throughput, and $N_{pneumatics}$ is the number of pneumatic devices per well. We base the emission factor ($F_{pneumatics}$), T_{heater} is the number operating hours per year, HHV_{heater} is the higher heating value, and N_{heater} is the number of heaters per well. We do not anticipate any time-varying parameters over the period of analysis.

Fugitive emissions ($E_{fugitives,j,t}$) associated with leaking valves, connectors, flanges, and open-ended lines (OEL), are estimated as follows:

$$E_{fugitives,j,t} = W_{producing,j,t} \cdot EF_{fugitives} \tag{21}$$

where $W_{producing,j,t}$ is the number of producing wells, and $EF_{fugitives}$ is the VOC emissions factor per well, as given by:

$$E_{fugitives,t} = T_{fugitives} \cdot VF_{fugitives} \cdot \sum_c \left(F_{fugitives,c} \cdot N_{fugitives,c}\right) \tag{22}$$

where $T_{fugitives}$ is the annual operating hours per well, $VF_{fugitives}$ is the VOC fraction, $F_{fugitives,c}$ is fugitive total organic carbon emission factor for each component type c, and $N_{fugitives,c}$ is the number of components of each type. We do not account for any time-varying parameters and assume point estimates for each input variable, given that this is a relatively very minor emission source. The NSPS subpart OOOOa for fugitive VOC emissions from production were promulgated but will not take effect (depending upon ongoing regulatory actions) until after the time horizon of this analysis. In addition, Ohio EPA has regulated fugitive emissions from production operations since 2014, requiring quarterly detection at some sites.

B3.1.1.3 Processing

Emissions from processing compressor stations ($E_{processing,i,j,t}$), including only those associated with natural gas liquids extraction facilities are estimated as follows:

$$E_{processing,i,j,t} = RE_{processing,state,i,j,t} \cdot U_{state,t} \tag{23}$$

where $RE_{processing,state,i,j,t}$ is the reported (or simulated) emissions, and $U_{state,t}$ is the proration factor based on the ratio of shale production to all production for each state. We use reported facility-level emissions and point source coordinate locations from the EPA National Emissions Inventory (NEI) for facilities with NAICS 4862 for years 2005, 2008, 2011, and 2014. For non-reporting years, we estimate emissions for each source location j by linearly interpolating between years or extrapolating across years, using reported emissions aggregated to a given source location resolution (i.e., county or 36 x 36 km grid cell). While the NEI dataset is inclusive of compressor stations associated with natural gas liquids extraction facilities, it does not include other types of processing, including ethane crackers or wellhead processing such as glycol dehydration; these other sources are likely non-trivial at present and will be even more significant in the coming years.

B3.1.1.4 Transmission

Emissions from transmission compressor stations ($E_{transmission,i,j,t}$) are estimated as follows:

$$E_{transmission,i,j,t} = RE_{transmission,state,i,j,t} \cdot U_{state,t} \tag{24}$$

where $RE_{transmission,state,i,j,t}$ is the reported (or simulated) emissions. We use reported facility-level emissions and point source coordinate locations from the EPA NEI for facilities with NAICS 211112 for years 2005, 2008, 2011, and 2014. For non-reporting years, we estimate emissions for each source location j by linearly interpolating between years or extrapolating across years, using reported emissions aggregated to a given source location resolution (i.e., county or 36 x 36 km grid cell). While the NEI dataset is inclusive of compressor stations, it does not include pipelines.

B3.1.1.5 Distribution

Emissions from distribution compressor stations ($E_{distribution,i,j,t}$) are estimated as follows:

$$E_{distribution,i,j,t} = RE_{distribution,state,i,j,t} \cdot U_{state,t} \tag{25}$$

where $RE_{distribution,state,i,j,t}$ is the reported (or simulated) emissions. We use reported facility-level emissions and point source coordinate locations from the EPA NEI for facilities with NAICS 22121 for years 2005, 2008, 2011, and 2014. For non-reporting years, we estimate emissions for each source location

j by linearly interpolating between years or extrapolating across years, using reported emissions aggregated to a given source location resolution (i.e., county or 36 x 36 km grid cell). While the NEI dataset is inclusive of compressor stations, it does not include distribution lines and mains.

B3.1.1.6 End use

We estimate emissions from electric utilities ($E_{electric,i,j,t}$) as follows:

$$E_{electric,i,j,t} = \sum_{state} RE_{electric,state,i,j,t} \cdot U_{state,t} \quad (26)$$

where $RE_{electric,state,i,j,t}$ is the reported (or simulated) emissions for each spatial unit j.

A variety of sources report NO_x emissions from the electric power sector, each with varying sectoral and temporal coverage, as shown in Figure B3. For NO_x emissions, we use reported plant-level emissions from the EPA Continuous Emissions Monitoring System (CEMS) because it provides estimates for all years from 2004 to 2016 and is reported at a high spatial resolution; we include facilities with a primary fuel type of pipeline natural gas and identified as electric utilities, including cogeneration. The CEMS dataset is inclusive of most electric power generation, as indicated by the relative heat input and generation across datasets (Table B4).

For VOC and PM2.5 emissions, we use reported facility-level emissions from NEI for years 2005,2008, 2011, and 2014. To extract natural gas facility emissions data from the NEI dataset and coordinate locations from the eGRID dataset, we do a crosswalk of the NEI, eGRID, and CEMS datasets. For non-reporting years, we estimate emissions for each source location j by linearly interpolating between years or extrapolating across years, using reported emissions aggregated to a given source location resolution (i.e., county or 36 x 36 km grid cell).

Table B3. Comparison of electric power emissions sector data sources.

Reporting system	Sector Coverage	Subsector Coverage	Years	Spatial Resolution
CEMS	Electric Power Industry (generators >25 MW)	Electric Utility; Industrial Boiler; Pulp & Paper Mill; Iron & Steel	2004 - 2016	plant
eGRID	Electric Power Industry (grid-connected)	-	2004, 2005, 2007, 2009, 2010, 2012, 2014, 2016	plant
EIA	Electric Power Industry	Electric Utility; Industrial; Commercial	2013 - 2016	plant
EIA	Electric Power Industry	Electric Utility; IPP NAICS-22 Non-Cogen; IPP NAICS-22 Cogen; Commercial Cogen; Commercial Non-Cogen; Industrial Cogen; Industrial Non-Cogen	2004 - 2016	state

Table B4. Comparison of 2014 electric utility data for Pennsylvania, Ohio, and West Virginia.

Reporting system	NOx Emissions (short tons)	Net Generation (TWh)	Heat Input (million mmBTU)
CEMS	3,996	71.0	574
eGRID	4,286	75.9	573
EIA	6,701	75.8	594

Emissions from industrial end use ($E_{industrial,i,j,t}$) are estimated as follows:

$$E_{industrial,i,j,t} = RE_{industrial,state,i,j,t} \cdot U_{state,t} \quad (27)$$

where $RE_{industrial,state,i,j,t}$ is the reported (or simulated) emissions. We use reported county-level emissions from the EPA NEI for commercial natural gas combustion for years 2008, 2011, and 2014; this includes natural gas that is combusted by industrial boilers and internal combustion engines.[44] For non-reporting years, we estimate emissions for each county by linearly interpolating between years or extrapolating across years; we use an area-weighting approach to develop estimates at the 36 x 36 km grid cell resolution.

Emissions from commercial end use ($E_{commercial,i,j,t}$) are estimated as follows:

$$E_{commercial,i,j,t} = RE_{commercial,state,i,j,t} \cdot U_{state,t} \quad (28)$$

where $RE_{commercial,state,i,j,t}$ is the reported (or simulated) emissions. We use reported county-level emissions from the EPA NEI for commercial natural gas combustion for years 2008, 2011, and 2014; this includes natural gas that is combusted by commercial and institutional boilers and internal combustion engines.[44] For non-reporting years, we estimate emissions for each county by linearly interpolating between years or extrapolating across years; we use an area-weighting approach to develop estimates at the 36 x 36 km grid cell resolution.

Emissions from residential end use ($E_{residential,i,j,t}$) are estimated as follows:

$$E_{residential,i,j,t} = RE_{residential,state,i,j,t} \cdot U_{state,t} \quad (29)$$

where $RE_{residential,state,i,j,t}$ is the reported (or simulated) emissions. We use reported county-level emissions from the EPA NEI for residential natural gas combustion for years 2008, 2011, and 2014; this includes natural gas that is combusted for residential household heating, grills, hot water heating, and dryers.[44] For non-reporting years, we estimate emissions for each county by linearly interpolating between years or extrapolating across years; we use an area-weighting approach to develop estimates at the 36 x 36 km grid cell resolution.

B3.1.2 Mortality and monetized damages

We estimate premature mortalities by combining the emissions model with three RCMs, AP3, APSCA, and InMAP. We use source-receptor versions of AP3 and APSCA to generate spatially-resolved marginal statistical mortalities for each receptor area associated with emissions from a source area. We use a source-resolved version of InMAP, whereby marginal statistical mortalities are allocated to the source location..

AP3 is an integrated assessment model which uses a dispersion model to link emissions to annual average concentrations, and then estimates exposures and mortalities based on predicted concentrations, demographic data, and dose-response functions.[45,46] Marginal impacts by receptor location associated with NO_x, $PM_{2.5}$, and VOCs emissions from source locations are resolved at the county level and assume a base year of 2014. To account for time-varying marginal mortality, we apply an annual population adjustment based on reported annual county-level population.

APSCA is a receptor-resolved version of the Estimating Air Pollution Social Impact Using Regression (EASIUR) model.[47] EASIUR marginal damages are derived using regression on simulations from CAMx, a state-of-the-art chemical transport model.[48] Marginal impacts by receptor location associated with NO_x and $PM_{2.5}$ emissions from source locations are resolved at the 36 x 36 km grid cell level and assume varying base years. We perform a single factor sensitivity analysis to account for uncertainty in the dispersion modeling in EASIUR using the 95% prediction interval multipliers reported Heo and Adams (2015).[48]

For all models, we parametrically vary the concentration-response relationship based on the American Cancer Society (ACS) and Harvard Six Cities (H6C) studies.[49,50] Specifically, the relative risk values assumed are 1.06 (ACS) and 1.14 (H6C), terms of increased mortality per 10 µg PM2.5/m^3 increase.

To develop monetized impact estimates, we use the value of a statistical life (VSL), which is "a summary measure for the dollar value of small changes in mortality risk experienced by a large number of people. and premature mortality estimates to generate monetary damages."[51] We simulate VSL based on the EPA-recommended default distribution for preparing economic analyses (Weibull distribution with a scale = 9.42 and shape = 1.51 converted to 2017 USD using the Consumer Price Index); the simulated mean (and standard deviation) VSL is $8.5M (+/- $5.7M).

B3.2 Additional results

The following are additional results from the air quality emissions, mortality, and damages modeling.

Table B5. Emissions per unit activity for 2004 and 2016 for preproduction and production processes. Mean emissions per unit activity are provided, in addition to the percent change in emissions from 2004 to 2016. Note that emissions reflect systems-level factors (e.g., fleet turnover of rate, percentage of wells with wellhead compressors). Comparison to unit-level emissions for 2009 reported in Roy et al. (2014).

Segment / process	NO_x				VOC				$PM_{2.5}$			
	2004	2016	%△	Roy et al. (2014)	2004	2016	%△	Roy et al. (2014)	2004	2016	%△	Roy et al. (2014)
Preproduction (tons/spud wells)												
Drilling	5.07	4.40	-13%	4.4	0.60	0.46	-23%	0.5	0.32	0.24	-23%	0.3
Hydraulic fracturing	1.88	1.20	-36%	2.2	0.23	0.08	-8%	0.25	0.14	0.05	-64%	0.16
Trucking	1.83	0.12	-94%	6.9	0.10	0.01	-1%	0.4	0.02	0.00	-94%	0.07
Well completion	-	-	-	-	0.79	0.08	-8%	3.8	-	-	-	-
Production (tons/producing well)												
Condensate tanks (tons/bbl)	-	-	-	-	0.01	0.00	0%	0.0003	-	-	-	-
Heaters	0.08	0.08	0%	0.0	0.00	0.00	0%	0.0	-	-	-	-
Pneumatics	-	-	-	-	0.00	0.00	0%	0.5	-	-	-	-
Fugitives	-	-	-	-	0.00	0.00	-2%	0.2	-	-	-	-
Wellhead compressors	0.02	0.01	-53%	1.1	0,01	0.00	0%	0.4	0.00	0.00	0%	0.01

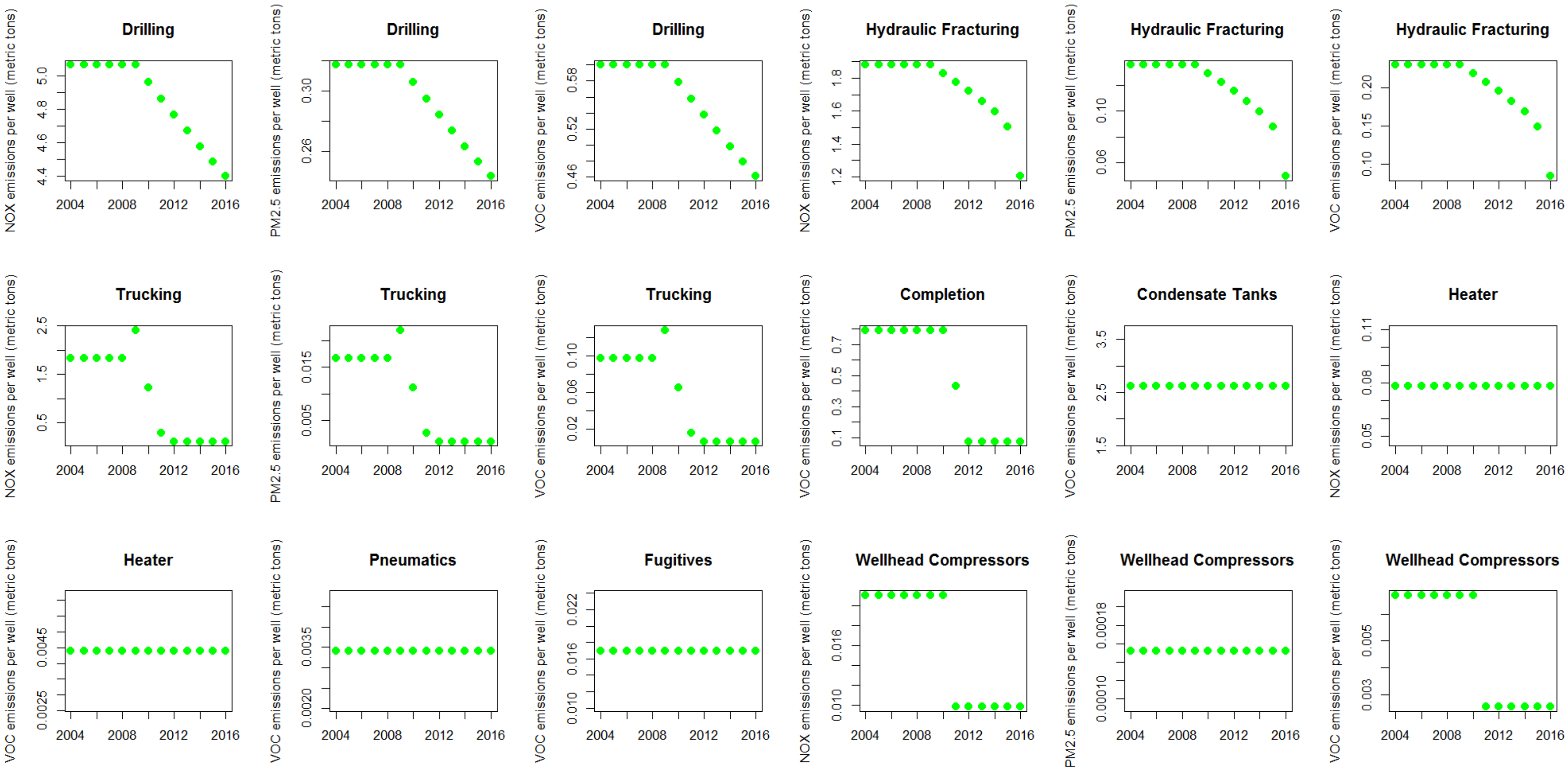

Figure B13. Emissions per unit activity from 2004 to 2016. Mean emissions per unit activity are provided.

Table B6. Cumulative air pollution emissions and percent attribution for each segment and process across the supply chain.

Segment / process	2004 to 2016 Cumulative Emissions					
	NO_x		VOC		$PM_{2.5}$	
	Emissions (thousand metric tons)	%	Emissions (thousand metric tons)	%	Emissions (thousand metric tons)	%
Preproduction	92	14%	14	10%	5.3	15%
Production	5	1%	71	51%	0.0	0%
Processing	0	0%	1	1%	0.0	0%
Transmission	121	18%	18	13%	3.8	11%
Distribution	4	1%	1	1%	0.3	1%
End use	447	67%	34	24%	25.7	73%
Total Supply Chain	**670**	**100%**	**140**	**100%**	**35.1**	**100%**

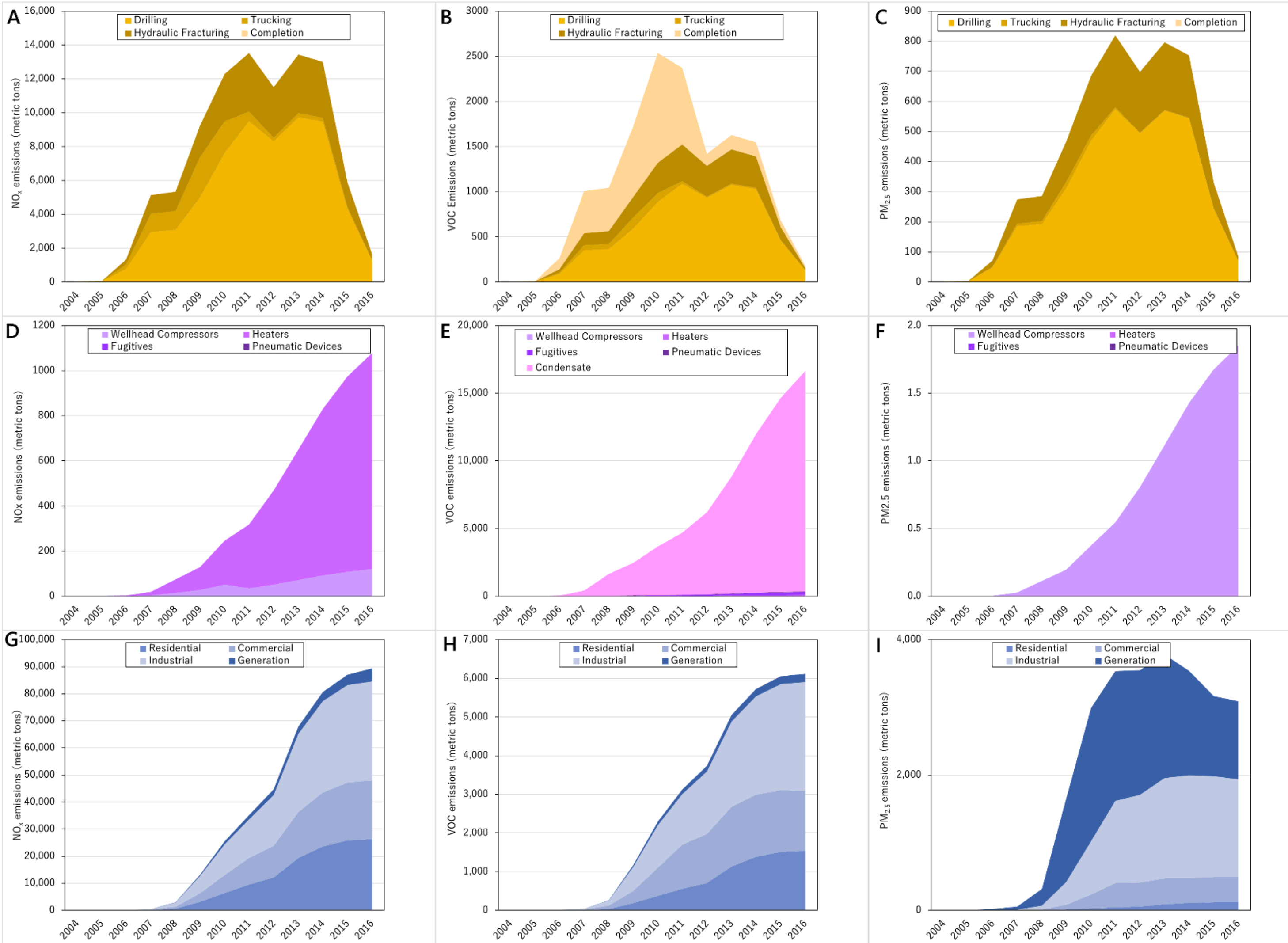

Figure B14. Emissions by process/segment over time. Preproduction (A) NO_x, (B) VOC, and (C) $PM_{2.5}$ emissions under baseline scenario. Production (D) NO_x, (E) VOC, and (F) $PM_{2.5}$ emissions under baseline scenario. End use (G) NO_x, (H) VOC, and (I) $PM_{2.5}$ emissions.

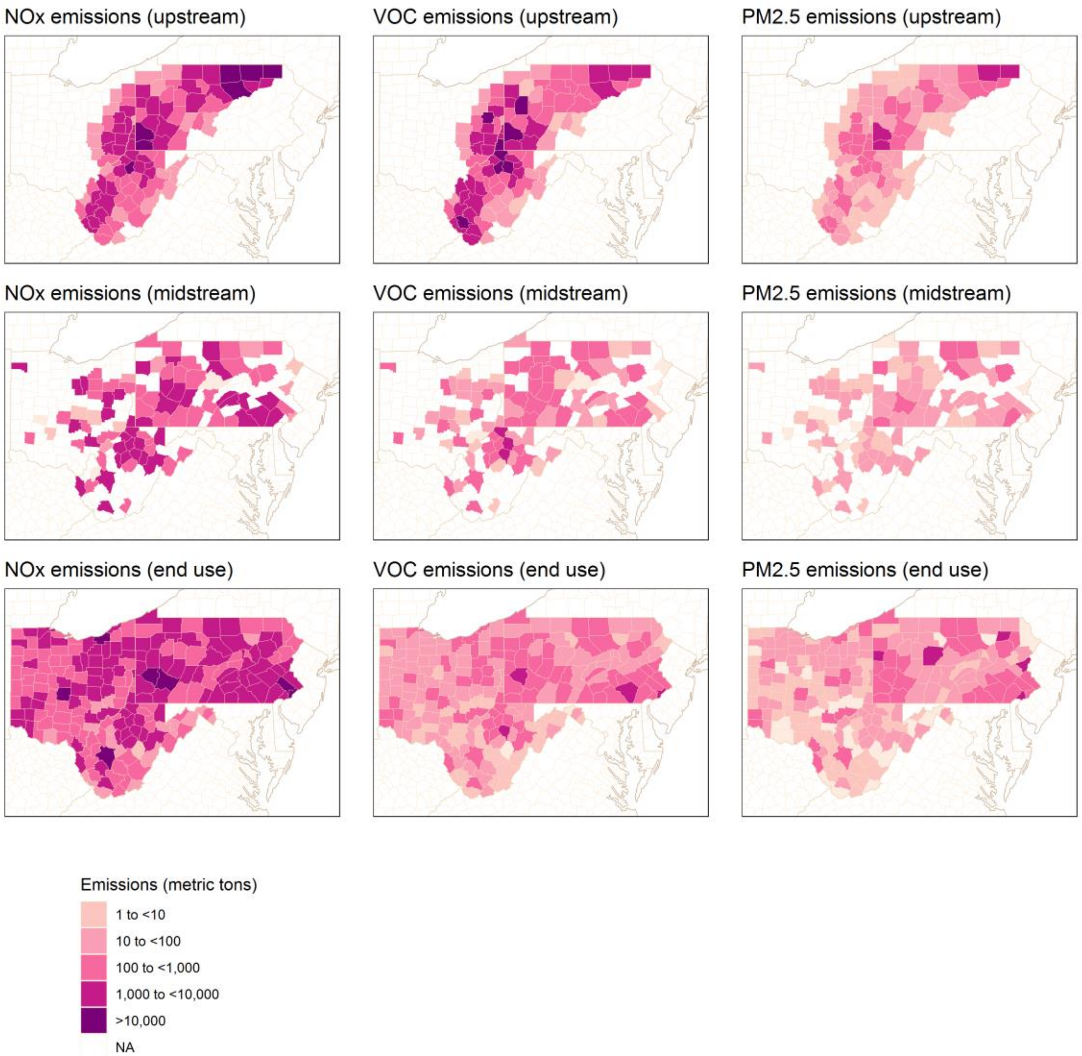

Figure B15. Maps of cumulative emissions for each county for different segments of the supply chain.

