ENVIRONMENTAL POLICY ACROSS TERRESTRIAL SPACE

by

JOHN MATTHEW MOREHOUSE

A DISSERTATION

Presented to the Department of Economics
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2022

DISSERTATION APPROVAL PAGE

Student: John Morehouse

Title: Environmental Policy Across Terrestrial Space

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Economics by:

Trudy Cameron	Co-Chair
Mark Colas	Co-Chair
Keaton Miller	Core Member
Edward Rubin	Core Member
Grant Jacobsen	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies

Original approval signatures are on file with the University of Oregon Division Graduate Studies.

Degree awarded June 2022

DISSERTATION ABSTRACT

John Matthew Morehouse

Doctor of Philosophy

Department of Economics

June 2022

Title: Environmental Policy Across Terrestrial Space

This dissertation examines spatial heterogeneity that results from various environmental policies. In Chapter 1, I provide a comprehensive overview of each dissertation chapter.

Chapter 2 (with Ed Rubin) demonstrates that most coal-fueled power plants are located on or near jurisdictional (county or state) borders. We find that coal-fired power plants are disproportionately sited on downwind borders (within county or state). Natural gas plants—much lower polluters—do not exhibit this behavior. Motivated by the inferred strategic siting, we use an atmospheric dispersion model developed by NOAA to estimate various aspects of the "pollution transport problem." We find that nearly 90% of coal-based particulate matter leaves its *state* of origin within 48 hours of release.

Chapter 3 (with Mark Colas) examines the effects of stringent land-use regulations on national carbon emissions. We develop and estimate a general equilibrium model of residential sorting and energy consumption. We find that relaxing land-use restrictions in California leads to a 0.6% drop in national carbon emissions. The mechanism behind this drop is straightforward. California cities have a temperate climate, carbon-efficient power plants, and high land-

use regulations. These land-use regulations inflate housing prices, thus keeping households out of California cities. When households live outside of California, they emit more carbon on average, and therefore national carbon emissions are higher due to California's land-use regulations.

In Chapter 4, I simulate the labor market effects of a carbon tax across the continental United States. To recover the welfare impacts of a carbon tax, I build and estimate a spatial equilibrium model that features heterogeneous households. I incorporate a rich level of heterogeneity into the model that allows me to answer: (1) *who* is most affected by a carbon tax, (2) *how much* the burden of a carbon tax is borne on different households, and (3) *where* the households are that bear the greatest burden from the tax. I find that workers without a college degree in manufacturing bear a disproportionate share of the tax incidence. Chapter 5 concludes this dissertation.

This dissertation includes previously both previously published and unpublished and co-authored material.

CURRICULUM VITAE

NAME OF AUTHOR: John Matthew Morehouse

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR, USA

DEGREES AWARDED:

Doctor of Philosophy, Economics, 2022, University of Oregon
Master of Science, Economics, 2018, University of Oregon
Bachelor of Science, Economics, 2016, University of Oregon

AREAS OF SPECIAL INTEREST:

Environmental Economics
Labor Economics
Urban Economics

PROFESSIONAL EXPERIENCE:

Freelance Economist, Vivid Economics, 2020
Intern, Vivid Economics, 2020
Research Assistant to Dr. Edward Rubin, University of Oregon, 2020
Finance Associate, Summit Bank, 2017

GRANTS, AWARDS AND HONORS:

Thomas J. Sargent Dissertation Fellow, Federal Reserve Bank of San Francisco, 2021
Kleinsorge Summer Research Award, University of Oregon, 2021
Best graduate student research paper award (Economics), University of Oregon, 2020
General University Scholarship, University of Oregon, 2019
Kleinsorge Summer Research Award, University of Oregon, 2019

Edward G. Daniel Scholarship, University of Oregon, 2018
Graduate Teaching Fellowship, University of Oregon, 2017-2022

PUBLICATIONS:

Colas, M., & Morehouse, J. (2022). The Environmental Cost of Land Use Restrictions *Quantitative Economics*, 13(1): 179-223.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my committee chairs, Trudy Cameron and Mark Colas, who provided endless support and guidance throughout my graduate school career. I would also like to thank Keaton Miller, Edward Rubin, and Grant Jacobsen for providing valuable feedback on various parts of this dissertation. I would also like to thank my undergraduate mentors, Andrew Schopieray and Ralph Mastromonaco, for their initial and continued belief in me. Lastly, I am grateful for the boundless support from my family and friends. I am particularly grateful to Annie and Olive.