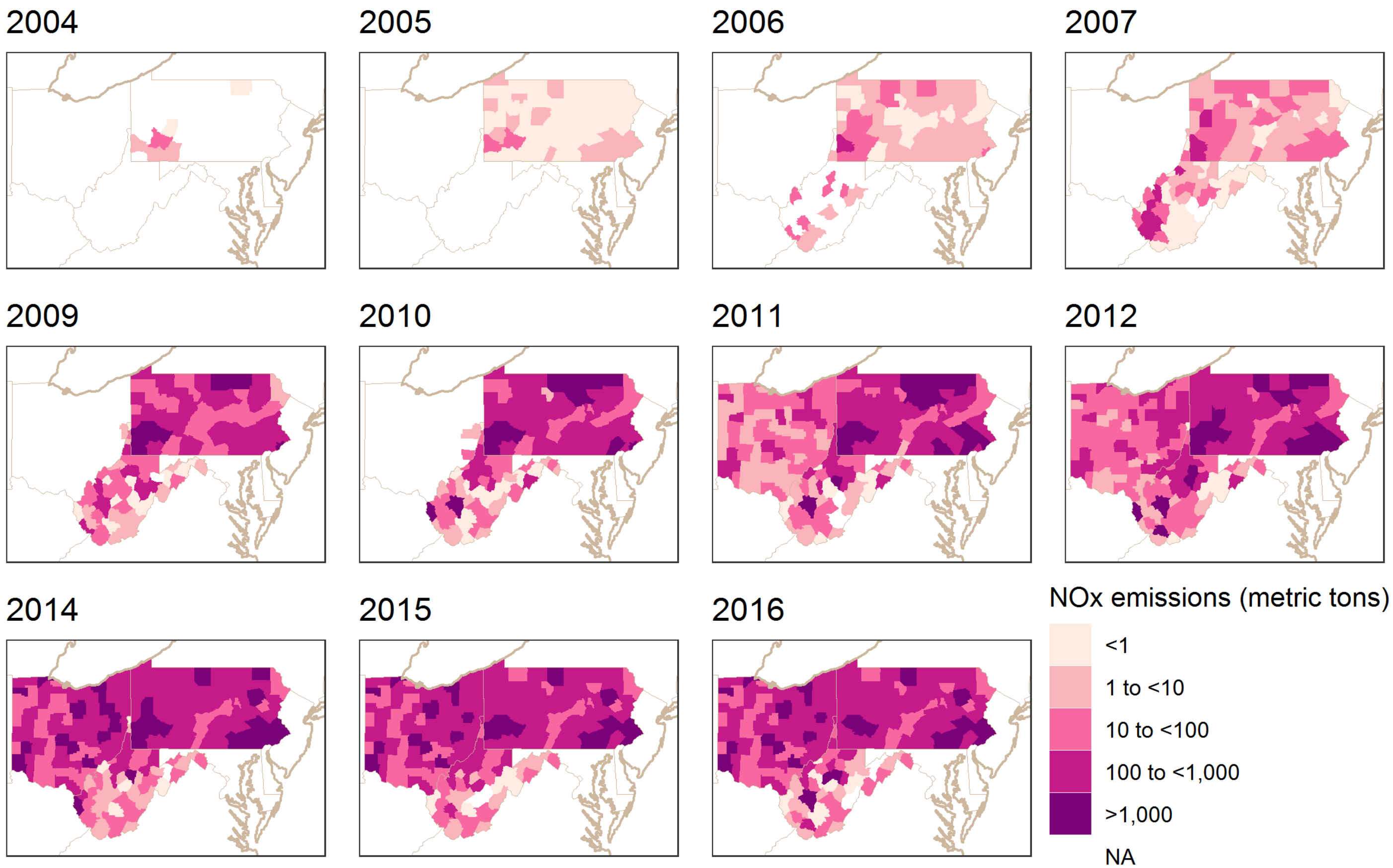

Figure B16. Maps of annual NO_x emissions for each county.

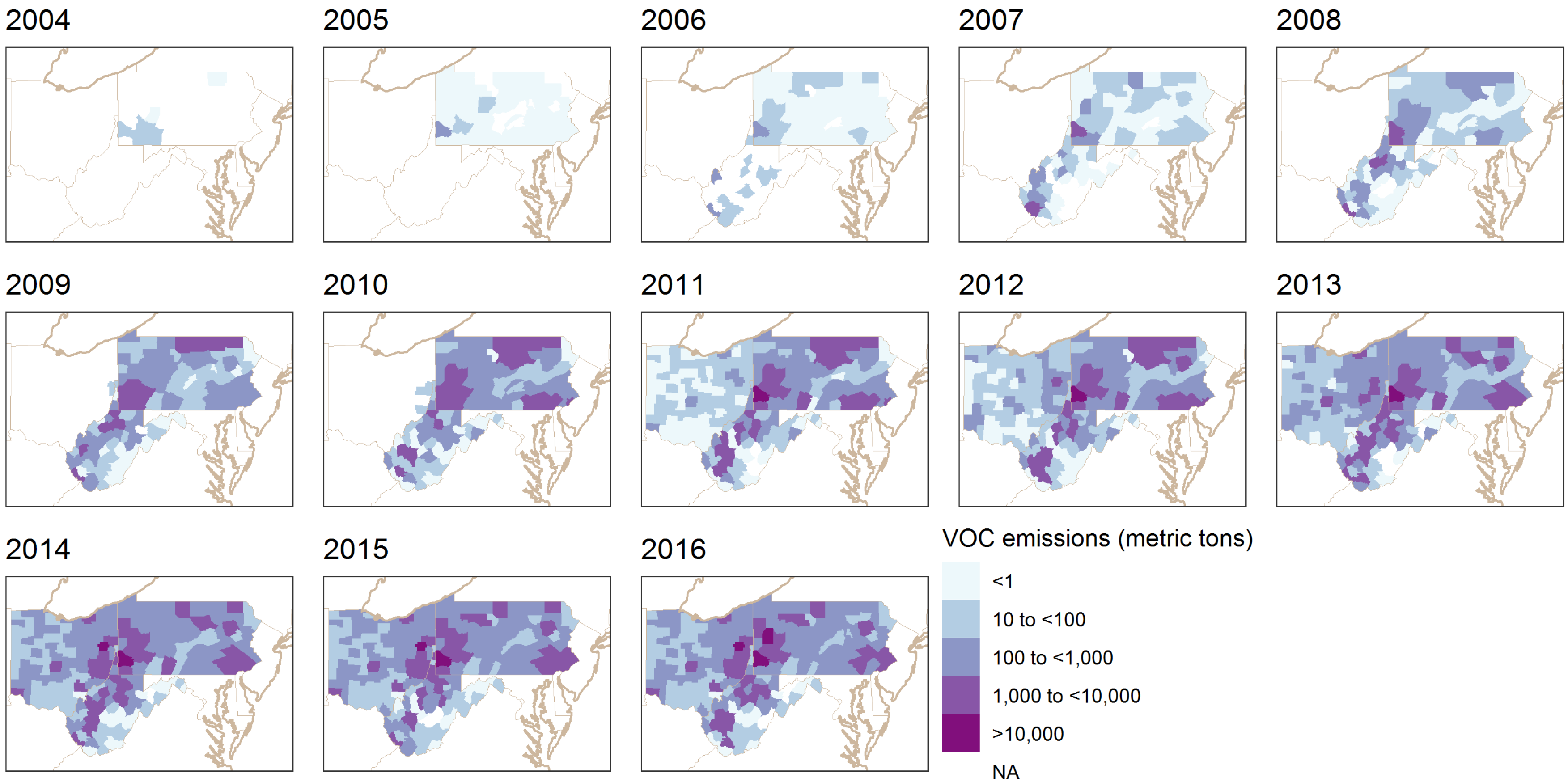

Figure B17. Maps of annual VOC emissions for each county.

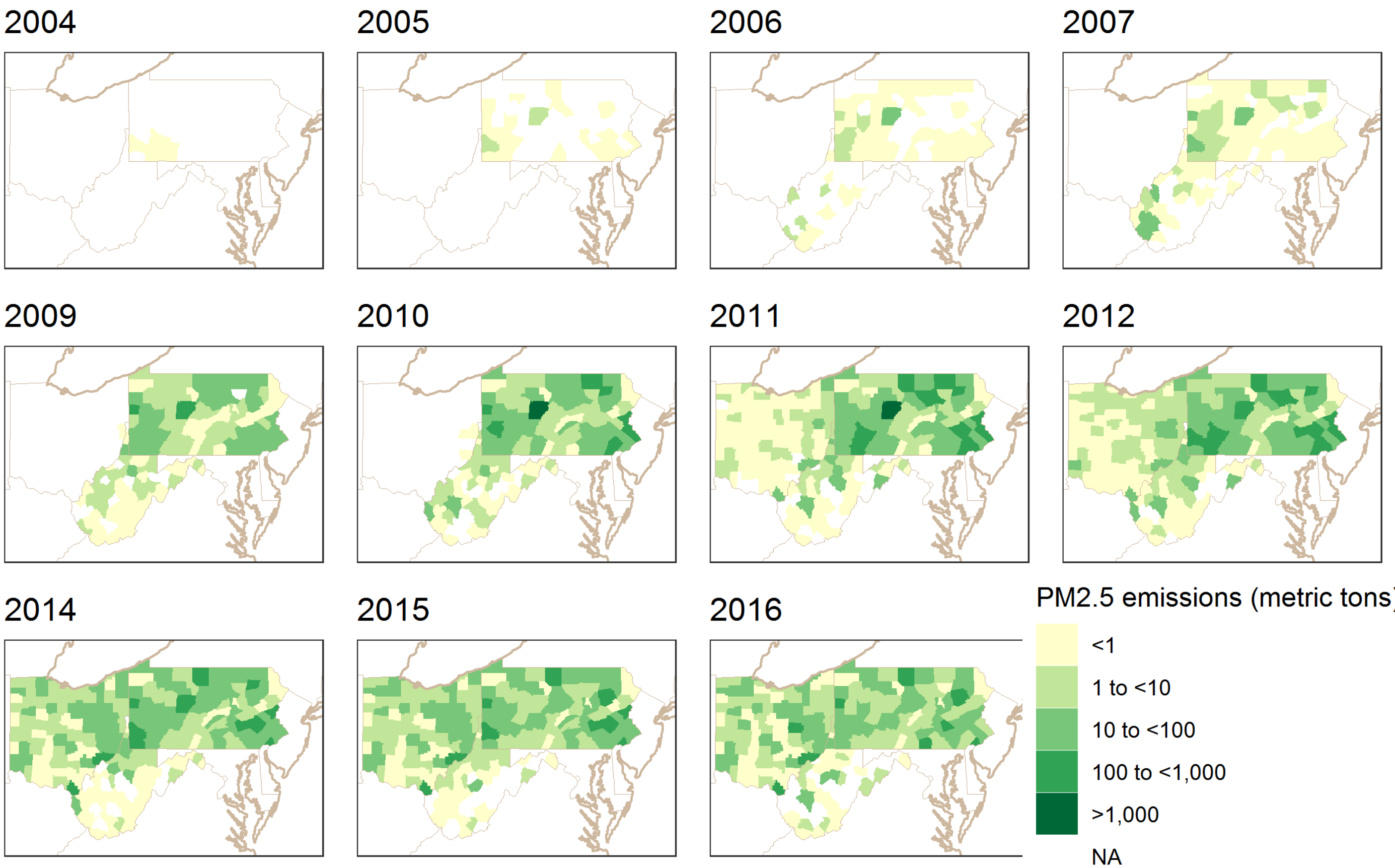

Figure B18. Maps of annual $PM_{2.5}$ emissions for each county.

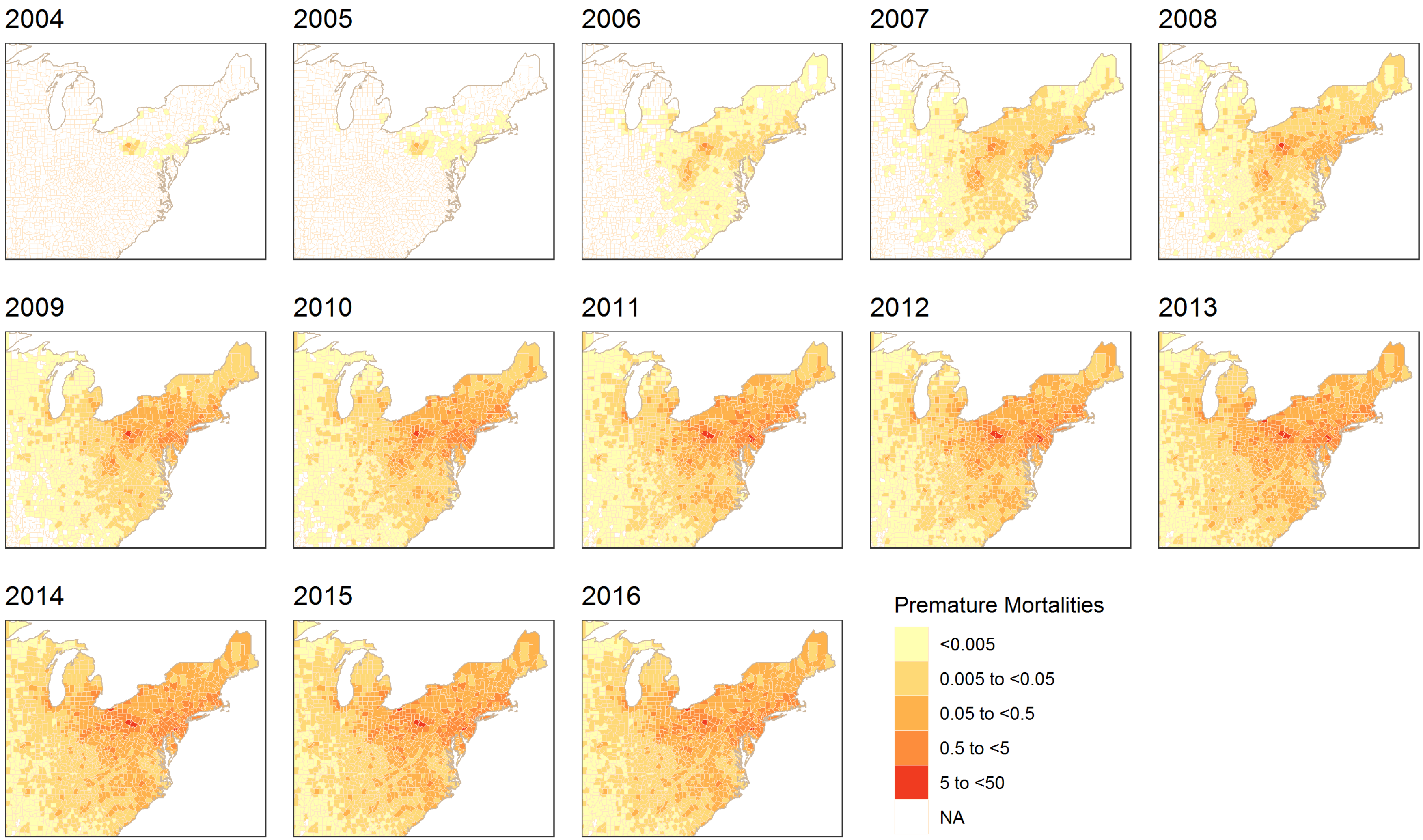

Figure B19. Maps of annual premature mortalities using AP3 and ACS concentration-response relationship.

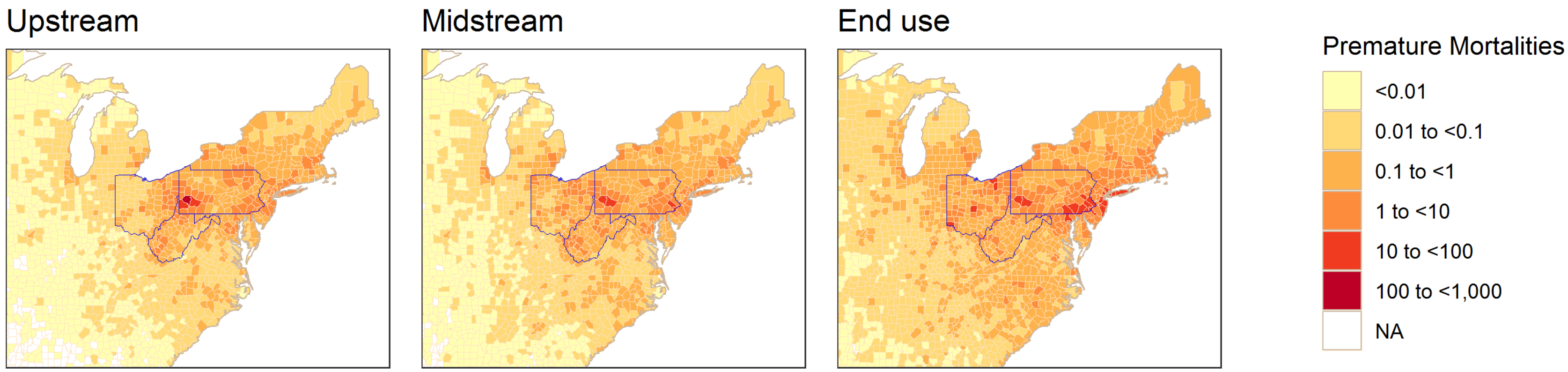

Figure B20. Maps of cumulative premature mortalities using AP3 and ACS concentration-response relationship for each segment of supply chain.

Table B7. Cumulative premature mortalities and monetized damages from 2004 to 2016. Damages in 2017 USD. Mortalities and damages based on base scenario emission assumptions. Mean damages are provided, as well as, 95% confidence intervals (in parentheses) reflecting uncertainty in the VSL. Estimates based on mean, all cause relative risk values from fine particulate matter reported in Pope et al. (2002) (ACS Cohort) and Lepeule et al. (2012) (Harvard Six Cities cohort).

Segment	Premature mortality		Damages (billion $)	
	ACS Cohort Study	Harvard Six Cities Study	ACS Cohort Study	Harvard Six Cities Study
AP3				
Upstream	461	942	4 (0.4 - 10)	8 (0.8 - 20)
Midstream	450	921	4 (0.4 - 10)	8 (0.8 - 20)
End use	1360	2766	12 (1 - 31)	23 (2 - 61)
Supply chain	**2270**	**4629**	**19 (2 - 51)**	**39 (4 - 103)**
APSCA				
Upstream	256	657	2 (0.2 - 6)	6 (0.5 - 15)
Midstream	266	684	2 (0.2 - 6)	6 (0.6 - 15)
End use	686	1763	6 (0.6 - 15)	15 (1 - 39)
Supply chain	**1208**	**3103**	**10 (1 - 27)**	**26 (3 - 69)**
InMAP				
Upstream	240	618	2 (0.2 - 5)	5 (0.5 - 14)
Midstream	248	638	2 (0.2 - 6)	5 (0.5 - 14)
End use	640	1643	5 (0.5 - 14)	14 (1 - 37)
Supply chain	**1129**	**2900**	**10 (0.9 - 25)**	**25 (2 - 65)**

B4 Climate change model

B4.1 Methods

We model temporally-resolved emissions of methane (CH_4) and carbon dioxide (CO_2) and climate impacts associated with shale gas activities across the supply chain from preproduction to end use from 2004 to 2016. We estimate climate impacts in terms of global temperature change and monetized damages. To account for major sources of uncertainty, we perform sensitivity analyses of key inputs, such as the absolute global temperature potential and the social cost of carbon and methane, and develop process-or site-level emission ranges for upstream and midstream processes.

B4.1.1 Emissions model

B4.1.1.1 Background

Climate change impacts of natural gas activity are a function of greenhouse gases (GHG) emitted across the supply chain, from preproduction through end use. Emissions are associated with various preproduction processes, such as well pad preparation, drilling, hydraulic fracturing, and well completion.[41] Emissions from other upstream and midstream segments, including production, gathering, processing, transmission and storage, and distribution, largely consist of vented and fugitive from fuel combustion.[43,52–54] Additional combustion emissions are associated with downstream processes, including electricity generation[43,53,55], commercial or residential heating, industrial end use, use as alternative transportation fuel[42,56], and liquefied natural gas exports[57].

According to the U.S. Greenhouse Gas Inventory (GHGI), the U.S natural gas sector in 2016 was the source of approximately a quarter of all GHG emissions and a third of all energy-related emissions.[58] Within the natural gas sector, almost 90% of emissions result from downstream combustion, a majority of which are from electric power and industrial end use segments, and the remaining 10% stem from upstream and midstream processes, as shown in

Table B8.[58] Annual emissions from upstream and midstream sources have slightly increased between 2004 and 2016, reflecting both increasing natural gas activity and decreasing emission factors from voluntary and regulatory reductions.

Several life cycle assessments (LCA) have been conducted to quantify the GHG emissions of current and emerging life cycle stages. A common finding is that life cycle emissions from domestic use of natural gas (including unconventional shale gas) are lower than coal, even when considering uncertainty in methane leakage rates.[41,43,53,54,59–61] The climate benefit of natural gas relative to coal is further supported at an energy systems level; the CO_2 intensity of U.S. electricity production has decreased between 2001 and 2017 by 30%, a trend reflective of declining coal generation and corresponding increases in natural gas and wind generation.[62]

Table B8. U.S. Greenhouse Gas Inventory and life cycle assessment GHG emissions across the natural gas supply chain.[41,53,58]

Segment	U.S. Greenhouse Gas Inventory				Life Cycle Assessment[a]	
	2004		2016		Emissions (CO_2eq / MJ)	% Emissions
	Emissions (CO_2eq mmt)	% Emissions	Emissions (CO_2eq mmt)	% Emissions		
Preproduction	11	1%	1	0%	1.8	3%
Production	57	4%	57	3%	9.7	14%
Gathering	31	2%	53	3%	b	b
Processing	31	2%	33	2%	4.3	6%
Transmission and storage	37	3%	33	2%	1.4	2%
Distribution	25	2%	12	1%	0.8	1%
End Use	1183	86%	1476	89%	50	73%
Total Supply Chain	1374	100%	1665	100%	68	100%

a These are life cycle emissions for domestic use of natural gas.
b Gathering emissions are included in production.

In the following sections, we describe regional and unit-level emissions model formulations and assumptions for each segment. Table B9 is a summary of emissions modeling input parameters, including definitions, values, and data sources.

Using a combination of top-down and bottom-up approaches, we derive both regionally aggregated emissions estimates, as well as, segment-specific emission factors. The inclusion of and detail in which we model each segment of and process within the natural gas sector varies based on the availability and quality of data, and the relative emissions contribution of each source, as indicted by national inventories and LCA studies. We also focus on representing uncertainty of systems-level emissions rather than unit-level variation. We use a combination of parametric and probabilistic methods for upstream processes, rather than employing a fully stochastic framework.

We account for time-varying parameters (where practicable), such as evolving regulation and changing activity factors. While we account for changing well completion regulation and corresponding practices over time, we do not explicitly differentiate between emission factors pre- and post- implementation of the 2016 NSPS subpart OOOO which limits vented and fugitive methane emissions from production wells, gathering stations, processing facilities, and transmission and storage compressor stations.[63] For the segments regulated under the NSPS subpart OOOO, we use data and modeling results from large-scale methane measurement studies conducted across the natural gas supply chain between 2013 and 2016[28,29,64–

[68]; these studies provide the best available methane loss rates for each segment and thus the most appropriate for representing operating practices over the period of our analysis. With respect to electricity generation, we use annually reported plant-level emissions, which we assume inherently incorporate time-varying regulatory, efficiency, and activity factors.

Unlike for the air quality impact model, we do not develop spatially-resolved source GHG emission estimates given that the species considered are well-mixed GHGs and global forcing per unit of emission are independent of geographic location of emission.[69] Spatial variability in GHG emission sources is potentially relevant from a regulatory perspective, namely with respect to classifying and identifying the regulated community or emitters for policy design and enforcement (e.g., methane superemitter policy).

For modeling the preproduction, production, and gathering segments, we explicitly use shale gas activity factors, such number of well. To attribute emissions from midstream and downstream segments to shale production, we use a prorate factor relative to total natural gas production (refer to section B2.4).

Table B9. Emissions modeling parameter values, definitions, and data sources.

Parameter	Parameter Definition	Parameter Value	Units	Source
$\theta_{1,production}$	fitted parameter for estimating mean methane emission factor for production	0.60 (0.44 to 0.81)	-	70
$\theta_{2,production}$	fitted parameter for estimating mean methane emission factor for production	1.4 (1.3 to 1.8)	-	70
$\sigma_{production}$	fitted parameter for estimating mean methane emission factor for production	1.3 (1.1 to 1.6)	-	70
$a_{production}$	fitted parameter for estimating mean methane emission factor for production	3.0 (2.6 to 3.2)	-	70
$b_{production}$	fitted parameter for estimating mean methane emission factor for production	-2.2 (-2.6 to -1.8)	-	70
$c_{production}$	fitted parameter for estimating mean methane emission factor for production	0.20 (0.050 to 0.42)	-	70
$C_{distribution,CH4}$	distribution methane content	93.4%	%	71
$C_{production,CH4,t}$	production methane content for northeast National Energy Modeling System (NEMS) region	time-varying (ranges from 83 to 84%)	%	71
$C_{production,CO2,t}$	production carbon dioxide content for unconventional natural gas in northeast NEMS region	3.5%	%	71
$C_{transmission,CH4}$	transmission and storage methane content	93.4%	%	71
$C_{processing,after,CO2}$	carbon dioxide content after processing	1%	%	71
$C_{processing,before,CO2}$	carbon dioxide content prior to processing	7.4%	%	71
$EF_{combustion}$	carbon dioxide emission factor for natural gas combustion	54	metric ton/mmcf	72,73
$EF_{completion,CH4,controlled}$	methane emissions factor for completions with emission controls	2.23 (base), 0.66 (low), 4.65 (high)	metric ton/spud well	28
$EF_{completion,CH4,uncontrolled}$	methane emissions factor for completions without emission controls	0.84 (base), 0.38 (low), 1.63 (high)	metric ton/spud well	28
$EF_{drilling}$	carbon dioxide emission factor for drilling	390 (base), 280 (low), 500 (high)	metric ton/spud well	41
$EF_{hydraulic}$	carbon dioxide emission factor for hydraulic fracturing pumping	460 (base), 230 (low), 690 (high)	metric ton/spud well	41
$EF_{main,type}$	methane emissions factors for each mile of main	cast iron: 1.16 (4.31), unprotected steel: 0.86 (2.32), protected steel: 0.10 (0.37), plastic: 0.03 (0.06)	metric ton/mile	58,67

Table B9 (continued). Emissions modeling parameter values, definitions, and data sources.

Parameter	Parameter Definition	Parameter Value	Units	Source
$EF_{meters,type}$	methane emissions factor for distribution meters and regulators	meter > 300 psi: 2.14 (3.59), meter 100 - 300 psi: 1.00, meter < 100 psi: 0.73, regulator > 300 psi: 0.87 (2.57), regulator 100 - 300 psi: 0.14 (0.39), regulator 40 - 100 psi: 0.16 (1.20), regulator < 40 psi: 0.02, vault: 0.05 (0.07)	metric ton/meter	[58,67]
$EF_{service,type}$	methane emissions factor for each service line	unprotected steel: 0.01 (0.04), protected steel: 0.001 (0.002), plastic: 0.0003 (0.0004), copper: 0.005	metric ton/service line	[58,67]
$EF_{wellpadpreparation}$	carbon dioxide emission factor for well pad preparation	340 (base), 300 (low), 360 (high)	metric ton/well	[41]
$F_{completion,controlled,t}$	fraction of well completions with emission controls	time-varying	-/-	[28,29]
$HHV_{reference}$	reference higher heating value used in AP-42 natural gas combustion carbon dioxide emission factors	1000	BTU/cf	[58]
$HHV_{s,t}$	higher heating value for delivered natural gas	time-varying	BTU/cf	[74]
$L_{gathering}$	methane loss rate for gathering facilities	0.40% (base), 0.36% (low), 0.45% (high)	%	[75]
$L_{processing}$	methane loss rate for processing segment	0.18% (base), 0.16% (low), 0.20% (high)	%	[75]
$L_{transmission}$	methane loss rate for transmission and storage segment	0.35% (base), 0.28% (low), 0.45% (high)	%	[64]
$M_{distribution,s,t,type}$	number of meters and regulators	time-varying (see Figure B21)	meters	[58,76]
$P_{main,s,t,type}$	length of main pipeline of each type	time-varying (see Figure B21)	miles	[76]
$P_{production,well,t}$	production per well	time-varying (refer to section B2.2)	mcf	[5–7]
$P_{service,s,t,type}$	number of service lines of each type	time-varying (see Figure B21)	service lines	[76]
$U_{s,t}$	fraction total unconventional production out of total production	time-varying (ranges from 0 to 99%)	-/-	[5–8]
$V_{commercial,s,t}$	volume of natural gas delivered to commercial consumers	time-varying (see Figure B9)	mmcf	[16]
$V_{completion}$	natural gas combusted during well completions for wells with emission controls	1.57 (base), 0.46 (low), 2.87 (high)	mmcf/well	[28]
$V_{distribution,s,t}$	volume of natural gas delivered to consumers	time-varying (see Figure B9)	mmcf	[77]
$V_{industrial,s,t}$	volume of natural gas delivered to industrial consumers	time-varying (see Figure B9)	mmcf	[15]

Table B9 (continued). Emissions modeling parameter values, definitions, and data sources.

Parameter	Parameter Definition	Parameter Value	Units	Source
$V_{leasefuel,s,t}$	natural gas lease fuel consumption	time-varying (see Figure B9)	mmcf	11
$V_{plantfuel,s,t}$	natural gas plant fuel consumption	time-varying (see Figure B9)	mmcf	12
$V_{processing,s,t}$	natural gas processing volume (i.e., unprocessed volume received by plant)	time-varying (see Figure B9)	mmcf	10
$V_{production,s,t}$	shale gas production volume	time-varying (refer to section B2.1)	mmcf	5–7
$V_{residential,s,t}$	volume of natural gas delivered to residential consumers	time-varying (see Figure B9)	mmcf	18
$V_{transmissionfuel,s,t}$	transmission and distribution fuel consumption	time-varying (see Figure B9)	mmcf	13
$W_{producing,t}$	number of producing wells	time-varying (refer to section B2.2)	wells	5–7
$W_{spud,t}$	number of spud wells	time-varying (refer to section B2.3)	wells	5–7

B4.1.1.2 Preproduction

We estimate well pad preparation emissions ($E_{wellpadpreparation,t}$) from land clearing and well pad construction for each year t as follows:

$$E_{wellpadpreparation,t} = \sum_s EF_{wellpadpreparation} \cdot W_{spud,s,t} \tag{30}$$

where $W_{spud,s,t}$ is the number of spud wells for each year t and state s. We similarly estimate drilling emissions ($E_{drilling,t}$) as follows:

$$E_{drilling,t} = \sum_s EF_{drilling} \cdot W_{spud,s,t} \tag{31}$$

We also estimate hydraulic fracturing emissions ($E_{hydraulic,t}$) from pumping as follows:

$$E_{hydraulic,t} = \sum_s EF_{hydraulic} \cdot W_{spud,s,t} \tag{32}$$

We use a range of emission factors for well pad preparation ($EF_{wellpadpreparation}$), drilling ($EF_{drilling}$), and hydraulic fracturing ($EF_{hydraulic}$) based on modeling results from Jiang et al. (2011). We do not account for changing practices and operating efficiencies related to well pad preparation, drilling, and hydraulic fracturing (e.g., a reduction in drilling time, a reduction in fracturing stages, and an increase in the number of wells per pad), which collectively are likely to have trivial impacts on total GHG emissions.

Emissions from well completions are highly uncertain, but are a relatively minor source, contributing less than 1% of life cycle GHG emissions.[41–43] We estimate methane losses from well completion ($E_{completion,CH4}$) as follows:

$$E_{completion,CH4,t} = \sum_s \left[F_{completion,controlled,t} \cdot EF_{completion,CH4,controlled} + \left(1 - F_{completion,controlled,t}\right) \cdot EF_{completion,CH4,uncontrolled}\right] \cdot W_{spud,s,t} \quad (33)$$

We attempt to reflect voluntary adoption of emission controls and rapidly evolving regulation, including the implementation of the NSPS subpart OOOO and National Emissions Standards for Hazardous Air Pollutants (NESHAP) standards, that require reduced emission completions (RECs) for hydraulically fractured wells by 2015.

We additionally derive methane emission factors for wells with emission controls ($EF_{completion,CH4,controlled}$), and without emission controls ($EF_{completion,CH4,uncontrolled}$) based on a measurement study conducted by Allen et al. (2013).[28] Given insufficient sample size, we do not further segregate types of emission controls. To derive a range of emission factor estimates, we employ a bootstrapping method used in other emission studies[28,78–80]; we resample with replacement *n* times from the dataset (where *n*=25 is the sample size), estimate the mean of the bootstrapped sample, iterate 100,000 times, and then find the mean and 95% confidence interval across the bootstrapped means.

We assume a changing fraction of well completions with emission controls ($F_{completion,controlled,t}$) and without emission controls ($F_{completion,uncontrolled,t}$). The penetration of emission controls over time is uncertain, although recent emission measurement studies suggest that most wells in Appalachia (for which measurements were taken) have emission controls.[28,29]

We also estimate carbon dioxide flaring emissions from well completion ($E_{completion,CO2}$) as follows:

$$E_{completion,CO2,t} = \sum_s F_{completion,controlled,t} \cdot V_{completion} \cdot EF_{combustion} \cdot W_{spud,s,t} \quad (34)$$

where $EF_{combustion}$ is the carbon dioxide emissions factor for natural gas combustion. The volume of natural gas combusted during controlled well completions ($V_{completion}$) is derived using the previously described bootstrapping method, using data from Allen et al. (2013).[28] Although we do not explicitly assume the flaring rate as in several previous studies, the combusted volume incorporates some observations in which flaring is employed.[41–43,81]

B4.1.1.3 Production

We develop a bottom-up estimate of methane losses from producing wells ($E_{production,s,t}$) based on a range of emissions factors ($EF_{production,CH4}$) conditional on site-level production as follows:

$$E_{production,CH4,t} = \sum_{well} EF_{production,CH4,well,t} \cdot W_{producing,t} \quad (35)$$

where the emission factor is:

$$EF_{production,CH4,well,t} = e^{\mu_{production,well},t+1/2\sigma_{production}{}^2} \tag{36}$$

where $\sigma_{production}$ is the standard deviation (fitted value) and $\mu_{production,well,t}$ is the mean given by:

$$\mu_{production,well,t} = a_{production} + b_{production} P_{production,well,t}^{\theta_{1,production}} + c_{producing} P_{production,well,t}^{\theta_{2,production}} \tag{37}$$

where $a_{production}$, $b_{production}$, $\theta_{1,production}$, and $\theta_{2,production}$ are fitted parameters, and $P_{production,well,t}$ is the annual production per well. The empirical relationship defining the emission factor conditional on site-level production described by equations 36 and 37 is that derived in Alvarez et al. (2018), a study which comprehensively evaluated methane emissions for production based on several recent measurement studies.[70] The study provides a range of fitted values for the parameters specific for the basin, defining the mean and 95% confidence interval of emission factors. While the empirical relationship is based on site-level rather than well-level data, we assume that the relationship applies at the well level; based on a GIS cluster analysis defining the number of wells per site in Alvarez et al. (2018), 96% of sites within the basin have one well.

We also estimate carbon dioxide emissions from combustion of lease fuel ($E_{production,CO2,t}$) as follows:

$$E_{production,CO2,t} = \sum_s V_{leasefuel,s,t} \cdot U_{s,t} \cdot EF_{combustion} \tag{38}$$

where $V_{leasefuel,s,t}$ is the lease fuel consumed (as reported by the EIA) and $U_{s,t}$ is the unconventional prorate factor.

B4.1.1.4 Gathering

We develop a top-down estimate of methane losses from gathering facilities ($E_{gathering,s,t}$) based on gathering volume and methane loss rate as follows:

$$E_{gathering,t} = \sum_s \left(V_{production,s,t} - V_{leasefuel,s,t} \cdot U_{s,t}\right) \cdot C_{production,t} \cdot L_{gathering} \tag{39}$$

where $V_{production,s,t}$ is the shale production volume. We develop scenarios by varying the main source of uncertainty, the methane loss rate ($L_{gathering}$). We use loss rate estimates from a study by Marchese et al. (2015), which is based on recent measurements of gathering facility emissions across the U.S., and specifically use loss rate estimates based on measurements in Pennsylvania, given that there is very high variability across regions; we further adjust central and high estimates to account for heavy-tailed distributions, as described in Alvarez et al. (2018).[65,70] We do not model gathering pipeline leaks, given outdated and otherwise insufficient activity and emissions data.[75] A recent study conducted by Zimmerle et al. (2017) measures gathering pipeline leaks, and while the study suggests that the GHGI may underestimate gathering pipeline leaks, study data were reported as insufficiently representative to develop emission factors.[82]

B4.1.1.5 Processing

We estimate fugitive and vented methane losses ($E_{processing,CH4,t}$) as follows:

$$E_{processing,CH4,t} = \sum_s V_{processing,s,t} \cdot U_{s,t} \cdot C_{production,t} \cdot L_{processing} \quad (40)$$

where $V_{processing,s,t}$ is the natural gas volume received by processing plants (as reported for each state by the EIA) and $C_{production,t}$ is the methane content of natural gas prior to processing. We develop scenarios by varying the main source of uncertainty, the methane loss rate ($L_{processing}$). We use loss rate estimates from a study by Marchese et al. (2015), which is based on recent measurements of gathering facility emissions across the U.S.; we further adjust central and high estimates to account for heavy-tailed distributions, as described in Alvarez et al. (2018).[65,70]

We estimate carbon dioxide vented ($E_{processing,CO2\ venting,t}$) for each year *t* as follows:

$$E_{processing,CO2\ venting,t} = \sum_s V_{processing,s,t} \cdot U_{s,t} \cdot \left(CB_{processing} - CA_{processing}\right) \quad (41)$$

where $CB_{processing}$ is the carbon dioxide content prior to venting and $CA_{processing}$ is the carbon dioxide content after venting to achieve a transmission grade composition of natural gas.

We also estimate carbon dioxide emissions from combustion of plant fuel ($E_{processing,CO2\ combustion,t}$) as follows:

$$E_{processing,CO2\ combustion,s,t} = \sum_s V_{plantfuel,s,t} \cdot U_{s,t} \cdot CE \quad (42)$$

where $V_{plantfuel,s,t}$ is the plant fuel consumed.

B4.1.1.6 Transmission and storage

We develop a top-down estimate of methane losses from transmission and storage infrastructure ($E_{transmission,CH4,t}$) as follows:

$$E_{transmission,CH4,t} = \sum_s \left(V_{distribution,s,t} + V_{transmissionfuel,s,t}\right) \cdot U_{s,t} \cdot C_{transmission} \cdot L_{transmission} \quad (43)$$

where $C_{transmission}$ is the methane content of transmission natural gas. Consistent with assumptions in Tong et al. (2015), we assume that the annual transmission volume is the sum of pipeline and distribution fuel use ($V_{distribution,s,t}$) and volume delivered to end use customers ($V_{transmissionfuel,s,t}$) (as reported by the EIA for each state). We model a range of methane loss rate ($L_{transmission}$) scenarios; we use methane loss rates for the entire transmission and storage segment derived in a study by Zimmerle et al. (2015) that combines recent measurements across the U.S. with pipeline losses estimated in the GHGI.

We also estimate carbon dioxide emissions from combustion of transmission and distribution fuel ($E_{transmission,CO2,t}$) as follows:

$$E_{transmission,CO2,t} = \sum_s V_{transmissionfuel,s,t} \cdot U_{s,t} \cdot EF_{combustion} \tag{44}$$

Given that the EIA reports transmission and distribution fuel consumption ($V_{transmissionfuel,s,t}$) in aggregate, we combine combustion emission estimates across segments.

B4.1.1.7 Distribution

We estimate methane losses from distribution infrastructure ($E_{distribution,CH4,t}$) as follows:

$$E_{distribution,CH4,t} = \sum_s \sum_{type} \big(P_{main,s,t,type} \cdot EF_{main,type} + P_{service,s,t,type} \cdot EF_{service,type} + M_{distribution,s,t,type} \cdot EF_{distribution,type}\big) \cdot U_{s,t} \tag{45}$$

where $P_{main,s,t,type}$ is the miles of mains of each material type and $P_{service,s,t,type}$ is the number of service lines of each material type, based on annual data reported by distribution operators to the Pipeline and Hazardous Materials Safety Administration (PHMSA). $M_{distribution,s,t,type}$ is the number of meters and regulators of each type. We employ emission factors ($EF_{main,type}$, $EF_{service,type}$, and $EF_{distribution,type}$) reported in the GHGI and in a measurement study by Lamb et al. (2015) of 13 urban distribution systems; we model base and high scenarios.[58,67] Rather than use the distribution system loss rate (0.10% to 0.22%) derived in Lamb et al. (2015), which accounts for the distribution of pipelines and meters of each type across the U.S., we use state-specific PHMSA activity data and emission factors by material type over time to reflect recent pipeline replacement efforts in urban areas, such as Pittsburgh, Pennsylvania, and Cincinnati, Ohio. The implied methane loss rate under base case assumptions is 0.21% in 2004 and decreases to 0.15% in 2016.

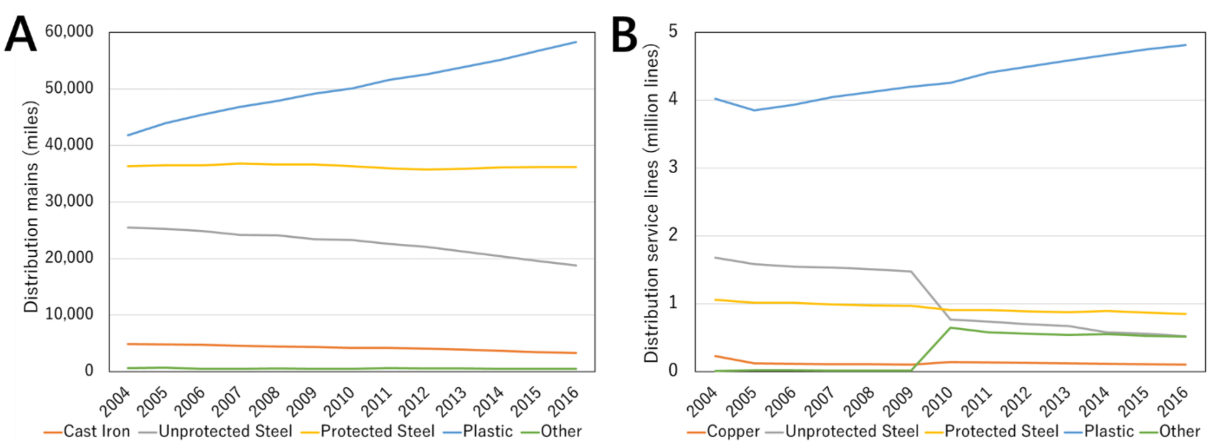

Figure B21. (A) Miles of distribution mains and (B) number of distribution service lines from 2004 to 2016 in Pennsylvania, Ohio, and West Virginia.[76]

B4.1.1.8 End Use

We consider emissions from electric power generation, industrial, commercial, and residential end uses.

A variety of sources report CO_2 emissions from the electric power sector, each with varying sectoral and temporal coverage but high correspondence between reported aggregate emissions, as shown in Table B10 and Table B11. We use the state-level EIA emissions dataset because it provides estimates for all years from 2004 to 2016, includes emissions from all facilities within the electric power sector, and allows for sub-sector segregation to prevent double-counting. We estimate annual emissions from electric power generation from utilities ($E_{electric\ utilities,s,t}$) as follows:

$$E_{electric\ utilities,s,t} = \sum_s RE_{electric\ utilities,s,t} \cdot U_{s,t} \qquad (46)$$

where $RE_{electric\ utilities,s,t}$ is the reported emissions. Electric power generation from large industrial and commercial facilities are included in emission estimates for those segments.

Table B10. Comparison of electric power emissions sector data sources.

Reporting system	Sector Coverage	Subsector Coverage	Years	Spatial Resolution
CEMS	Electric Power Industry (generators >25 MW)	Electric Utility; Industrial Boiler; Pulp & Paper Mill; Iron & Steel	2004 - 2016	plant
eGRID	Electric Power Industry (grid-connected)	-	2004, 2005, 2007, 2009, 2010, 2012, 2014, 2016	plant
EIA	Electric Power Industry	Electric Utility; Industrial; Commercial	2013 - 2016	plant
EIA	Electric Power Industry	Electric Utility; IPP NAICS-22 Non-Cogen; IPP NAICS-22 Cogen; Commercial Cogen; Commercial Non-Cogen; Industrial Cogen; Industrial Non-Cogen	2004 - 2016	state

Table B11. Comparison of 2014 electric utility data for Pennsylvania, Ohio, and West Virginia.

Reporting system	CO2 Emissions (million short tons)	Net Generation (TWh)	Heat Input (million mmBTU)
CEMS	33.6	71.0	574
eGRID	34.0	75.9	573
EIA	34.7	75.8	594

We estimate annual emissions from industrial facilities ($E_{industrial,s,t}$) as follows:

$$E_{industrial,s,t} = \sum_s CE \cdot V_{industrial,s,t} \cdot \frac{HHV_{s,t}}{HHV_{reference}} \cdot U_{s,t} \qquad (47)$$

where $V_{industrial,s,t}$ is the volume of industrial consumption of natural gas, $HHV_{s,t}$ is the higher heating value for each state s and year t, and $HHV_{reference}$ is the reference higher heating value. While the EIA reports emissions for a subset of large industrial facilities comprising less than 1% of industrial fuel consumption, we use the preceding estimation approach for all facilities.

We estimate annual emissions from commercial consumption ($E_{commercial,s,t}$) as follows:

$$E_{commercial,s,t} = \sum_s CE \cdot V_{commercial,s,t} \cdot \frac{HHV_{s,t}}{HHV_{reference}} \cdot U_{s,t} \qquad (48)$$

where $V_{commercial,s,t}$ is the volume of commercial consumption of natural gas, $HHV_{s,t}$ is the higher heating value for each state s and year t, and $HHV_{reference}$ is the reference higher heating value. The EIA reports emissions for a subset of large commercial facilities, which we do not segregate, given that they comprise less than 1% of commercial fuel consumption.

We estimate annual emissions from residential consumption ($E_{residential,s,t}$) as follows:

$$E_{residential,s,t} = \sum_s CE \cdot V_{residential,s,t} \cdot \frac{HHV_{s,t}}{HHV_{reference}} \cdot U_{s,t} \quad (49)$$

where $V_{residential,s,t}$ is the volume of residential consumption of natural gas, $HHV_{s,t}$ is the higher heating value for each state *s* and year *t*, and $HHV_{reference}$ is the reference higher heating value.

B4.1.2 Climate impact model

We use the metric, global temperature change, to assess the temporal trace and cumulative impact of natural gas activity on climate. We additionally monetize the impacts based on the social cost of carbon and methane.

We focus on global temperature change, a global metric which is useful for characterizing spatially and temporally smooth climate responses from emissions of well-mixed GHGs (including CO_2 and CH_4). However, other short-lived chemically active gases (NO_x, CO, and VOCs) that indirectly lead to changes in GHGs, as well as, aerosols and precursors (black carbon, organic carbon, SO_2) react on very different time-scales and have regionally heterogeneous effects; in addition, climate forcings may be both negative and positive, thereby contributing to both warming and cooling.[83] Thus, a global metric provides a rather limited view of potentially nontrivial impacts when multiple-pollutant emission scenarios are considered.[84] While we model emissions of NO_x and VOC emissions as they relate to health impacts, we do not incorporate these emissions into the climate impact model.

We use absolute rather than relative metrics that normalize impacts across species to a reference gas, such as global warming potential (GWP) and global temperature potential (GTP). GWP, the time-integrated radiative forcing due to a pulse emission relative to that of CO_2, has become a default metric because of its simplicity, despite well-documented criticisms of its formulation.[85] GWPs may be inappropriate for evaluating long-term effects, given that they do not take into account that if radiative forcing is applied for a short period, the climate system has time to relax back to equilibrium.[86] A related metric that incorporates additional physical processes and has a less ambiguous interpretation, GTP, is the ratio of a change in global mean surface temperature at a moment in time in response to an emission pulse relative to that of CO_2.[85,87] While these metrics lend themselves to a benefit-cost or cost-effectiveness framing of climate policy decisions[88], they do not facilitate understanding the temporal trace of impacts.

We focus on global temperature change, whereby the relationship between the equilibrium global mean surface temperature response (ΔT) and sustained radiative forcing (RF) has the general form[69]:

$$\Delta T = \lambda \cdot RF \tag{50}$$

where λ is the climate sensitivity parameter. Radiative forcing, typically expressed in units of watts per square meter, is a commonly used metric describing the net change in the energy balance of the climate system resulting from an imposed perturbation, such as emissions of GHGs from the natural gas sector.[71] Often radiative forcing and metrics derivative of radiative forcing are proportional to and describe the relation to the global mean temperature response, with fewer metrics describing the relation to other climate change phenomena. Furthermore, radiative forcing-based metrics do not incorporate climate effects unrelated to radiative forcing, such as the effects of land cover change on evapotranspiration. Thus, while there is utility in quantifying the global mean temperature change resultant of an emission, such an exercise is imperfect given its limited perspective on the factors driving broader climate change.

A more explicit form of the relationship is given by the following convolution of an emission scenario and the average global temperature potential (AGTP)[91,97,98]:

$$\Delta T(t) = \sum_i \int_0^t E_i(s) AGTP_i(t-s)ds \tag{51}$$

where s is the time of an emissions pulse and t is the time of an emissions response. $E_i(s)$ for species i is an emissions scenario (which we formulated and described in Section B4.1). AGTP is the temperature change at time t due to 1-kg emission at $t = 0$, typically expressed in units of K kg^{-1}. The general form of AGTP is given by[93,94,99]:

$$AGTP_i(t) = \int_0^t RF_i(s) R_T(t-s)ds \tag{52}$$

where RF_i is the radiative forcing due to a pulse emission and R_T is the temperature response to a unit of forcing, both of which are parameterized based on more complex models that explicitly include physical and chemical processes.[94] For well-mixed greenhouse gases, the general form of radiative forcing has the form[71]:

$$RF_i = A_i \cdot \exp\left(-\frac{t}{\tau_i}\right) \tag{53}$$

where τ is the perturbation lifetime for the removal of the gas from the atmosphere and A_i is the radiative forcing per unit mass increase in atmospheric concentration (i.e., radiative efficiency). It is assumed that A_i and τ_i are independent of the concentration of greenhouse gases; in reality, there are dependencies and nonlinearities in A_i and τ_i which can lead to systematic biases in the absolute value of these metrics.[71] The climate response function R_T is given by[94]:

$$R_T(t) = \sum_{j=1}^{M} \frac{c_j}{d_j} \exp\left(-\frac{t}{d_j}\right) \tag{54}$$

where c_j are climate sensitivity parameters and d_j are response time parameters for all terms $j = 1, \dots, M$. The absolute global temperature potentials for methane and carbon dioxide are given by[71,100,101]:

$$AGTP_{CH4}(t) = (1 + f_1 + f_2) A_{CH4} \sum_{j=1}^{2} \frac{\tau_{CH4} c_j}{\tau_{CH4} - d_j} \left[\exp\left(-\frac{t}{\tau_{CH4}}\right) - \exp(-\frac{t}{d_j})\right] \tag{55}$$

$$AGTP_{CO2}(t) = A_{CO2} \sum_{j=1}^{2} \left\{a_0 c_j \left[1 - \exp(-\frac{t}{d_j})\right] + \sum_{k=1}^{3} \frac{a_k \tau_{CO2,k} c_j}{\tau_{CO2,k} - d_j} \left[\exp\left(-\frac{t}{\tau_{CO2,k}}\right) - \exp\left(-\frac{t}{d_j}\right)\right]\right\} \tag{56}$$

where there are multiple exponential terms j. a_k for terms $k = 0, \dots, 3$ are coefficients describing the fraction of CO_2 remaining in the atmosphere after a pulse. The additional terms f_1 and f_2 are effects due to ozone and stratospheric water, respectively. The radiative efficiency of carbon dioxide and methane is given by[71]:

$$A_{CO2} = \alpha \left[\frac{\log(C_{0,t}+\Delta C)}{\Delta C}\right] \tag{57}$$

where α is the radiative transfer coefficient. $C_{0,t}$ is the reference CO_2 concentrations and ΔC is the change from the reference concentration (which we evaluate at $\Delta C = 1$ ppm$_v$). For emission pulses in years 2004 to 2016, we use observed CO_2 global atmospheric concentrations reported by the National Oceanic and Atmospheric Administration (NOAA), and for pulses after 2016, we use CO_2 concentrations for four IPCC RCP scenarios. The AGTP functions for 1-kg pulses from years 2004 to 2016 are depicted in Figure B22. We use a Monte Carlo simulation approach to reflect uncertainty in the AGTP values; we assign probability distributions to key input parameters and iterate 10,000 times. The mean and 95% confidence interval of the AGTP are provide in Figure B22.

Parameters, including their definitions, values or distributions, and sources, are provided in Table B12.

Table B12. Climate impact input parameters, definitions, values, units, and sources.