To my family

TABLE OF CONTENTS

Chapter Page

Chapter Page

Chapter Page

Chapter Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

LIST OF TABLES

Table Page

Table Page

CHAPTER I

INTRODUCTION

My dissertation seeks to provide insight into critical policy questions related to the regulation of air quality and climate. I examine spatial heterogeneity in response to (and as a result of) a set of environmentally related public policies. I use various methods such as structural models, atmospheric dispersion modeling, and reduced-form causal inference to answer these questions. This dissertation contains previously published and unpublished co-authored material. Chapter 2 is joint work with Edward Rubin, and Chapter 3 is joint work with Mark Colas.

Chapter 2 demonstrates that most coal-fueled power plants are located on or near jurisdictional (county or state) borders. Water is a crucial input for electricity production, and many borders are water (i.e., rivers), so it is entirely possible that this observed border siting results from minimizing input costs (and not regulatory avoidance). We develop a simple statistical test that uses variation in wind direction to distinguish strategic from non-strategic siting. We find that coal-fueled electricity generators have been sited strategically to export their emissions beyond the boundaries of their counties and states. Natural gas plants—much lower polluters—do not exhibit this behavior.

Motivated by this apparent strategic siting, we use an atmospheric dispersion model developed by NOAA to estimate various aspects of the "pollution transport problem." We find that nearly 90% of coal-based particulate matter leaves its *state* of origin within 48 hours of release. We also document that up to 20% of coal-based particulates in some counties can be attributed to coal-based electricity pollution in other counties. We then document the share of coal-based particulate matter in a given county by whether or not the county meets the

Environmental Protection Agencies' standards for local pollutants (under the Clean Air Act). Counties that are "out of attainment" (i.e., do not meet the standard) receive a substantial share of their total coal-based particulate matter emissions from plants in attainment counties in neighboring states. Our results highlight the importance of transport-focused regulation for local pollutants.

Chapter 3 shows that cities with restrictive land-use regulations tend to have low-carbon emissions per capita. We develop and estimate a spatial equilibrium model to understand the implications of these land-use regulations on national carbon emissions. Our model features heterogeneous households that consume energy and housing, and firms that employ college and non-college-educated labor as imperfect substitutes. The model also features locations that vary in the carbon efficiency of regional power plants, their amenities, and their marginal utility of energy consumption. We estimated the model on publicly available data.

We find that the relaxation of excessive land-use restrictions in California led to 0.6% lower overall national carbon emissions. The mechanism behind this drop is intuitive. California cities have more temperate climates, carbon-efficient power plants, and stringent land-use regulations. These land-use regulations inflate housing prices, thus lowering the equilibrium population levels in California cities (and raising the equilibrium population levels in states). When fewer people live in California, carbon emissions are higher because household-level carbon emissions are lower (on average) in California than in other states.

In Chapter 4, I simulate the labor market effects of a carbon tax across the continental United States. To recover the welfare impacts of a carbon tax, I build and estimate a spatial equilibrium model that features heterogeneous households. I incorporate a rich level of heterogeneity into the model that allows me to answer:

(1) *who* is most affected by a carbon tax, (2) *how much* of the burden of a carbon tax is borne by different households, and (3) *where* the households are that bear the greatest burden from the tax.

I find that manufacturing workers without a college degree can be expected to bear a disproportionate share of the tax incidence. Cities with mild climates, carbon-efficient power plants, and services-oriented economies experience modest population increases as households move in response to the carbon tax. Additionally, I use the model to demonstrate that, relative to flat transfers, progressive compensation leads to a decline in aggregate carbon emissions due to a reallocation of workers into less carbon-intensive cities and sectors.

CHAPTER II

DOWNWIND AND OUT: THE STRATEGIC DISPERSION OF POWER PLANTS AND THEIR POLLUTION

This chapter is co-authored with Edward Rubin. I had an essential role in developing the initial idea that led to this project. I also wrote a significant amount of code that generates the results in this paper. I wrote and edited many sections of the paper. Furthermore, I wrote a successful application for external funding for this project.