Parameter	Parameter Definition	Parameter Value / Distribution[a]	Units	Source
α	radiative transfer coefficient	5.35	W/m^2	71
A_{CH4}	radiative efficiency of CH_4[b]	3.62×10^{-4}	$W/m^2/ppb_v$	71
a_k	fraction of CO_2 remaining in the atmosphere after a pulse	$k = 0$: 0.2173 $k = 1$: 0.2240 $k = 2$: 0.2824 $k = 3$: 0.2763	unitless	100
$C_{0,t}$	reference mean atmospheric CO_2 concentration[c]	time-varying	ppm_v	102
c_j	climate sensitivity parameters	$j = 1$: Uniform(0.631 ± 0.2) $j = 2$: Uniform(0.429 ± 0.18)	K / W / m^2	90,94
d_j	response time parameters	$j = 1$: Uniform$(8.4 \pm 30\%)$ $j = 2$: Uniform$(409.5 \pm 30\%)$	year	90,94
f_1	ozone effect on CH_4 radiative forcing	Normal(0.5, 0.05)	-/-	71,90,101
f_2	stratospheric water effect on CH_4 radiative forcing	Normal(0.15, 0.05)	-/-	90,101
τ_{CH4}	CH_4 perturbation lifetime	Normal(12.4, 1)	years	90,94
$\tau_{CO2,k}$	CO_2 perturbation lifetime for each term	$\tau_{k=1} = 394.4$ $\tau_{k=2} = 36.54$ $\tau_{k=3} = 4.304$	years	100

a We represent normal distributions as Normal(mean, standard deviation). We represent uniform distributions as Uniform(base value ± error term), where the minimum is the base value minus error term and the maximum is the base value plus the error term.

b We convert from values given per ppm_v to kg, assuming the mean molecular weight of air is 28.97 kg/kmol, the molecular weight of methane is 16.04 g/mol, and the total mass of the atmosphere is 5.1352 x 1018 kg.

c We convert from values given in ppm_v to kg, assuming the mean molecular weight of air is 28.97 kg/kmol, the molecular weight of methane is 44.01 g/mol, and the total mass of the atmosphere is 5.1352×10^{18} kg.

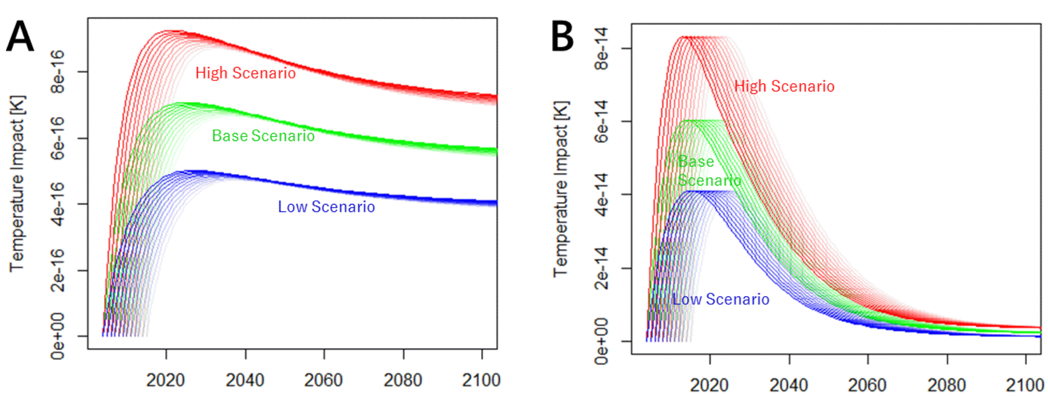

Figure B22. Absolute global temperature potential from 2004 to 2100 for 1-kg pulses in years 2004 to 2016. (A) CO_2 and (B) CH_4. Base scenario (based on simulated mean) indicated in green. Low scenario (based on simulated lower 95% confidence interval) indicated in blue. High scenario (based on simulated upper 95% confidence interval) indicated in red.

B4.2 Additional results

The following are additional results, including emissions, temperature impacts, and monetized damages.

Estimated emissions and the percent attribution of emissions across the supply chain are reasonably consistent with other studies. Table B13 provides cumulative emissions by supply chain segment and process and the percent attribution for each segment. Table B14 includes a comparison between this study, and the Alvarez et al. (2018) study and 2015 US GHGI inventory. The percent attribution of methane emissions across segments between this study and the Alvarez et al. (2018) study are similar, with a majority of emissions associated with production. However, the methane loss rates estimated in this study, although similar to the US GHGI and the source-based estimates in the Alvarez study, are lower than the site-based estimates in the Alvarez study. This may be attributed to differences in source inclusion, estimation methods, regional emission factors, and study scope (i.e., Appalachian basin versus U.S., shale gas versus O&G sector). Although we estimate a declining loss rate over time, this trend is largely a byproduct of estimation methods and increasing well productivity (Figure B23).

Table B13. Cumulative emissions and percent attribution for each segment and process across the supply chain. Base scenario emission estimates with low and high scenario estimates provided in the parentheses. Percent attribution based on base scenario estimates.

Segment / process	2004 to 2016 Cumulative Emissions			
	CO_2		CH_4	
	Emissions (mmt)	%	Emissions (mmt)	%
Preproduction	**15.5 (10.2 – 20.6)**	**2%**	**0.02 (0.01 - 0.05)**	**0%**
Well pad preparation	4.17 (3.68 - 4.42)	1%	-	-
Drilling	4.79 (3.44 - 6.14)	1%	-	-
Hydraulic fracturing	5.65 (2.82 - 8.46)	1%	-	-
Well completion	0.88 (0.26 - 1.61)	0%	0.02 (0.01 - 0.05)	0%
Production	**50.0**	**7%**	**3.69 (2.10 – 5.73)**	**63%**
Lease fuel consumption	50.0	7%	-	-
Fugitive / vented losses	-	-	3.69 (2.10 – 5.73)	63%
Gathering	**-**	**-**	**0.91 (0.65 - 1.17)**	**15%**
Fugitive / vented losses	-	-	0.91 (0.65 - 1.17)	15%
Processing	**12.2**	**2%**	**0.25 (0.22 - 0.27)**	**4%**
Plant fuel consumption	3.45	1%	-	-
Fugitive / vented losses	8.70	1%	0.25 (0.22 – 0.27)	4%
Transmission and storage	**26.0**	**4%**	**0.69 (0.55 - 0.89)**	**12%**
Fuel consumption	26.0	4%	-	-
Facility / pipeline fugitive losses	-	-	0.69 (0.55 - 0.89)	12%
Distribution	**-**	**-**	**0.30 (0.30 - 0.64)**	**5%**
Fugitive losses	-	-	0.30 (0.30 - 0.64)	5%
End use	**571**	**85%**	**-**	**-**
Electricity generation	175	26%	-	-
Industrial use	146	22%	-	-
Commercial use	96	14%	-	-
Residential use	152	23%	-	-
Total Supply Chain	**675 (669 - 680)**	**100%**	**5.87 (3.83 – 8.75)**	**100%**

Table B14. Comparison of 2015 methane emissions, percent emissions, and loss rates across this study, the Alvarez et al. (2008) study[72], and the U.S. GHGI.

Segment	This work		Alvarez et al. (2018) (site-based)[a]		Alvarez et al. (2018) (source-based)[b]		U.S. Greenhouse Gas Inventory	
	Appalachian basin		U.S.		U.S.		U.S.	
	Emissions [CH_4 mmt]	% Emissions	Emissions [CH_4 mmt]	% Emissions	Emissions [CH_4 mmt]	% Emissions	Emissions [CH_4 mmt]	% Emissions
Preproduction	0 (0-0.01)	0%	0.09 (0.08-0.12)[c]	1%	0.09 (0.08-0.12)[c]	1%	0.10[c]	1%
Production	0.76 (0.43-1.18)	61%	7.2 (5.6-9.1)[d]	56%	2.8 (2.7-2.9)[d]	33%	3.10[d]	40%
Gathering	0.22 (0.16-0.28)	17%	2.6 (2.4-3.2)	20%	2.6 (2.4-3.2)	31%	2.30	30%
Processing	0.07 (0.07-0.08)	6%	0.72 (0.65-0.92)	6%	0.72 (0.65-0.92)	9%	0.45	6%
Transmission and storage	0.14 (0.11-0.18)	11%	1.8 (1.6-2.1)	14%	1.8 (1.6-2.1)	21%	1.30	17%
Distribution	0.06 (0.06-0.12)	5%	0.44 (0.22-0.95)	3%	0.44 (0.22-0.95)	5%	0.44	6%
Total supply chain	1.25 (0.82-1.84)	100%	12.85 (10.6-16.4)	100%	8.45 (7.65-10.2)	100%	7.69	100%
Supply chain loss rate (%)	1.16 (0.76 – 1.71)[e]		2.3 (2.0 - 2.7)[f]		1.48 (1.34 – 1.78)[g]		1.34[g]	

a Values are bottom-up, site-based estimates derived in Alvarez et al. (2018). They include emissions across the oil and natural gas supply chain and are not exclusive of shale gas.

b Values are bottom-up, source-based estimates derived in Alvarez et al. (2018). They include emissions across the oil and natural gas supply chain and are not exclusive of shale gas.

c Includes emissions from both completions and workovers.

d Includes emissions from routine operations.

e Loss rate is estimated as a percentage of methane produced. Assumes state-reported shale production (refer to section B2.2), and time-varying methane content for the Northeast (~83-84%).[5]

f Reported loss rate, as a function of total methane produced (33 tcf, with average CH_4 content of 90 vol%), derived in Alvarez et al. (2018). The range represents the reported 95% confidence interval. That study also reported a loss rate of 2.9%, as a percentage of total methane delivered (25 tcf/y NG delivered, assuming an average CH4 content in NG of 95% by volume).

g Estimated loss rate, as a function of total methane produced (33 tcf, with average CH_4 content of 90 vol%), similar to the approach used in Alvarez et al. (2018).

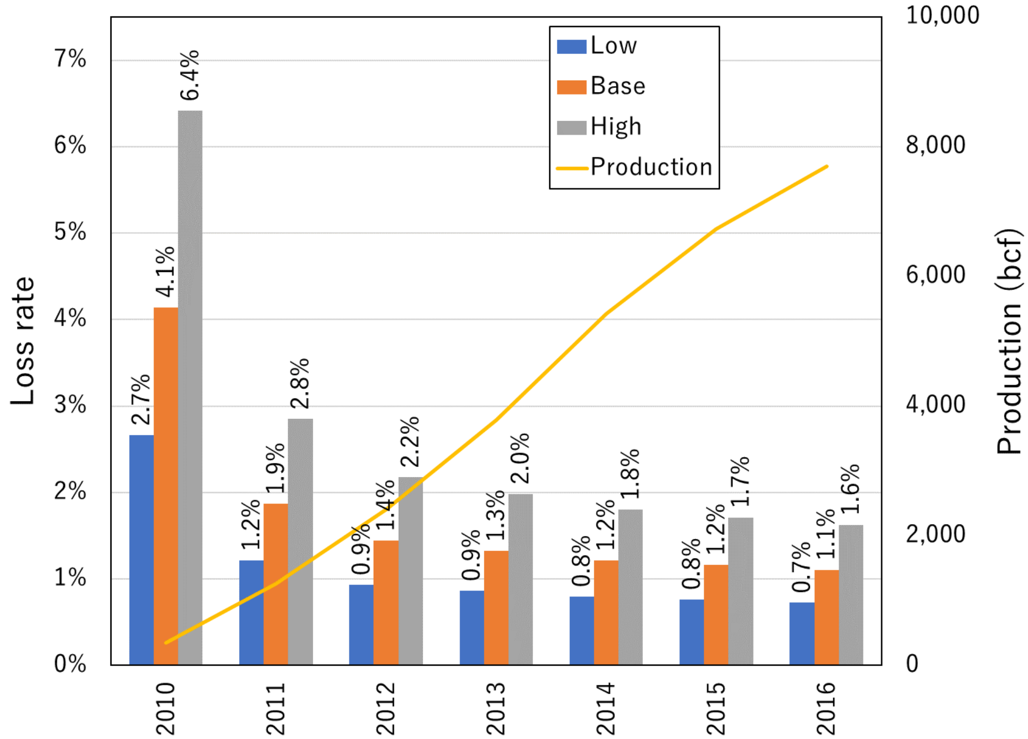

Figure B23. Methane loss rates over time. Blue, orange, and gray bars represent base, low, and high loss rate estimates, respectively, based on emission scenario estimates. Loss rates assume state-reported production and time-varying CH_4 content (~83-84%). Yellow line represents production over time.

Table B15. Cumulative climate change temperature impacts.

Time-Integrated Cumulative Temperature Impact from 2004 (Kelvin-years)	
to 2016	0.001 (0-0.001)
to 2050	0.023 (0.014-0.037)
to 2100	0.045 (0.029-0.067)
to 2200	0.082 (0.055-0.116)
Instantaneous Temperature Impact (Kelvin)	
2016	0.0003 (0.0002-0.0006)
2050	0.0005 (0.0004-0.0008)
2100	0.0004 (0.0003-0.0005)
2200	0.0004 (0.0003-0.0005)

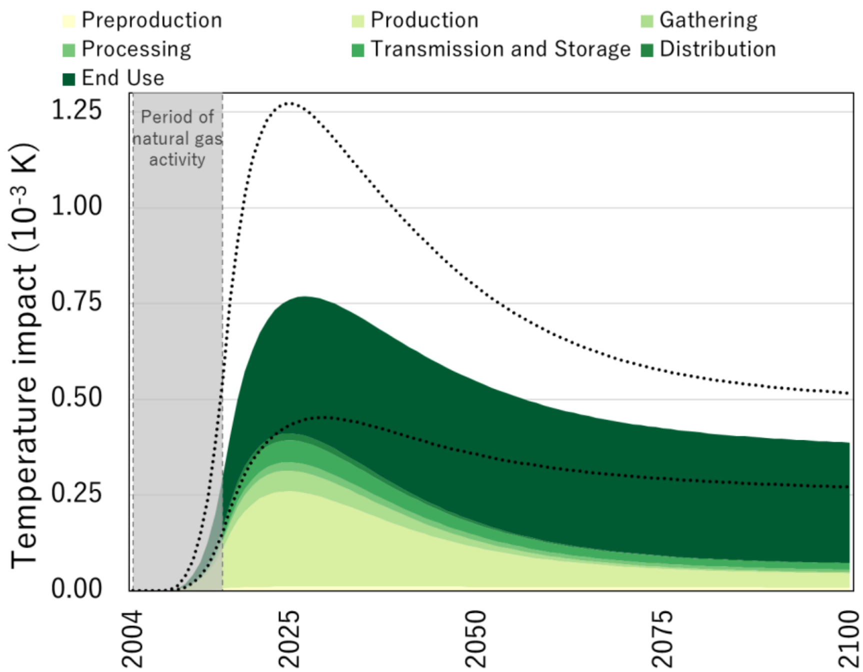

Figure B24. Annual temperature impact from sources within Appalachia indicating contributions from each segment of the supply chain. Dotted black lines depict temperature impact under low and high scenarios.

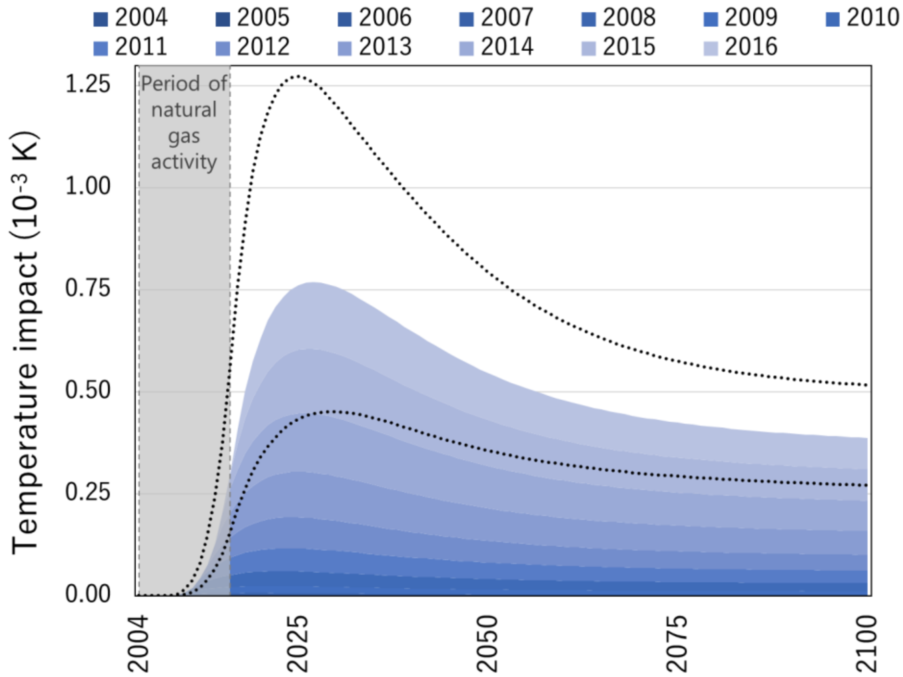

Figure B25. Annual temperature impact from sources within Appalachia indicating contributions from each year of natural gas activity. Dotted black lines depict temperature impact under low and high scenarios.

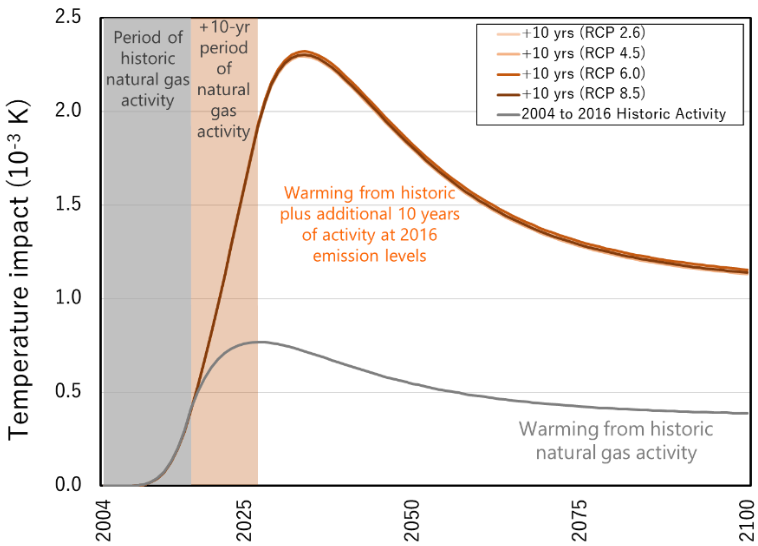

Figure B26. Annual temperature impact assuming an additional 10 years of natural gas activity (and emissions) at 2016 levels. Four Intergovernmental Panel on Climate Change (IPCC) Representative Climate Pathways (RCP) are modeled. Annual temperature impact from sources within Pennsylvania, Ohio, and West Virginia under baseline scenario assumptions.

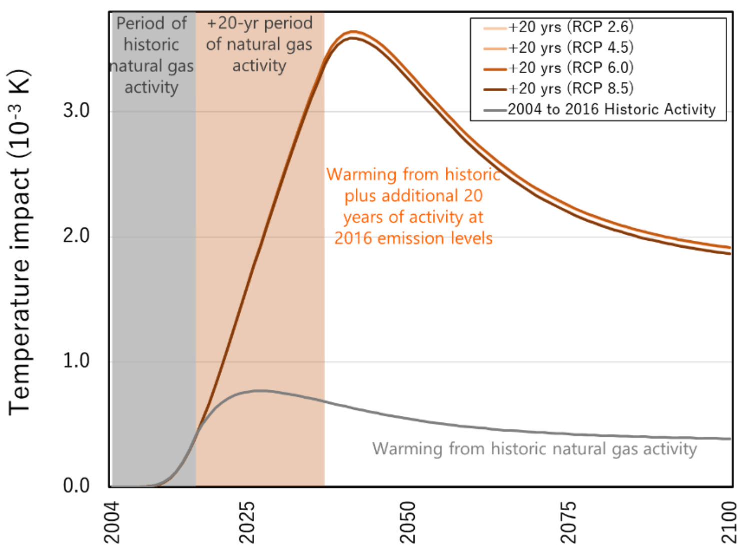

Figure B27. Annual temperature impact assuming an additional 20 years of natural gas activity (and emissions) at 2016 levels. Four IPCC RCP are modeled. Annual temperature impact from sources within Pennsylvania, Ohio, and West Virginia under baseline scenario assumptions.

B5 Employment model

We estimate county-level employment effects from shale gas activity in producing counties from 2004 to 2016. We build upon previous studies by expanding the geographic scope to include Pennsylvania, Ohio, West Virginia, and New York, as well as, incorporating additional years of data that may facilitate inference regarding learning within the industry that induces lower employment effects over time. In addition, we utilize a range of natural gas activity factors, including production, producing wells, and spud wells, that allows for the disaggregation of employment effects associated with initial well development and ongoing production.

Section B5.1 includes a brief review of labor market studies related to unconventional natural gas development. Section B5.2 describes the data compiled into a panel dataset and provides an overview of the empirical approach used to estimate the marginal employment effects from shale gas activity. We also describe the variable selection process in which we regress employment (and variants of the dependent variable) on various population, economic, and natural gas development explanatory variables and interaction terms. We additionally perform cross validation to facilitate comparisons across models, and bootstrapping to capture the uncertainty around the employment effects of natural gas activity. Combining results from the fixed effects modeling with natural gas activity, we develop estimates of aggregate employment effects over time.

B5.1 Background

Several empirical studies demonstrate that natural gas activity may impact local labor demand within the natural gas sector, have spillover effects into the non-resource economy, and alter the distribution of income, poverty rates, and educational attainment. A study by Weber (2012), focusing on shale plays in Colorado, Texas, and Wyoming that are more mature than the Marcellus and Utica formations, shows that the boom in drilling increased total employment in shale boom counties by 12% over an 8-year period. Maniloff and Mastromonaco (2015) similarly find that the boom caused a 24% increase in employment in boom counties from 2000 to 2010. On an annual basis, the shale boom is projected to have increased employment growth by 1-2% in boom counties across the United States (Fetzer 2014; Weinstein 2014). One million dollars in production growth created 2.35 total jobs in the county where production occurred, and an additional billion cubic feet of natural gas production created 7.3 to 18.5 total jobs (Brown 2014, Weber 2012, Weber 2014). Within the Marcellus, Komarek (2016) and Wrenn (2015) find the effects on overall employment range from a 3-7% increase for boom counties. Parades (2015) uses a panel-fixed effects regression approach to assess employment and income effects in the Marcellus from 2004 to 2011; results suggest statistically significant employment effects in the short-run, with each active well generating 6 to 16 jobs (Parades et al. 2015). These ex-post empirics of job creation in the Marcellus show that ex-ante estimates are greatly overstated. For example, one projection using IMPLAN, an input-out model, projected that the Marcellus industry in Pennsylvania would support more than 44,000 jobs in 2009 and 200,000 jobs in 2020 [103], whereas empirical evidence suggests that about 2000 jobs were created in 2009[104].

Natural gas development may generate spillover effects outside of the resource sector (Corden and Neary 1982). For example, drilling a natural gas well may require use of larger amounts of cement, which may lead to increased sales for cement companies. Additional income by natural gas and cement workers can have a ripple effect, with more money being spent on locally provided goods and services, leading to more jobs in those industries. It is estimated that each gas-related mining job is associated with 1.4 additional nonmining jobs in the county where production occurred, indicating that natural gas development has largely neutral effect on resource dependence as measured by employment (Weber 2014). Weinstein and Partridge (2011) find that each mining job from Marcellus development in Pennsylvania likely creates one additional job in the economy.

Growth in employment does not imply that median income will increase or poverty rates will decrease. Local labor market evidence linking natural resource booms to poverty and inequality are mixed. The distribution of the gains depends on the skills of local residents and where they fall in the income distribution, the extent of integration between local and regional labor markets, and the extent of spillover [105]. If income changes are equally distributed across the population, inequality would remain unchanged, while poverty would decline in a boom and rise in a bust. A majority of studies show that booms lower the poverty rate, at least in the short-run [104]. An analogue to the boom-and-bust cycle of natural gas is coal mining in the Appalachian region in which Black et al. (2005) found that the 1970s boom decreased poverty, but the 1980s bust reversed this reduction.

Most studies focus on short-term effects during a boom, with few studies analyzing the long-term effects of natural gas production over the boom-and-bust cycle and the potential negative spillovers into industries with high long-term growth potential (e.g., manufacturing) (Rodriguez and Sachs 1999, Corden and Neary 1983). Parades (2015) find weak evidence of long-term employment effects in the Marcellus, and Weber (2012) reports statistically discernible crowding out of manufacturing.

B5.2 Methods

B5.2.1 Data

Employment and Earnings

We use the Bureau of Economic Analysis (BEA) Local Area Personal Income and Employment (LAPI) datasets, which provide county-level estimates from 2005 to 2015.[1] The datasets are mainly derived from administrative records data of various federal and state government social insurance programs and tax codes. These data originate from the recipients of the income or from the payer of the income.

We use total employment estimates, which consist of wage and salary employment and proprietor's income. Wage and salary employment, as defined by the BEA, measures the average annual number of full- and part-time jobs in each area by place of work, including all jobs for which wages and salaries are paid.

We also estimate the share of total earnings from different sectors (i.e., farm, construction, manufacturing, retail, and mining), based on total and sector-level earnings. Earnings, as defined by the BEA, consist of compensation of employees and proprietors' income.

Population and Population Density

We use the BEA estimates of county-level population[2], and combining those data with county-level land area values from the Census Bureau, we estimate population density.[3]

Employment rate

We estimate the employment rate, which is the number of employed divided by the total labor force. The number employed and total labor force, as reported by the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics[4], are based on the Community Population Survey and unemployment insurance disaggregated from state-level statistics to the county-level; the processing of this data may not accurately reflect the county-level employed and labor force, thus limiting its utility for regression. The total labor force are those persons greater 16 years old, and the number employed are any persons that have worked (not necessarily full-time) or are on leave, vacation, illness, etc. during the survey.

Employment-to-population ratio

---

[1] BEA LAPI "Table CA1 Personal Income Summary_Personal Income, Population, Per Capita Personal Income", "Table CA4 Personal Income and Employment by Major Component", and "Table CA5N Personal Income by Major Component and Earnings by NAICS Industry" available at <https://www.bea.gov/>.

[2]BEA LAPI "Table CA1 Personal Income Summary_Personal Income, Population, Per Capita Personal Income" available at <https://www.bea.gov/>.

[3] US Census Bureau "Population, Housing Units, Area, and Density: 2010 - State -- County / County Equivalent more information 2010 Census Summary File 1" available at <https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk>.

[4]BLS Local Area Unemployment Statistics available at <https://www.bls.gov/lau/#cntyaa>.

The employment-to-population ratio is the total employment divided by the population (as reported in BEA LAPI). In some instances, the employment-to-population ratio is greater than, demonstrating the limitations of the underlying data.

Shale counties

We identify shale counties based on United States Geological Service (USGS) designations for Marcellus and Utica shale subunits. If at least 10% of a county's area is within the footprint of the shale subunits, we classify it as a shale county. Note that several producing shale formations other than the Marcellus and Utica exist within the Appalachian basin; however, they are largely within the footprint of the Marcellus and Utica formations, thus, we do not identify them separately. The shale classifier is used to subset the panel dataset for model fitting. Several other empirical studies, including Parades (2015), partition or exclude counties which are not within the Marcellus (or respective shale formation).

Nonmetropolitan counties

We identify nonmetropolitan counties based on a cross-county comparison of 2010 population estimates, whereby we classify counties as metropolitan if they are within the top 10 percentile based on population. Of the 272 counties within Pennsylvania, Ohio, West Virginia, and New York, 28 counties are classified as metropolitan, two of which produced (including Allegheny County, Pennsylvania, which had 49 producing wells in 2015, and Stark County, Ohio, which had only 2 producing wells in 2015). The metropolitan classifier is used to subset the panel dataset for model fitting. Including only nonmetropolitan counties creates a more homogenous sample, precluding counties with large cities from excessively influencing estimates from a linear model (Weber 2014). Moreover, nonmetropolitan counties with thin labor markets are presumably the population of interest.

B5.2.2 Modeling specification

County-level descriptive statistics of the panel data are provided in Table B16. We begin with a dataset consisting of all 272 counties within Pennsylvania, Ohio, West Virginia, and New York, of which 93 counties produced unconventional natural gas and 195 counties are within (or partially within) the Marcellus and/or Utica shale plays. The full dataset includes 2972 observations over the period 2005 to 2015, accounting for observations that were removed due to missing data. Sector-level earnings were obscured for some counties in some years, so we exclude these observations. In addition, Pennsylvania production (and resultingly well count) data is unavailable for 2010 due to reporting inconsistencies, further reducing the number of observations. The data subsets of shale, nonmetropolitan, and the intersection of shale and nonmetropolitan counties include 2096, 2654, and 2013 observations, respectively. Unless otherwise specified, we use the full dataset to fit the following models.

As part of the model selection process, we begin by fitting and comparing five types of models that can be used with panel datasets: pooled ordinary least squares (OLS), OLS fixed effects, general least squares (GLS) fixed effects, first-difference, and random effects. Using panel data allows us to control for several types of unobserved heterogeneity that could potentially confound the estimated effect of natural gas activity. The first specification measuring the contemporaneous effect of natural gas wells using pooled OLS is as follows:

$$Y_{ct} = \beta ProducingWells_{ct} + X_{ct}\theta + \alpha + \varepsilon_{ct} \tag{58}$$

Where Y_{ct} is a measure employment for each county c and year t, and α is a constant coefficient, assuming no fixed heterogeneity across counties or years. $ProducingWells_{ct}$ is the number of producing unconventional natural gas wells, and β can be interpreted as the average change in employment attributable to another producing well in the county. X_{ct} are labor market and demographic control variables, which are as follows:

$$X_{ct} = \{population_{ct}, populationdensity_{ct}, retailshareearnings_{ct}, \\ constructionshareearnings_{ct}, manufacturingshareearnings_{ct}, retailshareearnings_{ct}\} \tag{59}$$

These control variables are similar to those used in Paredes (2015) and Weber (2014). θ are the coefficients associated with the control variables. ε_{ct} is the random error term given by:

$$\varepsilon_{ct} = \mu_c + \lambda_t + \varepsilon_{ct} \tag{60}$$

where μ_c is the county error component, λ_t is the time error component, and ϵ_{ct} is the idiosyncratic error. The composition of the error term depends on whether the model incorporates county, time, or both types of effects. To test for the significance of county, year, and two-way effects, we use the Lagrange multiplier test for pooled OLS, finding that county effects are significant, but time effects are not.[5] We note that, as shown in Table B17, the average change in employment from a producing well β is not significant.

We then specify fixed effects models, including alternatively county, year, and two-way effects, as follows:

$$Y_{ct} = \beta ProducingWells_{ct} + X_{ct}\theta + \alpha_c + \gamma_t + \varepsilon_{ct} \tag{61}$$

[5] We employ the King/Wu and Breusch/Pagan statistics tests for one- and two-way unbalanced panel data as described in Baltagi (2013) and implemented in the plm R package. The following are the test statistics and p-values for each model form, where the null hypothesis is :

	King/Wu		**Breusch/Pagan**	
	Test Statistic	**p-value**	**Test Statistic**	**p-value**
Two-way	28.158	< 2.2e-16	16767	< 2.2e-16
County effects	129.49	< 2.2e-16	16767	< 2.2e-16
Year effects	0.36832	0.3563	0.13566	0.7126

where α_c are county fixed effects and γ_t are year fixed effects. The fixed effects model assumes that the county error component is correlated with the regressors. Comparing the two-way pooled OLS and fixed effects models using, we find (as anticipated) that the fixed effects model provides a better model fit.[6] For the fixed effects model, we apply OLS to transformed data, which provides consistent estimates of β, but not efficient estimates if the error is serially correlated. We find that the error is serially correlated, based on tests for serial correlation.[7] To control for serial correlation, the standard errors are corrected by clustering following the approach in Arellano (1987).

We also fit a fixed effects model using generalized least squares. A comparison of the OLS and GLS fixed effects models (accounting for county fixed effects), as provided in Table B19, suggests that the models are not substantially different; thus, further models use OLS, which allows for inclusion of additional control variables and accounting for both county and year fixed effects.

We additionally fit a first-difference model, which removes time-invariant individual components by lagging the model and subtracting:

$$\Delta Y_{ct} = \beta \Delta ProducingWells_{ct} + \Delta X_{ct}\theta + \Delta\varepsilon_{ct} \qquad (62)$$

The first-difference model is efficient and usually preferred if the error is persistent over time because $\Delta\varepsilon_{ct}$ will be serially uncorrelated. A summary of the coefficients, standard errors, and model fit between the fixed effects and first-difference models are summarized in Table B17.

For the proceeding model specifications, we use OLS fixed effects, including county and year fixed effects, and controlling for serial correlation by clustering standard errors by county.

[6] We employ the F-test for two-way effects, used for comparing within and pooling models, and implemented in the plm R package. The following are the test statistic and p-value, where the null hypothesis is that there are no significant fixed effects: F = 459.54, df1 = 268, df2 = 2770, p-value < 2.2e-16

[7] We employ the Breusch-Godfrey/Wooldridge test for serial correlation in panel models as described in Wooldridge (2010) and implemented in the plm R package. The following is the test statistic, where the null hypothesis is there is no serial correlation: chisq = 750.25, df = 1, p-value < 2.2e-16

We employ the Wooldridge's test for serial correlation, which is applicable to fixed effects panels, in particular to short panels with small T and large n. The following is the test statistic, where the null hypothesis is there is no serial correlation: F = 24.78, df1 = 1, df2 = 2952, p-value = 6.796e-07

Table B16. County-level descriptive statistics for different county subsets over the period 2004 to 2016.

Variable	Units	Pennsylvania, Ohio, West Virginia, and New York counties (269)		Pennsylvania, Ohio, West Virginia, and New York shale counties (193)		Pennsylvania, Ohio, West Virginia, and New York nonmetropolitan counties (241)		Pennsylvania, Ohio, West Virginia, and New York shale and nonmetropolitan counties (186)	
		Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Employment	jobs	101,979	244,794	61,785	125,765	42,303	44,139	40,591	43,052
Income	$'000	7,827,151	19,421,640	4,094,918	7,789,432	3,002,160	3,183,225	2,796,412	2,865,572
Unconventional production	Mmcf	7,727	59,887	10,951	71,048	8,599	63,303	11,330	72,455
Unconventional producing wells	wells	15	80	21	95	17	85	22	97
Unconventional spud wells	wells	3	18	5	21	4	19	5	21
Earnings	$'000	5,852,333	22,304,600	2,929,224	6,904,864	1,895,056	2,283,443	1,788,054	2,197,252
Farm share of earnings	$'000	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02
Construction share of earnings	$'000	0.07	0.03	0.07	0.04	0.07	0.03	0.07	0.04
Manufacturing share of earnings	$'000	0.16	0.11	0.15	0.10	0.17	0.11	0.16	0.10
Retail share of earnings	$'000	0.07	0.02	0.07	0.02	0.07	0.02	0.07	0.02
Population	persons	176,021	335,726	109,364	180,667	82,419	74,979	78,762	70,167
Population density	persons per mile^2	971	5,584	200	330	164	195	148	146
Employment rate	%	93	2	93	2	93	2	93	2
Employment ratio	jobs per person	0.49	0.14	0.48	0.12	0.48	0.11	0.47	0.11

Table B17. Employment effects for pooled OLS, first-differenced, and fixed effects (implemented using OLS).

<table>
<tr><th></th><th colspan="4">(1) County and year fixed effects (OLS)</th><th colspan="4">(2) Pooled OLS</th><th colspan="4">(3) First-differenced</th></tr>
<tr><th>Variables</th><th>Estimate</th><th>Std. Error</th><th>P(>|t|)</th><th></th><th>Estimate</th><th>Std. Error</th><th>P(>|t|)</th><th></th><th>Estimate</th><th>Std. Error</th><th>P(>|t|)</th><th></th></tr>
<tr><td>Producing wells</td><td>5.36</td><td>1.86</td><td>0.004</td><td>**</td><td>2.28</td><td>2.85</td><td>0.424</td><td></td><td>5.66</td><td>1.28</td><td>0.000</td><td>***</td></tr>
<tr><td>Population</td><td>0.47</td><td>0.09</td><td>0.000</td><td>***</td><td>0.63</td><td>0.00</td><td>< 2e-16</td><td>***</td><td>0.43</td><td>0.03</td><td>< 2e-16</td><td>***</td></tr>
<tr><td>Population density</td><td>88.91</td><td>31.63</td><td>0.005</td><td>**</td><td>-15.01</td><td>0.75</td><td>< 2e-16</td><td>***</td><td>106.73</td><td>7.42</td><td>< 2e-16</td><td>***</td></tr>
<tr><td>Farm earnings share</td><td>6.88E+03</td><td>3.30E+03</td><td>0.037</td><td>*</td><td>-1.67E+03</td><td>9.98E+03</td><td>0.867</td><td></td><td>-1.26E+03</td><td>2.45E+03</td><td>0.607</td><td></td></tr>
<tr><td>Construction earnings share</td><td>8.08E+03</td><td>3.79E+03</td><td>0.033</td><td>*</td><td>-1.26E+04</td><td>6.48E+03</td><td>0.052</td><td>.</td><td>9.77E+03</td><td>2.05E+03</td><td>0.000</td><td>***</td></tr>
<tr><td>Manufacturing earnings share</td><td>7.26E+03</td><td>3.24E+03</td><td>0.025</td><td>*</td><td>1.05E+02</td><td>2.01E+03</td><td>0.958</td><td></td><td>1.57E+04</td><td>1.86E+03</td><td>< 2e-16</td><td>***</td></tr>
<tr><td>Retail earnings share</td><td>-3.77E+04</td><td>1.25E+04</td><td>0.003</td><td>**</td><td>-8.76E+04</td><td>7.83E+03</td><td>< 2e-16</td><td>***</td><td>-3.13E+04</td><td>5.96E+03</td><td>0.000</td><td>***</td></tr>
<tr><td>Number of counties</td><td colspan="4">241</td><td colspan="4">241</td><td colspan="4">241</td></tr>
<tr><td>Years</td><td colspan="4">2004-2016</td><td colspan="4">2004-2016</td><td colspan="4">2004-2016</td></tr>
<tr><td>Sample</td><td colspan="4">Nonmetropolitan counties</td><td colspan="4">Nonmetropolitan counties</td><td colspan="4">Nonmetropolitan counties</td></tr>
<tr><td>R-squared</td><td colspan="4">0.55</td><td colspan="4">0.97</td><td colspan="4">0.28</td></tr>
<tr><td>Adjusted R-squared</td><td colspan="4">0.51</td><td colspan="4">0.97</td><td colspan="4">0.28</td></tr>
<tr><td colspan="13">Significance codes: ***p<0.001, **p<0.01, *p<0.05, .p<0.1</td></tr>
</table>

Table B18. Fixed effects assuming two-way, county, and time fixed effects.

	(1) County and year fixed effects (OLS)				(4) County fixed effects (OLS)				(5) Year fixed effects (OLS)			
Variables	Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)	
Producing wells	5.36	1.86	0.004	**	6.08	1.89	0.001	**	-15.76	6.64	0.018	*
Population	0.47	0.09	0.000	***	0.46	0.09	0.000	***	0.42	0.04	< 2e-16	***
Population density	88.91	31.63	0.005	**	92.34	31.69	0.004	**	21.36	1.19	< 2e-16	***
Farm earnings share	6.88E+03	3.30E+03	0.037	*	-8.06E+03	3.33E+03	0.016	*	-1.13E+05	5.76E+04	0.050	*
Construction earnings share	8.08E+03	3.79E+03	0.033	*	1.46E+04	4.26E+03	0.001	***	-2.09E+05	4.37E+04	0.000	***
Manufacturing earnings share	7.26E+03	3.24E+03	0.025	*	1.41E+04	3.86E+03	0.000	***	-1.06E+05	2.51E+04	0.000	***
Retail earnings share	-3.77E+04	1.25E+04	0.003	**	-3.73E+04	1.10E+04	0.001	***	-7.26E+05	1.63E+05	0.000	***
Number of counties	241				241				241			
Years	2004-2016				2004-2016				2004-2016			
Sample	Nonmetropolitan counties				Nonmetropolitan counties				Nonmetropolitan counties			
R-squared	0.55				0.54				0.85			
Adjusted R-squared	0.51				0.49				0.83			
Significance codes: ***p<0.001, **p<0.01, *p<0.05, .p<0.1												

Table B19. Employment effects for fixed effects implemented using OLS and fixed effects implemented using GLS.

	(4) County fixed effects (OLS)				(6) County fixed effects (GLS)			
Variables	Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)	
Producing wells	6.08	1.89	0.001	**	5.66E+00	1.28	0.000	***
Population	0.46	0.09	0.000	***	4.28E-01	0.03	< 2e-16	***
Population density	92.34	31.69	0.004	**	1.07E+02	7.42E+00	< 2e-16	***
Farm earnings share	-8.06E+03	3.33E+03	0.016	*	-1.26E+03	2446.80	0.607	
Construction earnings share	1.46E+04	4.26E+03	0.001	***	9.77E+03	2048.70	0.000	***
Manufacturing earnings share	1.41E+04	3.86E+03	0.000	***	1.57E+04	1856.10	< 2e-16	***
Retail earnings share	-3.73E+04	1.10E+04	0.001	***	-3.13E+04	5963.40	0.000	***
Number of counties	241				241			
Years	2004-2016				2004-2016			
Sample	Nonmetropolitan counties				Nonmetropolitan counties			
R-squared	0.54				0.999			
Adjusted R-squared	0.49							
Significance codes: ***p<0.001, **p<0.01, *p<0.05, .p<0.1								

Data subsets

We specify fixed effects models of the form described by equation (4) for various subsets of the full panel dataset, partitioned based on the classification of counties as nonmetropolitan and/or shale. Several other empirical studies, including Parades (2015), partition or exclude counties which are not within the Marcellus (or respective shale formation). Including only nonmetropolitan counties creates a more homogenous sample, precluding counties with large cities from excessively influencing estimates from a linear model (Weber 2014). Moreover, nonmetropolitan counties with thin labor markets are presumably the population of interest.

We find that the effect sizes on the producing well variable are fairly consistent across models, as shown in Table B20. The model fit using the shale county data subset has an overall model fit similar to that for model based on the full panel dataset, with a lower standard error for the producing well coefficient. The other models fit using data subsets do not provide as good of a fit as the model based on the full panel dataset.

Table B20. Employment effects for different subsets of data.

	(1) Nonmetropolitan counties				(7) All counties				(8) Shale counties				(9) Shale and nonmetropolitan counties			
Variables	Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)	
Producing wells	5.36	1.86	0.004	**	3.52	2.14	0.101		4.75	1.97	0.016	*	5.93	1.83	0.001	**
Population	0.47	0.09	0.000	***	0.47	0.16	0.003	**	0.87	0.44	0.046	*	0.64	0.17	0.000	***
Population density	88.91	31.63	0.005	**	74.80	21.10	0.000	***	-179.61	232.38	0.440		-43.13	85.39	0.614	
Farm earnings share	6.88E+03	3.30E+03	0.037	*	2.70E+04	9.71E+03	0.005	**	3.55E+03	4.56E+03	0.436		-1.24E+02	2.96E+03	0.967	
Construction earnings share	8.08E+03	3.79E+03	0.033	*	1.37E+04	6.89E+03	0.047	*	4.34E+03	5.10E+03	0.395		5.28E+03	2.60E+03	0.042	*
Manufacturing earnings share	7.26E+03	3.24E+03	0.025	*	9.06E+03	4.78E+03	0.058	.	8.99E+03	3.74E+03	0.016	*	9.38E+03	3.38E+03	0.006	**
Retail earnings share	-3.77E+04	1.25E+04	0.003	**	-3.37E+04	3.19E+04	0.291		-2.37E+04	1.62E+04	0.143		-1.44E+04	8.62E+03	0.096	.
Number of counties	241				269				193				186			
Years	2004-2016				2004-2016				2004-2016				2004-2016			
Sample	Nonmetropolitan counties				All counties				Shale counties				Shale and nonmetropolitan counties			
R-squared	0.55				0.76				0.45				0.56			
Adjusted R-squared	0.51				0.73				0.40				0.51			
Significance codes: ***p<0.001, **p<0.01, *p<0.05, .p<0.1																

B5.2.3 Variable selection

We specify various fixed effects models with alternative variables, including i) lag and lead variables to capture the dynamic effect of natural gas development, ii) natural gas activity variables (i.e., producing wells, production, and spud wells), iii) natural gas development and time interactions, and iv) employment-to-population ratio and employment rate (as alternative dependent variables).

The following specification includes lag and lead variables to assess the dynamic effect of natural gas development:

$$Y_{ct} = \beta ProducingWell_{ct-i} + X_{ct}\theta + \alpha_c + \gamma_t + \varepsilon_{ct} \tag{63}$$

Including lag and lead variables allows the estimation of long-run effects which may result from support industries entering the area. Given that unconventional natural gas activity is highly correlated over time, including lags and leads of producing wells induces multicollinearity and makes inference more challenging (Paredes 2015). As shown in Table B21. , variants of the distributed lag model show that the aggregated effect of the contemporaneous and lag/lead variables yield a similar effect size as that for the contemporaneous specification. In addition, the lag and lead variables are statistically significant.

We also fit models for different natural gas activity variables, including producing wells, production, and spud wells. As shown in Table B23, the effects associated with alternatively incorporating a single type of activity variable are all positive and significant. When including only the producing wells variable, we find the mean marginal effect size is 5 jobs per producing wells; Paredes (2015) similarly finds that 6.8 to 16.8 jobs are supported per producing well. Given that producing wells and production are highly correlated, including both terms is redundant. We find that including both producing and spud wells reveals positive and significant effects associated with both variables. This can be interpreted as a larger employment effect associated with well development than that associated with an already producing well.

Incorporating an interaction term between gas activity and a dummy variable indicating whether a well was producing before or after 2012, we find decreasing marginal employment effects from natural gas activity over time, with 16 jobs per producing well prior to 2012 and 5 jobs thereafter. The intuition is that learning occurs and the industry becomes more efficient over time, which translates into decreasing marginal employment effects from natural gas activity over time. Table B24 shows the interaction between natural gas activity and time.

Finally, we specify models with alternative dependent variables, including the employment-to-population ratio and employment rate. The purpose of using rate-based dependent variables is because they may have a more meaningful interpretation than absolute employment; for example the creation of 100 jobs in a given county may or may not be significant, depending on the size of the total labor market. As previously noted, the underlying data used to develop the employment rate may not accurately reflect annual changes disaggregated at the county-level, thus limiting their utility for regression analysis. We find that the models incorporating these alternative dependent variables do not provide a good overall model fit; however, there is a positive and significant effect on the producing well variable for the employment-to-population ratio model.

Table B21. Employment effects for distributed lag compared to contemporaneous models.

	(1) No lag				(10) 1-year lag				(11) 2-year lag			
Variables	Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)	
Producing wells	5.36	1.86	0.004	**	19.64	5.45	0.000	***	11.13	3.58	0.002	**
1-year lag producing wells	-				-16.68	4.60	0.000	***	-			
2-year lag producing wells	-				-				-9.18	3.21	0.004	**
Population	0.47	0.09	0.000	***	0.49	0.15	0.001	***	0.52	0.15	0.001	***
Population density	88.91	31.63	0.005	**	72.64	70.43	0.302		61.02	76.14	0.423	
Farm earnings share	6.88E+03	3.30E+03	0.037	*	7.91E+03	3.69E+03	0.032	*	8.99E+03	4.07E+03	0.027	*
Construction earnings share	8.08E+03	3.79E+03	0.033	*	5.86E+03	3.91E+03	0.134		4.87E+03	3.87E+03	0.208	
Manufacturing earnings share	7.26E+03	3.24E+03	0.025	*	6.76E+03	3.48E+03	0.052	.	5.66E+03	3.78E+03	0.135	
Retail earnings share	-3.77E+04	1.25E+04	0.003	**	-3.30E+04	1.44E+04	0.022	*	-3.37E+04	1.69E+04	0.046	*
Number of counties	241				241				241			
Years	2004-2016				2004-2016				2004-2016			
Sample	Nonmetropolitan counties				Nonmetropolitan counties				Nonmetropolitan counties			
R-squared	0.55				0.50				0.46			
Adjusted R-squared	0.51				0.45				0.39			
Significance codes: ***p<0.001, **p<0.01, *p<0.05, .p<0.1												

Table B22. Employment effects for distributed lead compared to contemporaneous models.

	(1) No lead				(12) 1-year lead				(13) 2-year lead			
Variables	Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)	
Producing wells	5.36	1.86	0.004	**	-12.27	4.04	0.002	**	-3.31	2.69	0.220	
1-year lead producing wells	-				17.39	5.12	0.001	***	-			
2-year lead producing wells	-				-				9.46	3.53	0.007	**
Population	0.47	0.09	0.000	***	0.46	0.16	0.004	**	0.40	0.18	0.027	*
Population density	88.91	31.63	0.005	**	79.58	70.35	0.258		94.17	78.55	0.231	
Farm earnings share	6.88E+03	3.30E+03	0.037	*	7.25E+03	4.17E+03	0.082	.	6.22E+03	5.46E+03	0.255	
Construction earnings share	8.08E+03	3.79E+03	0.033	*	5.11E+03	4.51E+03	0.257		5.83E+03	5.58E+03	0.297	
Manufacturing earnings share	7.26E+03	3.24E+03	0.025	*	8.61E+03	4.04E+03	0.033	*	9.30E+03	4.97E+03	0.062	.
Retail earnings share	-3.77E+04	1.25E+04	0.003	**	-3.18E+04	1.38E+04	0.021	*	-3.30E+04	1.61E+04	0.041	*
Number of counties	241				241				241			
Years	2004-2016				2004-2016				2004-2016			
Sample	Nonmetropolitan counties				Nonmetropolitan counties				Nonmetropolitan counties			
R-squared	0.55				0.54				0.51			
Adjusted R-squared	0.51				0.49				0.45			
Significance codes: ***p<0.001, **p<0.01, *p<0.05, .p<0.1												

Table B23. Employment effects for different natural gas activity variables.

	(1) Producing wells				(14) Spud wells				(15) Production			
Variables	Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)	
Producing wells	5.36	1.86	0.004	**	-				-			
Spud wells	-				2.08E+01	6.32	0.001	**	-			
Production	-				-				5.01E-03	0.00	0.037	*
Population	0.47	0.09	0.000	***	4.63E-01	8.65E-02	0.000	***	4.66E-01	8.68E-02	0.000	***
Population density	88.91	31.63	0.005	**	8.92E+01	31.51	0.005	**	8.87E+01	31.63	0.005	**
Farm earnings share	6.88E+03	3.30E+03	0.037	*	7.35E+03	3430.70	0.032	*	5.99E+03	3316.30	0.071	.
Construction earnings share	8.08E+03	3.79E+03	0.033	*	7.61E+03	4.02E+03	0.058	.	9.38E+03	3.81E+03	0.014	*
Manufacturing earnings share	7.26E+03	3.24E+03	0.025	*	6.44E+03	3175.00	0.043	*	6.79E+03	3202.30	0.034	*
Retail earnings share	-3.77E+04	1.25E+04	0.003	**	-4.02E+04	12326.00	0.001	**	-3.97E+04	12389.00	0.001	**
Number of counties	241				241				241			
Years	2004-2016				2004-2016				2004-2016			
Sample	Nonmetropolitan counties				Nonmetropolitan counties				Nonmetropolitan counties			
R-squared	0.55				0.55				0.55			
Adjusted R-squared	0.51				0.51				0.50			
Significance codes: ***p<0.001, **p<0.01, *p<0.05, .p<0.1												

Table B23 (continued). Employment effects for different natural gas activity variables.