2.1 Introduction

Federalist systems offer potential efficiencies in many settings, but they also may incentivize strategic responses from local governments—whose focus tends to emphasize the provision of locally enjoyed goods (Oates, 1972, 1999). As a consequence, local administrative units may seek to export the negative externalities generated by locally beneficial economic activities. The extent to which the local units can export their costs—and increase *local* net welfare—will depend on the degree to which local actors can separate the externalities from the productive activities themselves.

Consequently, federalist regulatory systems face two important challenges when governing air quality. First, local governments face few incentives to internalize the costs of pollution once it leaves their jurisdictions Monogan, Konisky, and Woods (2017); Oates (1972); Revesz (1996); Tiebout (1956). Second, air pollution can travel long distances (i.e., crossing city, county, and state borders) (Oates, 2002; Sergi, Azevedo, Davis, & Muller, 2020; United States Senate, Committee on Public Works, Staff Report, 1963). In the U.S., the spatially discontinuous patchwork of local and state authorities presents many

opportunities for local decision-makers to strategically site major polluters in locations that reduce air-pollution exposure within the county/state. As a result, local governments have incentives to site polluters where the jurisdiction can simultaneously enjoy the benefits of production (e.g., increased jobs and wages) while exporting the pollution costs. This hypothesized behavior is, in a sense, a variation on NIMBY-ism: the property owner wants the activity on her property but wishes to export the negative externalities.[1]

In two steps, we empirically substantiate the hypothesis that decision-makers attempt to capture local benefits and export their negative externalities. First, we identify strategic siting within a significant group of air polluters in the United States—demonstrating that decision-makers sited coal-fired power plants to reduce *downwind* pollution exposure within their own counties and states. Establishing this result is necessary to demonstrate strategic exporting of externalities, but it is not sufficient—externalities must be sufficiently exportable. We document the extreme mobility of the pollution generated by these plants. We first show that governments tend to site these major polluters near the downwind border of administrative units, and then we quantify the extent to which these polluters' emissions are carried downwind from their source counties and states. Together, these two tendencies support our hypothesis that local decision-makers make siting choices that take advantage of polluters' benefits while minimizing the costs to their own constituencies.

[1] NIMBY is "Not In My BackYard," as used by (Gates, 1980; Livezey, 1980; Mitchell & Carson, 1986)—and many individuals since.

We focus on coal-fired power plants, which historically have accounted for a substantial share of air pollution in the U.S.[2] Coal-fired electricity production offers a classic example of a negative externality: The plant's operators and immediate community enjoy positive economic benefits, while counties and states downwind of the plant bear the costs of the plant's pollution. The context of coal-fired electicity generating units (EGUs) provides several advantages. First, natural-gas-fired power plants—which produce much less pollution than their coal-fired counterparts—provide a helpful 'control group' in our empirical framework. Second, while coal and natural gas EGUs both use water as an input, neither type of EGU uses do not use *areas downwind or upwind* as (non-strategic) inputs—a fact we exploit in our empirical tests. Third, electricity generators are required to record important emissions data—unlike many other major polluters. These emissions records are advantageous when we model pollution transport for coal-fired power plants.

We first document that electricity generators tend to be sited near administrative borders. Given that water both (a) forms many administrative borders and (b) is a key input to electricity generation—thus affecting EGU siting—we develop a simple, non-parametric test that shows localities (states and counties) sited coal plants to reduce within-unit, downwind exposure. Natural gas EGUs do not exhibit this behavior. In other words, while water may explain coal power plants' proximity to borders, it does not explain their tendency to be sited on *downwind* borders. Our natural gas placebo test corroborates this finding. Finally, using a state-of-the-art particle-trajectory model, we illustrate the extreme *exportabilitiy* of coal plants' pollution: within six hours, 50% of coal

[2]E.g., in 2014, U.S. coal EGUs accounted for approximately 65.7% of SO_2 emissions, 44.0% of mercury emissions, 39.1% of arsenic emissions, and 10.6% of NO_x emissions in the United States U.S. Environmental Protection Agency (2018)—while contributing only 39% of total electricity generation (EIA, 2021).

plants' emissions leave their source states—and 99% depart their source counties.[3] Together, these results illustrate critical challenges facing decentralized, federalist approaches to administration and regulation. More broadly, we find significant evidence that local decision-makers strategically respond to the spatial patchwork of jurisdictions created by the federalist system in the U.S.