	(1) Producing wells				(16) Producing and spud wells				(17) Producing wells and production				(18) Producing wells, spud wells, and production			
Variables	Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)	
Producing wells	5.36	1.86	0.004	**	4.50	1.59	0.005	**	-8.30E-03	0.00	0.063	*	1.07E+01	3.74	0.004	**
Spud wells	-				1.64E+01	4.20	0.000	***	-				1.66E+01	3.35	0.000	***
Production	-				-				1.14E+01	4.70	0.015	*	-8.49E-03	0.00	0.017	*
Population	0.47	0.09	0.000	***	4.68E-01	8.67E-02	0.000	***	4.70E-01	8.66E-02	0.000	*	4.69E-01	8.65E-02	0.000	***
Population density	88.91	31.63	0.005	**	8.94E+01	31.58	0.005	**	8.90E+01	31.61	0.005		8.95E+01	31.56	0.005	**
Farm earnings share	6.88E+03	3.30E+03	0.037	*	8.70E+03	3374.10	0.010	**	7.17E+03	3335.50	0.032		9.01E+03	3389.60	0.008	**
Construction earnings share	8.08E+03	3.79E+03	0.033	*	5.75E+03	3.98E+03	0.149		7.58E+03	3.87E+03	0.050		5.21E+03	4.04E+03	0.198	
Manufacturing earnings share	7.26E+03	3.24E+03	0.025	*	7.52E+03	3280.50	0.022	*	7.28E+03	3226.20	0.024	*	7.54E+03	3264.50	0.021	*
Retail earnings share	-3.77E+04	1.25E+04	0.003	**	-3.53E+04	12598.00	0.005	**	-3.80E+04	12516.00	0.002		-3.56E+04	12612.00	0.005	**
Number of counties	241				241				241				241			
Years	2004-2016				2004-2016				2004-2016				2004-2016			
Sample	Nonmetropolitan counties				Nonmetropolitan counties				Nonmetropolitan counties				Nonmetropolitan counties			
R-squared	0.55				0.56				0.56				0.56			
Adjusted R-squared	0.51				0.52				0.52				0.52			
Significance codes: ***p<0.001, **p<0.01, *p<0.05, .p<0.1																

Table B24. Employment effects including time and natural gas activity interactions.

	(1) No interaction				(19) Interaction between producing wells and before 2012 dummy			
Variables	Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)	
Producing wells	5.36	1.86	0.004	**	5.62	1.91	0.003	**
Population	0.47	0.09	0.000	***	4.69E-01	0.09	0.000	***
Population density	88.91	31.63	0.005	**	8.90E+01	3.16E+01	0.005	**
Farm earnings share	6.88E+03	3.30E+03	0.037	*	7.83E+03	3321.50	0.018	*
Construction earnings share	8.08E+03	3.79E+03	0.033	*	7.49E+03	3709.40	0.043	*
Manufacturing earnings share	7.26E+03	3.24E+03	0.025	*	7.21E+03	3228.60	0.026	*
Retail earnings share	-3.77E+04	1.25E+04	0.003	**	-3.64E+04	12459.00	0.004	**
Producing wells x Before 2012	-				1.08E+01	1.82	0.000	***
Number of counties	241				241			
Years	2004-2016				2004-2016			
Sample	Nonmetropolitan counties				Nonmetropolitan counties			
R-squared	0.55				0.56			
Adjusted R-squared	0.51				0.51			
Significance codes: ***p<0.001, **p<0.01, *p<0.05, .p<0.1								

Table B25. Effects for models with alternative dependent variables.

	(1) Employment				(20) Employment rate				(21) Employment-to-population ratio			
Variables	Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)		Estimate	Std. Error	P(>|t|)	
Producing wells	5.36	1.86	0.004	**	-1.55E-03	0.00	0.000	***	5.79E-05	0.00	0.000	***
Population	0.47	0.09	0.000	***	1.20E-05	8.31E-06	0.148		-2.40E-07	1.85E-07	0.193	
Population density	88.91	31.63	0.005	**	-2.89E-03	0.00	0.124		2.08E-04	0.00	0.000	***
Farm earnings share	6.88E+03	3.30E+03	0.037	*	-1.13E+01	2.43	0.000	***	-1.50E-01	0.05	0.004	**
Construction earnings share	8.08E+03	3.79E+03	0.033	*	7.27E+00	2.31	0.002	**	1.28E-01	4.73E-02	0.007	**
Manufacturing earnings share	7.26E+03	3.24E+03	0.025	*	2.18E+00	2.21	0.324		8.81E-02	0.05	0.058	.
Retail earnings share	-3.77E+04	1.25E+04	0.003	**	-9.50E+00	6.37	0.136		-6.91E-01	0.21	0.001	***
Number of counties	241				241				241			
Years	2004-2016				2004-2016				2004-2016			
Sample	Nonmetropolitan counties				Nonmetropolitan counties				Nonmetropolitan counties			
R-squared	0.55				0.04				0.17			
Adjusted R-squared	0.51				-0.05				0.09			
Significance codes: ***p<0.001, **p<0.01, *p<0.05, .p<0.1												

B5.2.4 Model checking

Actual versus predicted employment

We compare the actual versus predicted employment for some of the models specified in the previous sections. All of the model predictions closely fit the actual employment, as shown in Figure B28.

Bootstrapping

To capture the uncertainty around the employment effects of natural gas activity, we estimate confidence intervals through bootstrapping. We use block sampling, where a single draw consists of all observations for a county, thereby maintaining the dependence structure between years. We also resample cases, rather than residuals, which does not assume that the shape of regression function or the distribution of the error of the original model are correct; resampling cases treats each bootstrap sample as an observation and refits the fixed effects model.

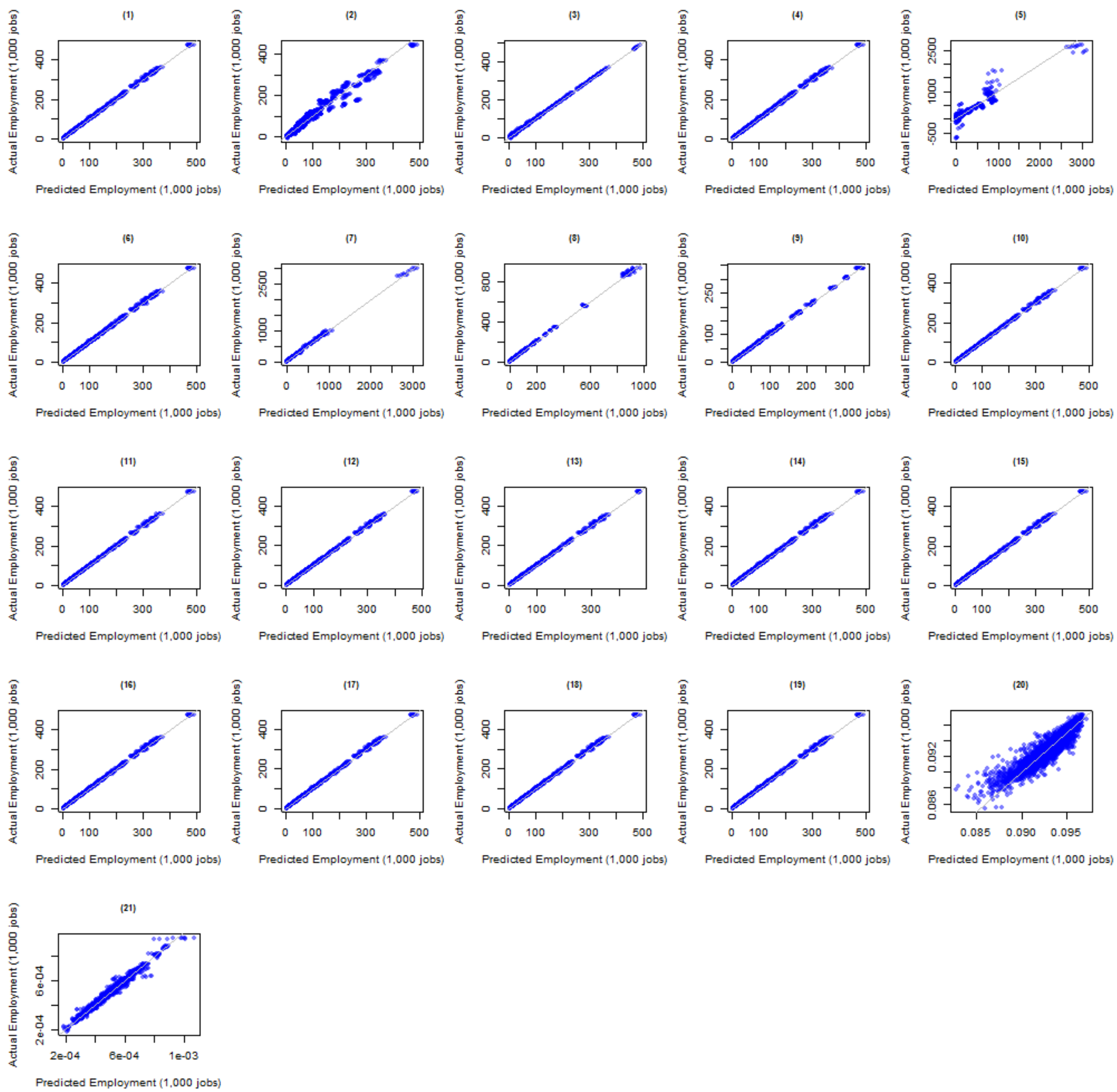

Figure B28. Actual versus predicted employment for different model specifications.

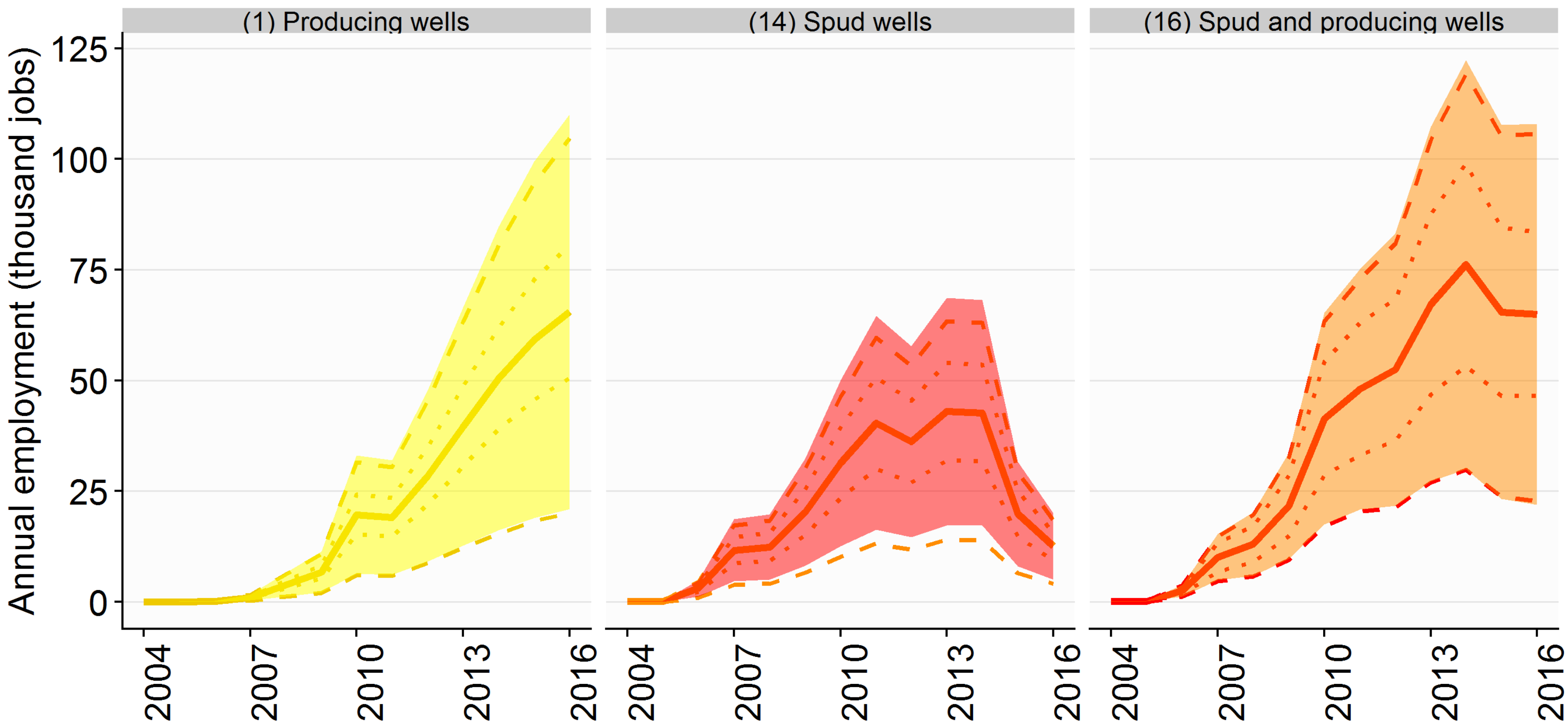

Figure B29. Annual employment based on model specifications 1, 14, and 16. Solid line represents mean. Shaded region represents 95% confidence interval based on robust standard errors clustered by county. Dashed lines represent bootstrapped 95% confidence interval. Dotted lines represent within model 95% confidence interval.

B6 References

1. U.S. Energy Information Adminstration. *U.S. Crude Oil and Natural Gas Proved Reserves, Year-end 2016*. (2018).
2. U.S. Geological Survey. National Assessment of Oil and Gas Project - Appalachian Basin Province (067) Boundary. (2002).
3. U.S. Geological Survey. National Assessment of Oil and Gas Project Devonian Marcellus Shale of the Appalachian Basin Province (067) Assessment Units. (2011).
4. U.S. Geological Survey. National Assessment of Oil and Gas Project - Utica Shale Unconventional Assessment. (2015).
5. Pennsylvania Department of Environmental Protection. PA DEP Oil & Gas Reporting Website. Available at: https://www.paoilandgasreporting.state.pa.us/publicreports/Modules/Welcome/Welcome.aspx.
6. Ohio Department of Natural Resources. Oil & Gas Resources. Available at: http://oilandgas.ohiodnr.gov/production.
7. West Virginia Department of Environmental Protection. Oil and Gas Production Data. Available at: https://dep.wv.gov/oil-and-gas/databaseinfo/Pages/default.aspx.
8. U.S. Energy Information Adminstration. Natural Gas Gross Withdrawals from Gas Wells. Available at: http://www.eia.gov/dnav/ng/ng_prod_sum_a_epg0_fgg_mmcf_a.htm%0A.
9. U.S. Energy Information Adminstration. International and Interstate Movements of Natural Gas by State. Available at: http://www.eia.gov/dnav/ng/ng_move_ist_a2dcu_swv_a.htm%0A.
10. U.S. Energy Information Adminstration. Natural Gas Processed. Available at: http://www.eia.gov/dnav/ng/ng_prod_pp_a_epg0_ygp_mmcf_a.htm%0A.
11. U.S. Energy Information Adminstration. Natural Gas Lease Fuel Consumption. Available at: http://www.eia.gov/dnav/ng/ng_cons_sum_a_epg0_vcl_mmcf_a.htm%0A.
12. U.S. Energy Information Adminstration. Natural Gas Plant Fuel Consumption. Available at: http://www.eia.gov/dnav/ng/ng_cons_sum_a_epg0_vcf_mmcf_a.htm%0A.
13. U.S. Energy Information Adminstration. Natural Gas Pipeline & Distribution Use. Available at: http://www.eia.gov/dnav/ng/ng_cons_sum_a_epg0_vgp_mmcf_a.htm%0A.
14. U.S. Energy Information Adminstration. Natural Gas Delivered to Electric Power Consumers. Available at: http://www.eia.gov/dnav/ng/ng_cons_sum_a_epg0_veu_mmcf_a.htm%0A.
15. U.S. Energy Information Adminstration. Natural Gas Delivered to Industrial Consumers. Available at: http://www.eia.gov/dnav/ng/ng_cons_sum_a_epg0_vin_mmcf_a.htm%0A.
16. U.S. Energy Information Adminstration. Natural Gas Delivered to Commercial Consumers. Available at: http://www.eia.gov/dnav/ng/ng_cons_sum_a_epg0_vcs_mmcf_a.htm%0A.
17. U.S. Energy Information Adminstration. Natural Gas Delivered to Vehicle Fuel Consumers. Available at: http://www.eia.gov/dnav/ng/ng_cons_sum_a_epg0_vdv_mmcf_a.htm%0A.
18. U.S. Energy Information Adminstration. Natural Gas Delivered to Residential Consumers. Available at: http://www.eia.gov/dnav/ng/ng_cons_sum_a_epg0_vrs_mmcf_a.htm%0A.
19. Allen, D. T. Atmospheric Emissions and Air Quality Impacts from Natural Gas Production and Use. *Annu. Rev. Chem. Biomol. Eng.* **5,** 55–75 (2014).
20. U.S. Environmental Protection Agency. *Part II AGENCY :Oil and Natural Gas Sector: New Source Performance Standards and National Emission Standards for Hazardous Air Pollutants Reviews;*

Final Rule. **77,** 1–27 (2012).

21. Roy, A. A., Adams, P. J. & Robinson, A. L. Air pollutant emissions from the development, production, and processing of Marcellus Shale natural gas. *J. Air Waste Manage. Assoc.* **64,** 19–37 (2014).

22. Roohani, Y. H., Roy, A. A., Heo, J., Robinson, A. L. & Adams, P. J. Impact of natural gas development in the Marcellus and Utica shales on regional ozone and fine particulate matter levels. *Atmos. Environ.* **155,** 11–20 (2017).

23. Jiang, M., Hendrickson, C. T. & Vanbriesen, J. M. Life cycle water consumption and wastewater generation impacts of a Marcellus shale gas well. *Environ. Sci. Technol.* **48,** 1911–1920 (2014).

24. Rahm, B. G. *et al.* Wastewater management and Marcellus Shale gas development: Trends, drivers, and planning implications. *J. Environ. Manage.* **120,** 105–113 (2013).

25. Bar-Ilan, A., Parikh, R., Grant, J., Shah, T. & Pollack, A. K. *Recommendations for Improvements to the CENRAP States' Oil and Gas Emissions Inventories*. (2008).

26. Roy, A. A. *et al.* Air pollutant emissions from the development , production , and processing of Marcellus Shale natural gas Air pollutant emissions from the development , production , and processing of Marcellus Shale natural gas. **2247,** (2014).

27. Allen, D. T. *et al.* Methane emissions from process equipment at natural gas production sites in the United States: liquid unloadings. *Env. Sci Technol* **49,** 641–648 (2014).

28. Allen, D. T. *et al.* Measurements of methane emissions at natural gas production sites in the United States. *Proc. Natl. Acad. Sci.* **110,** 17768–17773 (2013).

29. Omara, M. *et al.* Methane Emissions from Conventional and Unconventional Natural Gas Production Sites in the Marcellus Shale Basin. *Environ. Sci. Technol.* **50,** 2099–2107 (2016).

30. U.S. Environmental Protection Agency. Median Life , Annual Activity , and Load Factor Values for Nonroad Engine Emissions Modeling. *EPA-420-R-10-016. NR-005d. U.S. Environ. Prot. Agency* 1–47 (2010).

31. U.S. Environmental Protection Agency. Nonroad Engine Population Estimates Nonroad Engine Population Estimates. (2010).

32. U.S. Environmental Protection Agency. Calculation of Age Distributions in the Nonroad Model: Growth and Scrappage. 16 (2005).

33. Baker, R. & Pring, M. Drilling Rig Emission Inventory for the State of Texas. **2009,** (2014).

34. Armendariz, A., Ph, D. & Alvarez, R. Emissions from Natural Gas Production in the Barnett Shale Area and Opportunities for Cost-Effective Improvements report by : 45 (2009). doi:10.1656/058.012.s413

35. U.S. Environmental Protection Agency. Control Techniques Guidelines for the Oil and Natural Gas Industry. 343 (2016).

36. U.S. Environmental Protection Agency. National Emissions Inventory. (2018). Available at: https://www.epa.gov/air-emissions-inventories/2014-national-emissions-inventory-nei-data.

37. U.S. Environmental Protection Agency. Continuous Environmental Monitoring System. Available at: https://ampd.epa.gov/ampd/.

38. U.S. Environmental Protection Agency. Emissions & Generation Resource Integrated Database. (2018). Available at: https://www.epa.gov/energy/emissions-generation-resource-integrated-database-egrid.

39. Behrer, A. P. & Mauter, M. S. Allocating Damage Compensation in a Federalist System: Lessons

from Spatially Resolved Air Emissions in the Marcellus. *Environ. Sci. Technol.* **51,** 3600–3608 (2017).

40. Lutz, B. D., Lewis, A. N. & Doyle, M. W. Generation , transport , and disposal of wastewater associated with Marcellus Shale gas development. **49,** 647–656 (2013).

41. Jiang, M. *et al.* Life cycle greenhouse gas emissions of Marcellus shale gas. *Environ. Res. Lett.* **6,** (2011).

42. Tong, F., Jaramillo, P. & Azevedo, I. M. L. Comparison of Life Cycle Greenhouse Gases from Natural Gas Pathways for Medium and Heavy-Duty Vehicles. *Environ. Sci. Technol.* **49,** 7123–7133 (2015).

43. Weber, C. L. & Clavin, C. Life Cycle Carbon Footprint of Shale Gas: Review of Evidence and Implications. *Environ. Sci. Technol.* **46,** 5688–5695 (2012).

44. U.S. Environmental Protection Agency. *2014 National Emissions Inventory, version 2 - Technical Support Document*. (2018).

45. Muller, N. Z. & Mendelsohn, R. *The Air Pollution Emission Experiments and Policy Analysis Model (APEEP) Technical Appendix*. (2006).

46. Muller, N. Z. Boosting GDP growth by accounting for the environment. *Science (80-.).* **345,** 873–874 (2014).

47. Heo, J., Adams, P. J. & Gao, H. O. Public health costs accounting of inorganic PM2.5pollution in metropolitan areas of the United States using a risk-based source-receptor model. *Environ. Int.* **106,** 119–126 (2017).

48. Heo, J. & Adams, P. J. EASIUR User 's Guide Version 0 . 2. (2015).

49. Krewski, D. *et al.* Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality. *Res Rep Heal. Eff Inst* 5–36 (2009).

50. Lepeule, J., Laden, F., Dockery, D. & Schwartz, J. Chronic exposure to fine particles and mortality: An extended follow-up of the Harvard six cities study from 1974 to 2009. *Environ. Health Perspect.* **120,** 965–970 (2012).

51. U.S. EPA. *Guidelines for Preparing Economic Analyses*. **2010,** (2014).

52. Litovitz, A., Curtright, A., Abramzon, S., Burger, N. & Samaras, C. Estimation of regional air-quality damages from Marcellus Shale natural gas extraction in Pennsylvania. *Environ. Res. Lett.* **8,** (2013).

53. Cohon, J. L. *et al. Hidden costs of energy: Unpriced consequences of energy production and use*. *AIP Conference Proceedings* **1401,** (2011).

54. Jaramillo, P., Griffin, M. W. & Matthews, S. H. Comparative Life-Cycle Air Emissions of Coal , Domestic Natural Gas , LNG , and SNG for Electricity Generation. *Environ. Sci. Technol.* **41,** 6290–6296 (2007).

55. Venkatesh, A., Jaramillo, P., Griffin, W. M. & Matthews, H. S. Uncertainty in life cycle greenhouse gas emissions from United States natural gas end-uses and its effects on policy. *Environ. Sci. Technol.* **45,** 8182–8189 (2011).

56. Burnham, A. *et al.* Life-cycle greenhouse gas emissions of shale gas, natural gas, coal, and petroleum. *Environ. Sci. Technol.* **46,** 619–27 (2012).

57. Burnham, A., Han, J., Elgowainy, A. & Wang, M. Updated Fugitive Greenhouse Gas Emissions for Natural Gas Pathways in the GREET1_2014 Model. (2014).

58. Tong, F., Jaramillio, P. & Azevedo, M. L. Comparison of Life Cycle Greenhouse Gases from

Natural Gas Pathways for Light-Duty Vehicles. *Energy Fuels* 6008–6018 (2015). doi:10.1021/acs.energyfuels.5b01063

59. Abrahams, L. S., Samaras, C. & Matthews, H. S. Life Cycle Greenhouse Gas Emissions From U.S. Lique fi ed Natural Gas Exports: Implications for End Uses. (2015). doi:10.1021/es505617p

60. U.S. Environmental Protection Agency. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2016*. (2018).

61. Heath, G. A., Donoughue, P. O., Arent, D. J. & Bazilian, M. Harmonization of initial estimates of shale gas life cycle greenhouse gas emissions for electric power generation. (2014). doi:10.1073/pnas.1309334111

62. Stephenson, T., Valle, J. E. & Riera-Palou, X. Modeling the relative GHG emissions of conventional and shale gas production. *Environ. Sci. Technol.* **45,** 10757–10764 (2011).

63. Jaramillo, P., Griffin, W. M. & Matthews, H. S. Comparative life cycle air emissions of coal, domestic natural gas, LNG, and SNG for electricity generation. *Environ. Sci. Technol.* **41,** 6290 (2007).

64. Schivley, G., Azevedo, I. M. L. & Samaras, C. Assessing the evolution of power sector carbon intensity in the United States. *Environ. Res. Lett.* (2018). doi:https://doi.org/10.1088/1478-3975/aa9768

65. U.S. Environmental Protection Agency. *Oil and Natural Gas Sector: Emission Standards for New, Reconstructed, and Modified Sources*. **81,** 35824–35942 (2016).

66. Zimmerle, D. J. *et al.* Methane Emissions from the Natural Gas Transmission and Storage System in the United States. *Environ. Sci. Technol.* **49,** 9374–9383 (2015).

67. Marchese, A. J. *et al.* Methane emissions from United States natural gas gathering and processing-SI. *Environ. Sci. Technol.* (2015).

68. Lamb, B. K. *et al.* Direct and Indirect Measurements and Modeling of Methane Emissions in Indianapolis, Indiana. *Environ. Sci. Technol.* **50,** 8910–8917 (2016).

69. Lamb, B. K. *et al.* Direct measurements show decreasing methane emissions from natural gas local distribution systems in the United States. *Environ. Sci. Technol.* **49,** 5161–9 (2015).

70. Mitchell, A. L. *et al.* Measurements of Methane Emissions from Natural Gas Gathering Facilities and Processing Plants: Measurement Results_supplemental information. *Environ. Sci. Technol.* **49,** 3219–3227 (2015).

71. Myhre, G. *et al.* Anthropogenic and Natural Radiative Forcing. *Clim. Chang. 2013 Phys. Sci. Basis. Contrib. Work. Gr. I to Fifth Assess. Rep. Intergov. Panel Clim. Chang.* 659–740 (2013). doi:10.1017/ CBO9781107415324.018

72. Alvarez, R. A. *et al.* Assessment of methane emissions from the U . S . oil and gas supply chain. **7204,** 1–9 (2018).

73. U.S. Environmental Protection Agency. Annexes to the Inventory of U.S. GHG Emissions and Sinks. 1–458 (2018).

74. U.S. Environmental Protection Agency. *AP 42, Fifth Edition, Volume I Chapter 1: External Combustion Sources*. (1998).

75. U.S. Energy Information Adminstration. Carbon Dioxide Emissions Coefficients. (2016). Available at: https://www.eia.gov/environment/emissions/co2_vol_mass.php.

76. U.S. Energy Information Adminstration. Heat Content of Natural Gas Delivered to Consumers. Available at: http://www.eia.gov/dnav/ng/ng_cons_heat_a_epg0_vgth_btucf_a.htm%0A.

77. Marchese, A. J. *et al.* Methane Emissions from United States Natural Gas Gathering and Processing. *Environ. Sci. Technol.* **49,** 10718–10727 (2015).

78. Pipeline and Hazardous Materials Safety Administration. Gas Distribution, Gas Gathering, Gas Transmission, Hazardous Liquids, Liquefied Natural Gas (LNG), and Underground Natural Gas Storage (UNGS) Annual Report Data. Available at: https://www.phmsa.dot.gov/data-and-statistics/pipeline/gas-distribution-gas-gathering-gas-transmission-hazardous-liquids.

79. U.S. Energy Information Adminstration. Natural Gas Delivered to Consumers. Available at: http://www.eia.gov/dnav/ng/ng_cons_sum_a_epg0_vgt_mmcf_a.htm%0A.

80. Frey, H. C. & Zhao, Y. Quantification of variability and uncertainty for air toxic emission inventories with censored emission factor data. *Environ. Sci. Technol.* **38,** 6094–6100 (2004).

81. Brandt, A. R., Heath, G. A. & Cooley, D. Methane Leaks from Natural Gas Systems Follow Extreme Distributions. *Environ. Sci. Technol.* **50,** 12512–12520 (2016).

82. Romano, D., Bernetti, A. & De Lauretis, R. Different methodologies to quantify uncertainties of air emissions. *Environ. Int.* **30,** 1099–1107 (2004).

83. Burnham, A. *et al.* Life-cycle greenhouse gas emissions of shale gas, natural gas, coal, and petroleum. *Environ. Sci. Technol.* **46,** 619–627 (2012).

84. Zimmerle, D. J. *et al.* Gathering pipeline methane emissions in Fayetteville shale pipelines and scoping guidelines for future pipeline measurement campaigns. *Elem Sci Anth.* **5,** 70 (2017).

85. O'Neill, B. C. The jury is still out on Global Warming Potentials. *Clim. Change* **44,** 427–443 (2000).

86. Fuglestvedt, J. S. *et al.* Metrics of climate change: Assessing radiative forcing and emission indices. *Clim. Change* **58,** 267–331 (2003).

87. Reilly, J. M. & Richards, K. R. Climate change damage and the trace gas index issue. *Environ. Resour. Econ.* **3,** 41–61 (1993).

88. Schmalensee, R. Comparing Greenhouse Gases for Policy Purposes. *Energy J.* **14,** 245–255 (2018).

89. Hammitt, J. K., Jain, A. K., Adams, J. L. & Wuebb. A welfare-based index for assessing environmental effects of GHG emissions. *Nature* **381,** 301–303 (1996).

90. Boucher, O. Comparison of physically- and economically-based CO2- equivalences for methane. *Earth Syst. Dyn.* **3,** 49–61 (2012).

91. Berntsen, T. & Fuglestvedt, J. Global temperature responses to current emissions from the transport sectors. *Proc. Natl. Acad. Sci.* **105,** 19154–19159 (2008).

92. Shindell, D. & Faluvegi, G. The net climate impact of coal-fired power plant emissions. *Atmos. Chem. Phys.* **10,** 3247–3260 (2010).

93. Shine, K. P., Fuglestvedt, J. S., Hailemariam, K. & Stuber, N. Alternatives to the Global Warming Potential for comparing climate impacts of emissions of greenhouse gases. *Clim. Change* **68,** 281–302 (2005).

94. Boucher, O. & Reddy, M. S. Climate trade-off between black carbon and carbon dioxide emissions. *Energy Policy* **36,** 193–200 (2008).

95. Shine, K. P., Berntsen, T. K., Fuglestvedt, J. S., Skeie, R. B. & Stuber, N. Comparing the climate effect of emissions of short- and long-lived climate agents. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **365,** 1903–1914 (2007).

96. Tol, R. S. J., Berntsen, T. K., Oneill, B. C., Fuglestvedt, J. S. & Shine, K. P. A unifying framework for metrics for aggregating the climate effect of different emissions. *Environ. Res. Lett.* **7,** (2012).

97. Shindell, D. *et al.* Climate, health, agricultural and economic impacts of tighter vehicle-emission

standards. *Nat. Clim. Chang.* **1,** 59–66 (2011).

98. Peters, G. P., Aamaas, B., Berntsen, T. & Fuglestvedt, J. S. The integrated global temperature change potential (iGTP) and relationships between emission metrics. *Environ. Res. Lett.* **6,** 044021 (2011).

99. Fuglestvedt, J. S. *et al.* Transport impacts on atmosphere and climate: Metrics. *Atmos. Environ.* **44,** 4648–4677 (2010).

100. Joos, F. *et al.* Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: A multi-model analysis. *Atmos. Chem. Phys.* **13,** 2793–2825 (2013).

101. Collins, W. J. *et al.* Global and regional temperature-change potentials for near-term climate forcers. *Atmos. Chem. Phys.* **13,** 2471–2485 (2013).

102. National Oceanic and Atmospheric Administration. Global atmospheric carbon dioxide concentration. Available at: ftp://aftp.cmdl.noaa.gov/products/trends/co2/co2_annmean_mlo.txt.

103. Considine, T. J., Watson, R. & Blumsack, S. *The economic impacts of the Pennsylvania Marcellus Shale natural gas play : An update The Economic Impacts of the Pennsylvania Marcellus Shale Natural Gas Play : An Update*. (2010).

104. Weber, J. G. The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming. *Energy Econ.* **34,** 1580–1588 (2012).

105. Marchand, J. & Weber, J. Local labor markets and natural resources: A synthesis of the literature. *J. Econ. Surv.* **00,** 1–22 (2017).

Appendix C: Supplemental Information for Chapter 5

Table of Contents

List of Figures

List of Tables

C1 Supplementary methods and results

C1.1 Spatial and temporal equity metrics

Table C1. Summary of inequality indicators (adapted from Levy et al 2006).[17]

	Gini index	Variance of logarithms[13]	Squared coefficient of variation	Atkinson index	Mean log deviation	Theil's entropy index [14]
Description	Measure of dispersion scaled by twice the value of the mean. Measures relative difference between actual and uniform distributions.	Measure of dispersion of the log values.	Entropy measure. Includes sensitivity measure, ϑ, set to 2.	Entropy measure. Explicitly incorporates normative judgments about social welfare.	Entropy measure.	Entropy measure. Average if the reciprocals of weightd probabilities.
Formula	$\frac{\frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n}\lvert x_k - x_l\rvert}{\mu}$	$\frac{1}{n}\sum_{k=1}^{n}(z_k - \bar{z})^2$ where $z_k = \ln(x_k)$	$\frac{\sigma^2}{\mu^2}$	$1 - \left[\frac{1}{n}\sum_{k=1}^{n}\left[\frac{x_k}{\bar{x}}\right]^{1-\epsilon}\right]^{\frac{1}{1-\varepsilon}}$ where ε = inequality aversion (range from 0 to infinity)	$\frac{1}{n}\sum_{k=1}^{n}\ln\left(\frac{\mu}{x_k}\right)$	$\frac{1}{n}\sum_{k=1}^{n}\left(\frac{x_k}{\mu}\right)\ln\left(\frac{x_k}{\mu}\right)$
Approach for comparisons	Relative to all those better off	Relative to average	Relative to average	Relative to average	Relative to average	Relative to average
Method for aggregation	Additive	Weighted additive	Weighted additive	Weighted additive	Weighted additive	Weighted additive
Principle of transfers?[a,15]	Yes	No (fails for transfers at high levels)	No (fails for transfers at high levels)	Yes	Yes	Yes
Subgroup decomposable?[b,16]	No (unless subgroups strictly ordered)	No	Yes (within-group and between-group not independent)	Yes (although not strictly additive)	Yes	Yes

a Axiom that an indicator must not decrease when income (or other parameter) is transferred from a poorer to a richer person (or from a person worse off to a person better off). The indicator must decrease when income (or other parameter) is transferred from a richer to a poorer person (or from a person better off to a person worse off).
b Axiom that an indicator must be subgroup decomposable (or additive separable). Total inequality can be divided into constituent parts of the distribution.

Table C2. Spatial and temporal equity metrics for air quality impacts. Based on premature mortality estimates.

Equity Unit[a]	Gini Coefficient	Atkinson Index (ε =0.5)	Maximum Absolute Difference (premature mortality)	Theil Index	Coefficient of Variation	Squared Coefficient of Variation
Spatial equity – all grid cells (with mortality >0.1)[b]						
Grid cell (2010)	0.81	0.56	10	1.73	4.15	17.3
Grid cell (2016)	0.74	0.46	13	1.33	3.21	10.3
Grid cell (cumulative)	0.76	0.50	98	1.48	3.55	12.6
Spatial equity – all counties (with mortality >0.1)[c]						
County (2010)	0.78	0.52	20	1.51	3.49	12.2
County (2016)	0.66	0.37	22	0.96	2.33	5.15
County (cumulative)	0.69	0.40	201	1.12	2.80	7.82
Spatial equity – counties within Pennsylvania, Ohio, and West Virginia[d]						
County (2010)	0.60	0.30	20	0.82	2.27	5.15
County (2016)	0.62	0.32	22	0.83	2.13	4.52
County (cumulative)	0.77	0.50	201	1.42	3.58	12.9
Temporal equity[c]						
Year	0.47	0.29	439	0.45	0.83	0.70

a We estimate spatial equity between counties or 36 x 36 km grid cells for 2010, 2016, or cumulatively from 2004 to 2016.
b Based on mortality estimates from APSCA source-receptor RCM and ACS C-R relationship model specification.
c Based on mortality estimates from AP3 source-receptor RCM and ACS C-R relationship model specification.
d Based on mean mortality estimates across six model specifications using different RCMs (i.e., AP3, APSCA, InMAP) and C-R relationships (i.e., ACS, H6C).

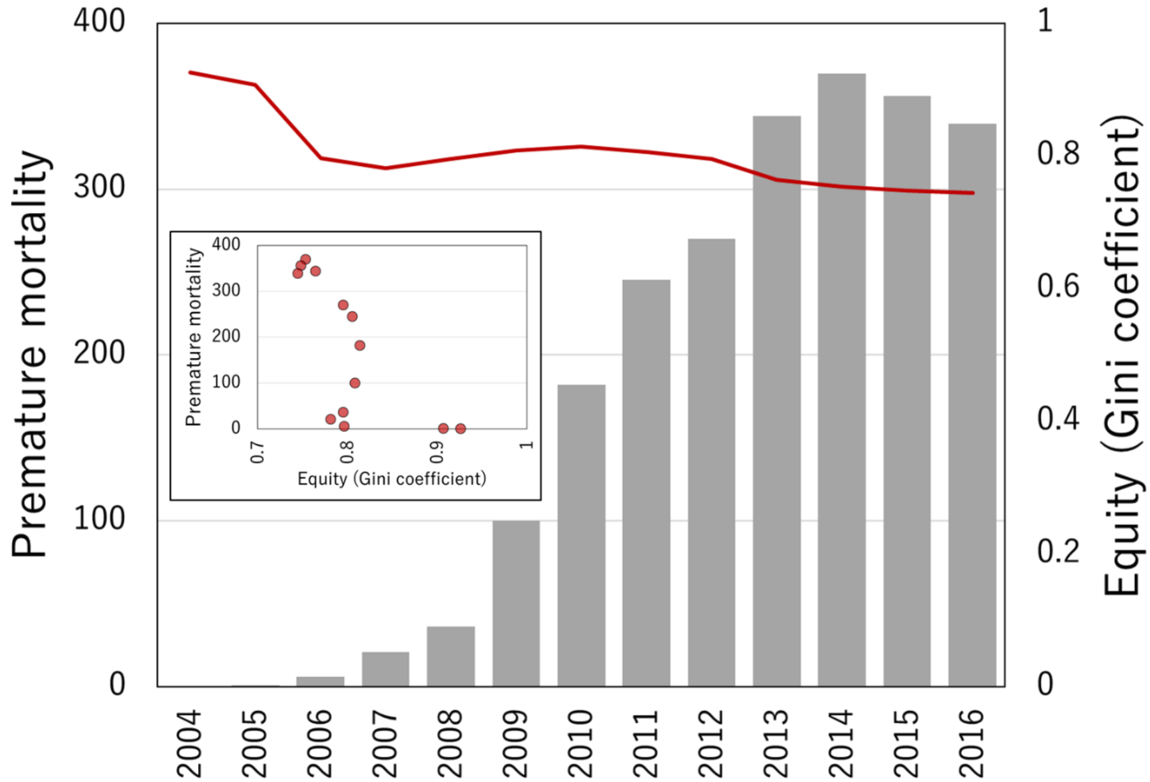

Figure C1. Air quality and equity tradeoff. Gini coefficients are used as a measure for spatial equity between all counties (with mortality >0.1), which are represented by the red line. Grey bars are annual premature mortality estimates. Based on mortality estimates from APSCA source-receptor RCM and ACS C-R relationship model specification.

Table C3. Spatial and temporal equity metrics for employment impacts.

Equity Unit	Gini Coefficient	Atkinson Index (ε=0.5)	Maximum Absolute Difference	Theil Index	Coefficient of Variation	Squared Coefficient of Variation
Producing counties						
County (2010)	0.89	0.75	7,000	1.08	2.91	10.6
County (2016)	0.88	0.76	6,000	1.01	3.26	7.57
County (cumulative)	0.88	0.75	46,000	1.08	2.75	8,48
Year	0.44	0.24	76,000	0.38	0.77	0.60
Appalachia counties						
County (2010)	0.72	0.46	7,000	1.01	1.82	3,31
County (2016)	0.72	0.46	6,000	0.96	1.65	2.72
County (cumulative)	0.72	0.45	46,000	1.00	1.79	3.20
Year	0.44	0.24	76,000	0.38	0.77	0.60

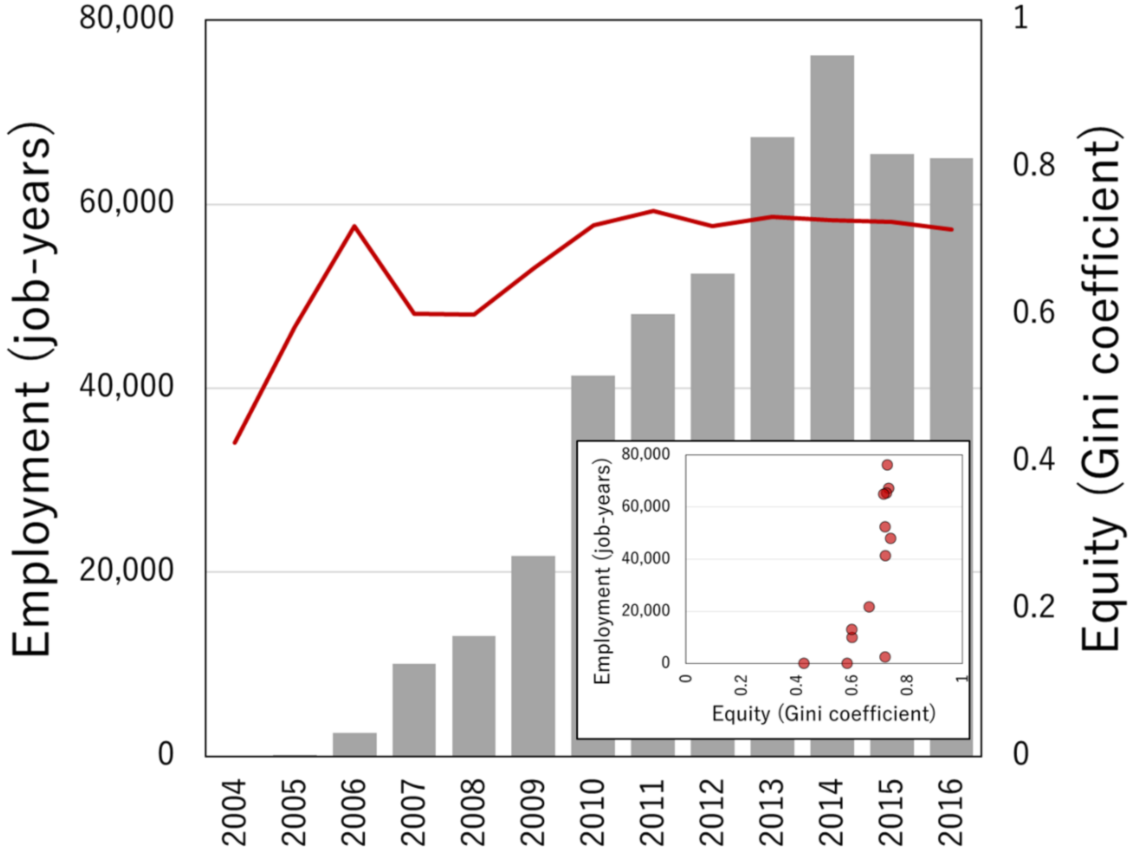

Figure C2. Employment and equity tradeoff. Gini coefficients are used as a measure for equity, which are represented by the red line. Gini coefficients are determined for each year and are estimated across all producing counties in a given year. Gray bars represent mean annual employment estimates.

Table C4. Climate change equity metrics. Based on global temperature change estimates.

Equity Unit	Gini Coefficient	Atkinson Index (ε=0.5)	Maximum Absolute Difference	Theil Index	Coefficient of Variation	Squared Coefficient of Variation
Equity Horizon: 2004 to 2016						
Year	0.72	0.55	0.000	0.93	1.53	2.35
Equity Horizon: 2004 to 2100						
Year	0.23	0.10	0.001	0.12	0.43	0.19
Decade	0.21	0.08	0.01	0.12	0.41	0.16
Generation	0.12	0.01	0.01	0.03	0.24	0.06
Equity Horizon: 2004 to 2200						
Year	0.17	0.06	0.001	0.08	0.36	0.13
Decade	0.17	0.05	0.01	0.08	0.35	0.13
Generation	0.11	0.01	0.01	0.02	0.23	0.05
Century	0.05	0.003	0.01	0.01	0.10	0.01

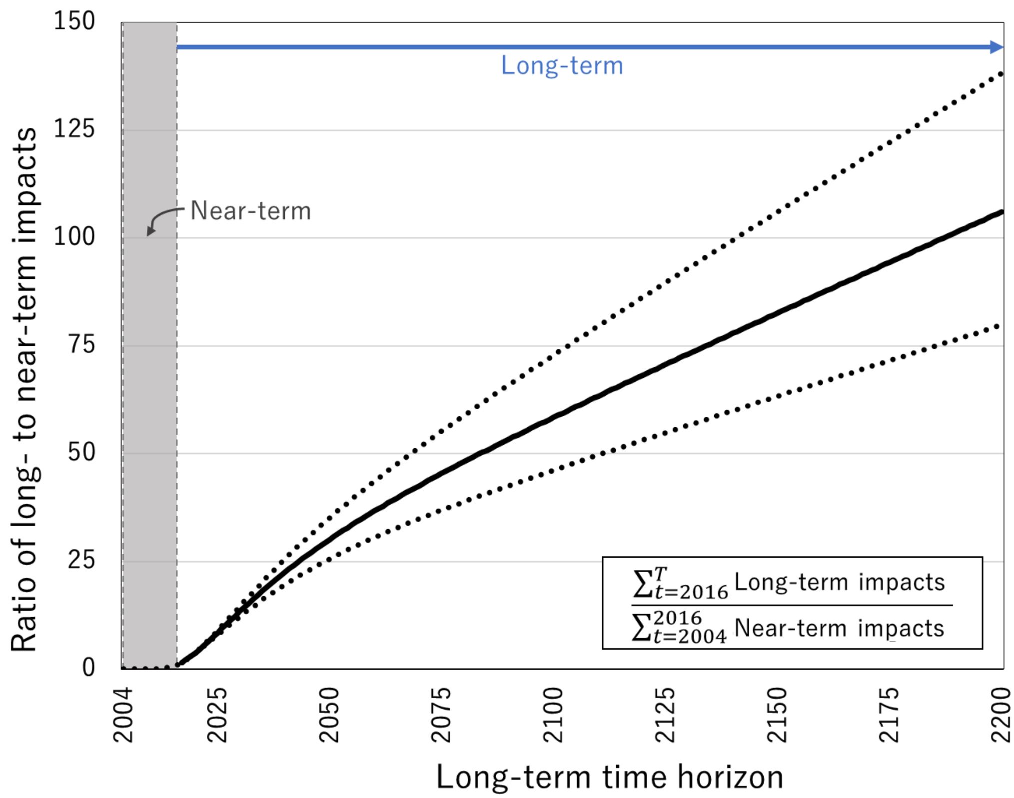

Figure C3. Ratio of long- to near-term impacts under varying time horizons for estimating long-term impacts.

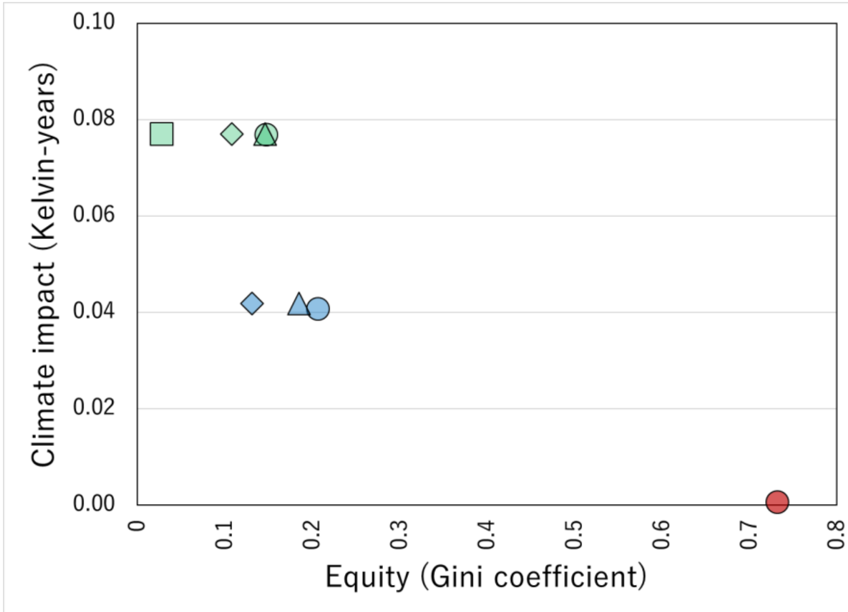

Figure C4. Climate change impacts and equity tradeoff. Point shapes represent equity unity: year (circle), decade (triangle), generation (diamond), and century (square). Point colors represent equity horizon: 2004 to 2016 (red), 2004 to 2100 (blue), and 2004 to 2200 (green).

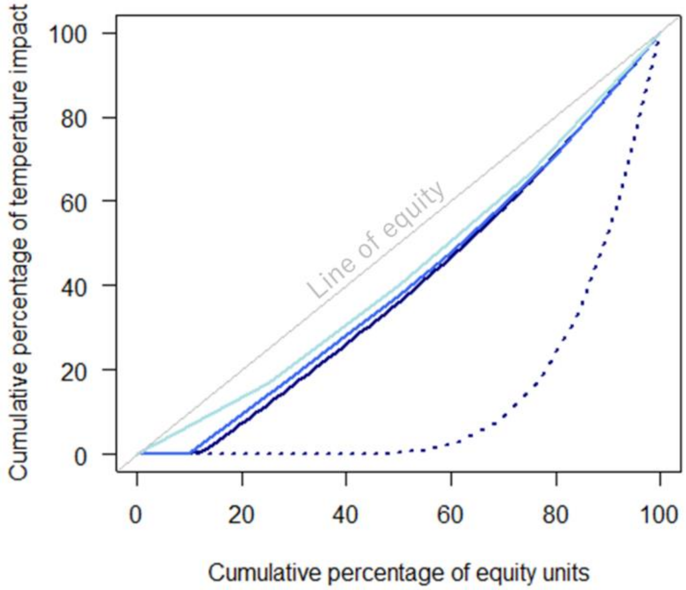

Figure C5. Lorenz curves representing the temporal equity of temperature impacts. The color of the lines represents the equity unit: year (dark blue), decade (medium blue), or generation (light blue). An equity unit is the period over which impacts are integrated. The type of line represents the equity horizon: dotted (2004 to 2016) and solid (2004 to 2100). The thin grey line is the line of equity.

C1.2 Distributive equity with respect to racial and socioeconomic subpopulations

C1.2.1 Population-weight mortality and mortality rates by race, income, and poverty level

Table C5. Summary of parameters, data limitations, and data sources for air quality equity analysis by subpopulation.

Subpopulation	Modified parameters	Data limitations	Data sources
Race (white, black or African American, Asian or Pacific Islander, Native American or Alaskan Native)	Mortality rates by race, year, county AND Population by race, year, county	59% reporting of mortality rates by race, year, county, with other values suppressed; to fill in missing data, used similar hierarchical approach as used in BenMap guidance. ~17% reporting mortality by age, race, county, year:. did not use age stratification.	CDC[18]
Income level (<$15,000, >$15,000 to <$35,000, >$35,000 to <$75,000, >$75,000 to <$150,000, >$150,000)	Population by income, year, county	-	US Census Bureau[19]
Poverty level (above, below poverty line)	Population by poverty level, age, year, and county	Poverty level age classes differ from mortality rate age classes; directly match age classes or use an age-weighted average.	US Census Bureau[19]

Table C6. Estimated population-weighted mortality for 2010, 2016, and cumulatively from 2009 to 2016 by race, household income, and poverty level. Population-weighted estimates based on annual, county-level premature mortality derived from AP3 and ACS study dose-response. Estimates in terms of premature mortality and values in parentheses are percent attribution. Note that total premature mortality estimates for each subpopulation are not necessarily the same, given differences in population estimates.

	2010	2016	Cumulative from 2009 to 2016
Total	182 (100%)	339 (100%)	2206 (100%)
Race			
White	157 (86%)	533 (85%)	1948 (86%)
Black or African American alone	22 (12%)	92 (13%)	286 (13%)
American Indian and Alaska Native	2 (1%)	6 (1%)	32 (1%)
Asian or Pacific Islander	1 (0%)	1 (0%)	7 (0%)
Household income			
<$15,000	24 (13%)	43 (13%)	285 (13%)
>$15,000 to <$35,000	39 (21%)	69 (20%)	462 (21%)
>$35,000 to <$75,000	58 (32%)	104 (31%)	686 (31%)
>$75,000 to <$150,000	45 (25%)	87 (26%)	554 (25%)
>$150,000	16 (9%)	37 (11%)	219 (10%)
Poverty level			
Below poverty line	22 (13%)	49 (14%)	304 (14%)
Above poverty line	149 (87%)	290 (86%)	1868 (86%)

C1.2.2 Income distribution and poverty level regression

Table C7. Comparison of boom and non-boom counties with respect to changes in poverty level and income Gini coefficients from 2005 to 2015.

Variables	Boom Counties (n = 90)		Non-Boom Counties (n = 168)		p-value[a]
	Mean	SD	Mean	SD	
Change in % below poverty level from 2005 to 2015	0.599	2.157	1.683	1.519	0.000
Log of change in % below poverty level from 2005 to 2015	-0.205	0.080	-0.169	0.071	0.000
Change in Gini coefficient from 2005 to 2015	-0.091	0.029	-0.080	0.026	0.013
Log of change in Gini coefficient from 2005 to 2015	-0.205	0.080	-0.169	0.071	0.004

a This p-value associated with testing if the means are different from each other, where the null hypothesis assumes that they are not different.

C1.2.3 Spatial coincidence of natural gas activity and demographic variables

Table C8. Demographic comparisons of mean and population-weighted mean populations in producing and nonproducing counties in 2010.

Variable	Population-Weighted Mean		Mean		T-Test	Significance
	Producing Counties (n=68)	Nonproducing Counties (n=204)	Producing Counties (n=68)	Nonproducing Counties (n=204)		
Population (1000s)	401	851	71	200	4.090	0.000
Percent black	5.2	13.8	2.3	4.9	4.624	0.000
Percent nonwhite	8.5	25.7	4.3	9.6	5.996	0.000
Per capita personal income ($)	39,507	45,860	33,120	37,071	4.185	0.000
Median household income ($)	44,033	53,245	39,222	47,350	8.147	0.000
Percent below poverty line	15.2	15.6	17.7	15.2	-4.204	0.000

Table C9. Demographic comparisons of mean and population-weighted mean populations in producing and nonproducing counties in 2016.

Variable	Population-Weighted Mean		Mean		T-Test	Significance
	Producing Counties (n=91)	Nonproducing Counties (n=181)	Producing Counties (n=91)	Nonproducing Counties (n=181)		
Population (1000s)	329	918	77	215	4.263	0.000
Percent black	5.5	14.4	2.9	5.1	3.388	0.000
Percent nonwhite	9.2	28.3	5.5	10.9	5.441	0.000
Per capita personal income ($)	42,687	54,159	36,710	42,775	5.370	0.000
Median household income ($)	49,375	61,907	44,461	54,701	8.951	0.000
Percent below poverty line	14.4	14.3	16.6	13.7	-5.226	0.000

Table C10. Spatial coincidence of production and demographic variables, using a logistic regression approach.

Variable	2010			2016			2010 to 2016		
	Estimate	Odds Ratio	P-value	Estimate	Odds Ratio	P-value	Estimate	Odds Ratio	P-value
Percent nonwhite	-0.132	0.876	0.060.	-0.046	0.955	0.108	-0.077	0.926	0.000***
Log of median household income	-8.447	0.000	0.009**	-8.685	0.000	0.000***	-5.213	0.005	0.000***
Percent below poverty line	-0.026	0.975	0.800	-0.118	0.888	0.100.	-0.036	0.964	0.147
Constant	30.672	2.09E+13	0.022.	35.069	1.70E+15	0.000***	20.031	5.01E+08	0.000***
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1									

Table C11. Median compensation, chief executive officer (CEO) compensation, and CEO-to-median employee compensation ratio for top producing firms in Appalachia. Based on 2017 Security and Exchange Commission Schedule 14A filings.[20]

Firm	2015 production (bcf)	2016 production (bcf)	CEO compensation ($)	Median employee compensation[a] (%)	CEO-to-median employee compensation ratio
Chesapeake Appalachia[a]	1,027	983	14,903,906	118,761	125
EQT Production Company	797	573	8,254,140	102,470	81
Cabot Oil and Gas Corp.	685	638	12,401,033	75,891	163
Antero Resources Corp.	683	581	9,925,217	87,186	114
Southwestern Energy Production Co,	541	538	8,687,476	108,458	80
Range Resources Appalachia LLC[c]	474	415	8,700,000	123,500	70
Rice Drilling LLC[d]	317	214			
Gulfport Energy Corporation	314	235	4,738,156	90,439	52
CNX Gas Company LLC[e]	302	217	10,585,778	129,390	82
Chief Oil and Gas LLC[f]	284	272			
Talisman Energy USA Inc.[g]	192	196			
Chevron Appalachia LLC[h]	177	181			
Anadarko E&P Onshore LLC[i]	166	148	16,959,896	160,251	106
Seneca Resources Corp.[j]	150	133			
XTO Energy, Inc.	146				
Noble Energy Inc.	797	139	11,262,048	127,488	88
Average			10,641,765	112,383	96

a Each firm has discretion over how the median employee compensation is determined. The median compensation is based on salary, taxable income, or earnings, and often includes stock incentives, bonuses, retirement benefits, employer matching, and/or overtime.
b Values are based on filings for Chesapeake Energy, the parent company of Chesapeake Appalachia.
c Values are based on filings for Range Resources, the parent company of Range Resources Appalachia.
d Acquired by EQT and did not file as Rice Drilling in 2017.
e Values are based on filings for CNX Resources, the parent company of CNG Gas Company LLC.
f Privately held.
g Acquired by Repsol. Canadian company.
h Private subsidiary of Chevron Corp.
i Values are based on filings of Anadarko, the parent company of Anadarko E&P Onshore LLC.
j Subsidiary of National Fuel which would not be representative of Seneca.

C1.3 Pairwise tradeoffs

C1.3.1 Air quality and employment tradeoffs

Table C12. Dataset descriptions for regression model to derive marginal effect of premature morality on job-years. Descriptions of the mortality estimates and county and year aggregations. These datasets are used both in the air quality and employment elasticity modeling and regressions for determining marginal effect of premature mortality on employment.

Model	Mortality estimates	County and year aggregation
1	Receptor-resolved mortality from supply chain activity	Observations (n=101) for each producing county aggregated over years
2	Receptor-resolved mortality from upstream activity	Observations (n=101) for each producing county aggregated over years
3	Source-resolved mortality from supply chain activity	Observations (n=210) for each source county aggregated over years
4	Source-resolved mortality from upstream activity	Observations (n=210) for each source county aggregated over years
5	Mortality from all supply chain activity	Observations (n=13) for each year aggregated over counties
6	Mortality from upstream activity	Observations (n=13) for each year aggregated over counties

Table C13. Linear model fits for premature mortality and job-years tradeoff. Linear fits ($y = \beta 0+\beta x+\varepsilon$) for 6 different models (as described in Table C12). β are in units of job-years per premature mortality.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
β (mean ± se)	17.8 ±28.7	55.7 ±57.7	94.1 ±14.2	352 ±19.8	157 ±4.68	826 ±163
Constant (mean ± se)	4453 ±916	4401 ±901	677 ±470	1047 ±290	2982 ±1274	2586 ±7939
Adjusted R^2	-0.006	-0.001	0.17	0.60	0.99	0.67
n	101	101	210	210	13	13

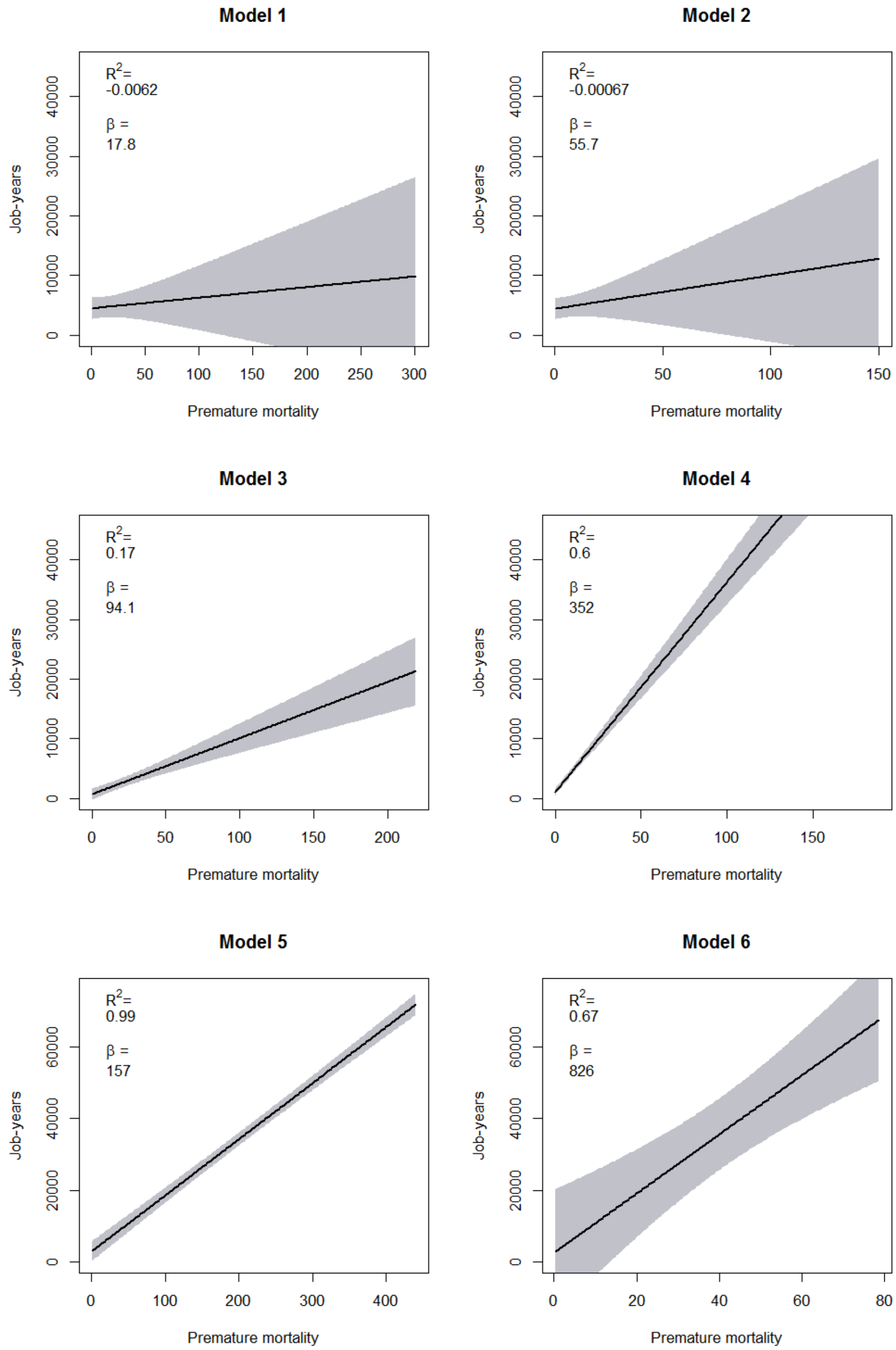

Figure C6. Linear models for premature mortality and job-years tradeoff. Linear fits for 6 different models (as described in Table C12). Linear models ($y = \beta_0 + \beta x + \varepsilon$) are indicated by black lines and within model 95% confidence intervals are indicated by grey shaded regions. The adjusted R^2 and mean value of β (in units of job-years per premature mortality) are provided.

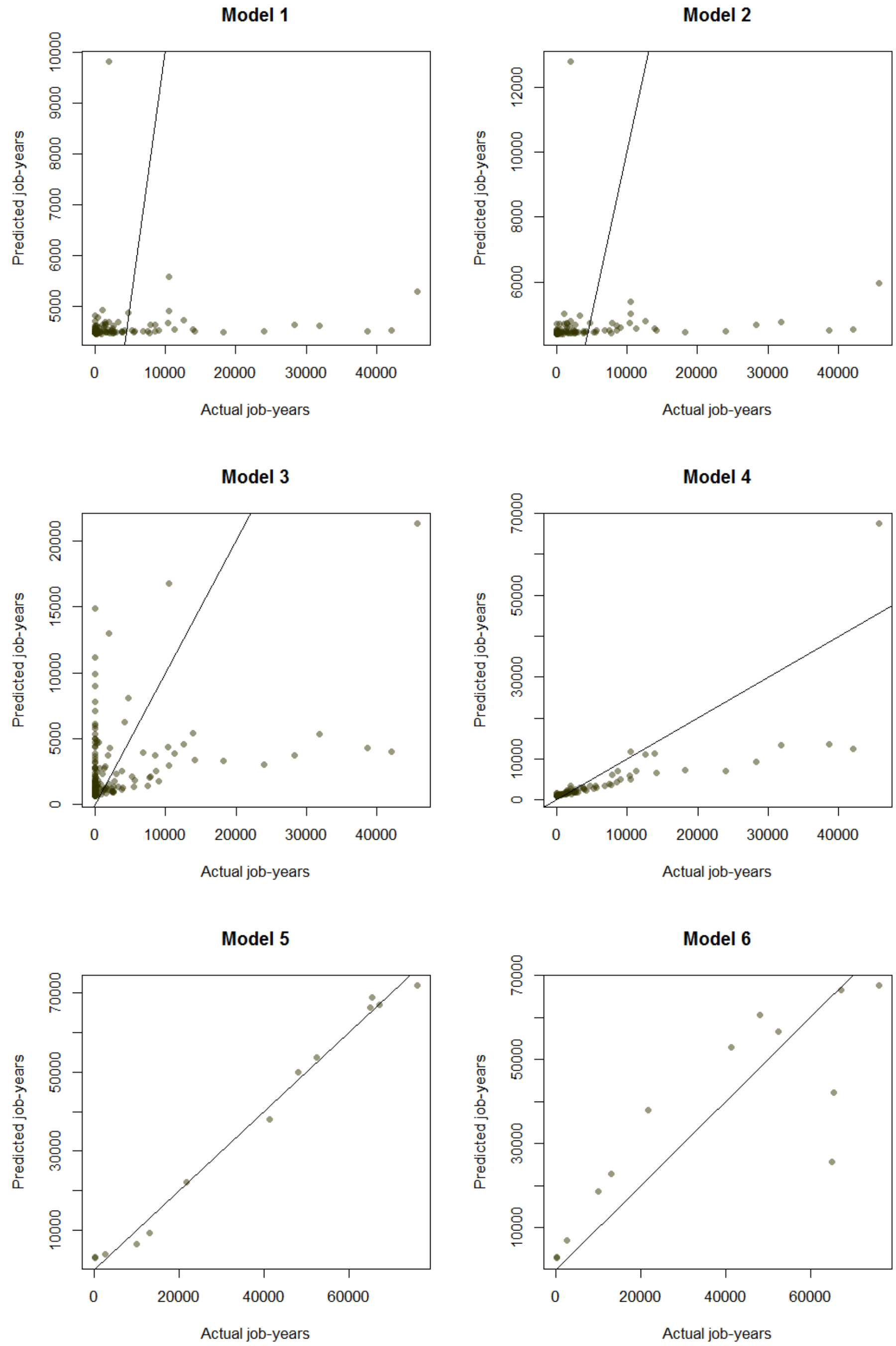

Figure C7. Actual versus predicted values for linear models for premature mortality and job-years tradeoff. Linear fits for 6 different models (as described in Table C12). Dots represent actual and predicted job-years, and black line is a line with slope = 1 and intercept = 0 (indicating perfect fit).

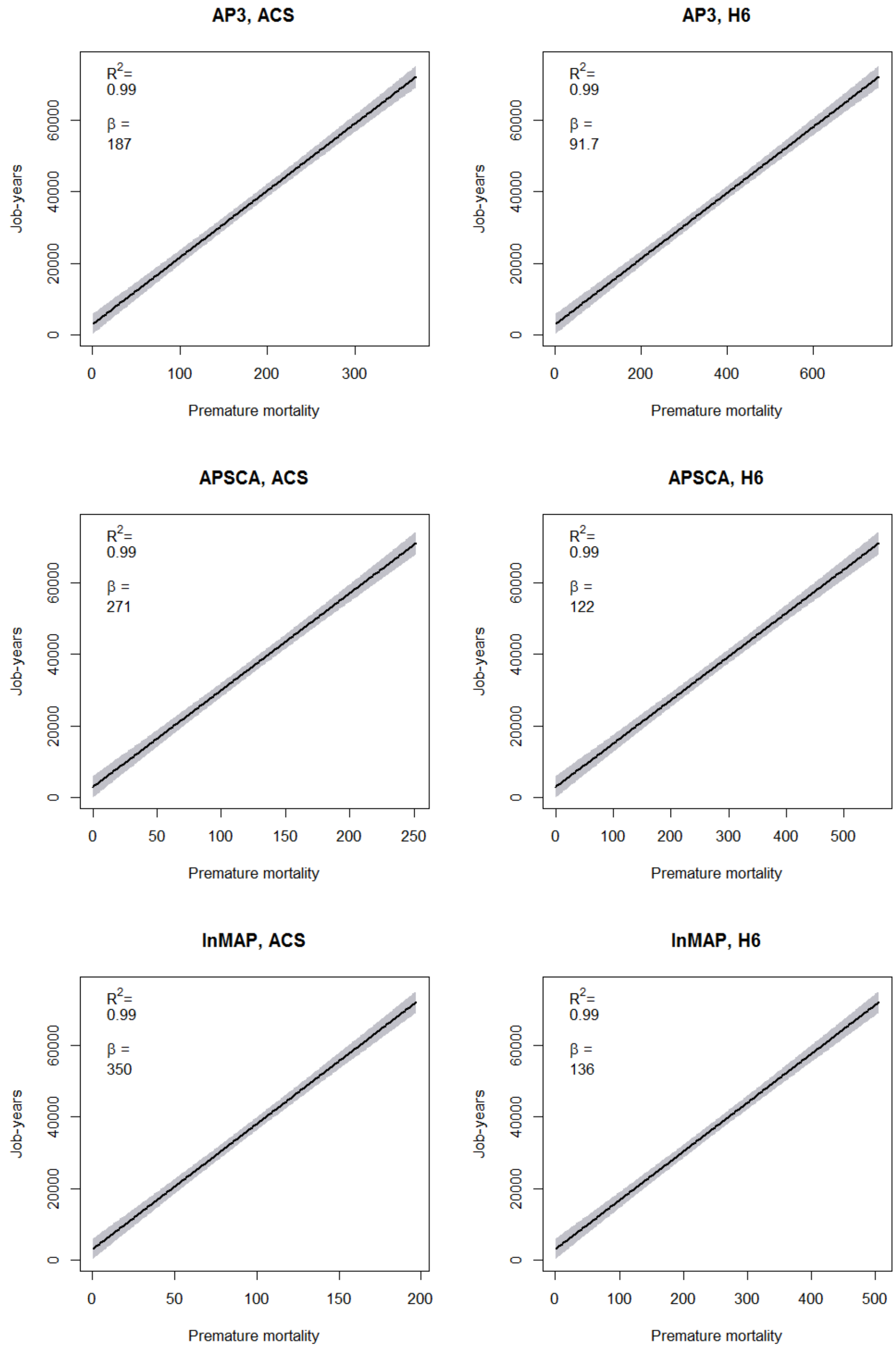

Figure C8. Sensitivity of linear models for premature mortality and job-years tradeoff with respect to air quality model and concentration-response (C-R) relationship. These are all linear fits for model specification 5 (as described in Table C12). We assume two different C-R relationships, American Cancer Society (ACS) and Harvard Six Cities (H6C), and three different reduced complexity air quality models, AP3, APSCA, and InMAP. Linear models ($y = \beta_0 + \beta x + \varepsilon$) are indicated by black lines and within model 95% confidence intervals are indicated by grey shaded regions. The adjusted R^2 and mean value of β (in units of job-years per premature mortality) are provided.

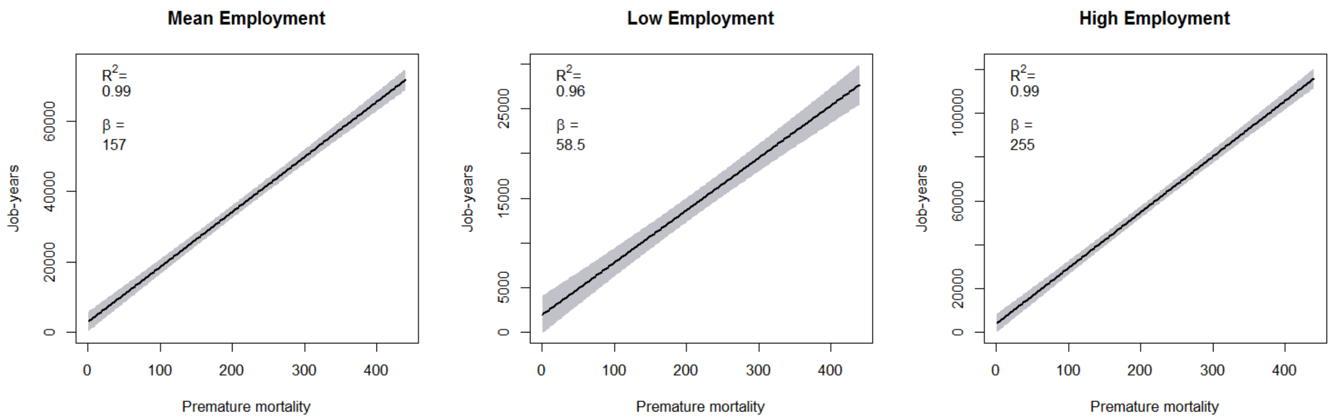

Figure C9. Sensitivity of linear models for premature mortality and job-years tradeoff with respect to employment estimates. These are all linear fits for model specification 5 (as described in Table C12). We assume mean, upper 95% CI, and lower 95% CI estimates of employment, as reported in Mayfield et al. (forthcoming). Linear models ($y = \beta_0 + \beta x + \varepsilon$) are indicated by black lines and within model 95% confidence intervals are indicated by grey shaded regions. The adjusted R^2 and mean value of β (in units of job-years per premature mortality) are provided.

C2 References

1. U.S. Energy Information Adminstration. *U.S. Crude Oil and Natural Gas Proved Reserves, Year-end 2016*. (2018).
2. U.S. Geological Survey. National Assessment of Oil and Gas Project - Appalachian Basin Province (067) Boundary. (2002).
3. U.S. Geological Survey. National Assessment of Oil and Gas Project Devonian Marcellus Shale of the Appalachian Basin Province (067) Assessment Units. (2011).
4. U.S. Geological Survey. National Assessment of Oil and Gas Project - Utica Shale Unconventional Assessment. (2015).
5. Pennsylvania Department of Environmental Protection. PA DEP Oil & Gas Reporting Website. Available at: https://www.paoilandgasreporting.state.pa.us/publicreports/Modules/Welcome/Welcome.aspx.
6. Ohio Department of Natural Resources. Oil & Gas Resources. Available at: http://oilandgas.ohiodnr.gov/production.
7. West Virginia Department of Environmental Protection. Oil and Gas Production Data. Available at: https://dep.wv.gov/oil-and-gas/databaseinfo/Pages/default.aspx.
8. U.S. Energy Information Adminstration. Natural Gas Gross Withdrawals from Gas Wells. Available at: http://www.eia.gov/dnav/ng/ng_prod_sum_a_epg0_fgg_mmcf_a.htm%0A.
9. U.S. Energy Information Adminstration. International and Interstate Movements of Natural Gas by State. Available at: http://www.eia.gov/dnav/ng/ng_move_ist_a2dcu_swv_a.htm%0A.
10. Fullerton, D. Distributional Effects of Environmental and Energy Policy : An Introduction. (2009).
11. U.S. Environmental Protection Agency. Technical Guidance for Assessing Environmental Justice in Regulatory Analysis. (2016).
12. Klinsky, S. & Dowlatabadi, H. Conceptualizations of justice in climate policy. **9,** 88–108 (2009).
13. Allison, P. D. Measures of Inequality Author (s): Paul D . Allison Source : American Sociological Review , Vol . 43 , No . 6 (Dec ., 1978), pp . 865-880 Published by : American Sociological Association Stable URL : http://www.jstor.org/stable/2094626. *Am. Sociol. Rev.* **43,** 865–880 (1978).
14. Theil, H. The Information Approach to Demand Analysis Author (s): H . Theil Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/1911889 Accessed : 07-06-2016 14 : 08 UTC Your use of the JSTOR archive indicates your acceptance of the Te. *Econometrica* **33,** 67–87 (1965).
15. Dalton, H. The Measurement of the Inequality of Incomes Author (s): Hugh Dalton Published by : Wiley on behalf of the Royal Economic Society Stable URL : http://www.jstor.org/stable/2223525. *Econ. J.* **30,** 348–361 (1920).
16. Litchfield, J. a. Inequality : Methods and Tools. *Entropy* (1999). doi:http://siteresources.worldbank.org/INTPGI/Resources/Inequality/litchfie.pdf
17. Levy, J. I., Chemerynski, S. M. & Tuchmann, J. L. Incorporating concepts of inequality and inequity into health benefits analysis. *Int. J. Equity Health* **5,** 2 (2006).
18. Centers for Disease Control and Prevention, N. C. for H. S. Underlying Cause of Death 1999-2016 on CDC WONDER Online Database, released December, 2017. Data are from the Multiple Cause of Death Files, compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Prog. (2018). Available at: http://wonder.cdc.gov/.

19. US Census Bureau. American Community Survey. (2018). Available at: factfinder.census.gov.

20. Security and Exchange Commission. Schedule 14A filings. Available at: Security and Exchange Commision.