Our results parallel a growing literature documenting strategic pollution-related responses in federalist systems. This nascent literature has so far identified three main varieties of strategic responses by local decision-makers and polluters: (1) strategic siting of polluting plants (Monogan et al., 2017), (2) strategic production or abatement decisions (Zou, 2021), and (3) strategic monitoring (Grainger, Schreiber, & Chang, 2018; Mu, Rubin, & Zou, 2021). Each of these strategic responses implies different costs and requires different remedies. For example, Zou (2021) provides evidence that scheduled intermittent monitoring leads to significantly lower pollution levels on monitored days (relative to unmonitored days). Consequently, air-quality levels near intermittent monitors are likely worse than monitoring data would suggest. Grainger et al. (2018) find that the siting of air-quality monitors is vulnerable to strategy—again, resulting in an underestimate of local ambient air pollution. Mu et al. (2021) detect a set of monitors that appear to shut down in anticipation of high-pollution events—also biasing air-quality estimates downward. Broadly, this literature suggests that current regulatory and

[3]One might wonder whether this degree of exportability of emissions makes local strategic siting irrelevant. A plant's pollution tends to leave its county quickly, siting is still relevant within the county. For local decision-makers and residents, there is a substantial difference between (a) a plant's pollution passing through/over the major city within the county and (b) the plant's pollution immediately exiting the county. The same reasoning also applies at the state level.

political structures create opportunities for polluters and local decision-makers to avoid fully internalizing pollution-based costs.[4]

Our paper most closely relates to Monogan et al. (2017). Like us, these authors find significant evidence that industrial facilities with large emissions systematically locate closer to states' downwind borders relative to lower-emissions industrial facilities. However, our analysis differs from Monogan et al. (2017) in four important ways. First, we define "strategic siting" (within a jurisdiction, i.e., state or county) as choosing a plant location where the downwind area is less than the upwind area (in the given jurisdiction)—based upon the location's prevailing wind.[5] Comparing the area downwind to the area upwind—within the same jurisdiction—implicitly controls for the size of the jurisdiction. In contrast, Monogan et al. (2017) focus on polluters' *distance* to the state's "downwind border." Second, we study strategic siting at both the county and the state level, while Monogan et al. (2017) focus only on state-level siting. We are unaware of any existing analyses that detect within-county strategic siting. We believe both levels warrant consideration. State and county governments each potentially face incentives to mitigate pollution exposure—e.g., counties are often the most basic unit of air-quality regulation, while state agencies coordinate county-level responses to regulation. Beyond regulation, politicians at every level face political incentives

[4]Our paper also broadly relates to a large literature on the pollution-haven hypothesis (PHH), which posits that polluters tend to locate in areas with less stringent environmental regulation. Much of the PHH literature investigates this hypothesis at the international level—focusing on how emissions-intensive production shifts towards countries with lax environmental regulation. Cherniwchan, Copeland, and Taylor (2017); Cole (2004); Levinson (2008); Millimet and Roy (2015) provide helpful overviews and discussions of the PHH literature. Our main hypothesis—that local decision-makers site polluters to capture economic benefits while exporting pollution's costs—follows a similar line of reasoning as the PHH but focuses on within-unit spatial siting decisions (enabling the export of pollution) rather than variation in regulatory stringency.

[5]As we explain below, we define downwind/upwind area using 30-year averages for prevailing wind directions from NOAA (North American Regional Reanalysis, 2006).

to increase economic activity while maintaining some degree of environmental quality (producing health and amenity values). Third, we focus exclusively on electricity generators—and specifically compare coal EGUs to natural-gas EGUs. As described above, coal EGUs account for a substantial share of local and national air pollution ($PM_{2.5}$, NO_x, SO_2, mercury, lead, ozone, and CO).[6] Finally, we extend beyond both Monogan et al. (2017) and the current literature by including additional descriptions of the geography of power plants *and*, importantly, descriptions of the transport of coal EGUs' emissions across the United States.

Methodologically, our empirical test of strategic siting overlaps with a growing literature that uses wind direction for identification. For example, Zivin, Liu, Song, Tang, and Zhang (2020) use the difference between upwind and downwind agricultural fires in China to identify the effect of fire smoke on cognitive test performance. Rangel and Vogl (2019) use a similar approach to estimate the effects of fire smoke on infant health at birth. Schlenker and Walker (2016) and Anderson (2019) use upwind and downwind exposure to traffic-induced pollution (from planes and automobiles, respectively) to measure the effects of pollution on local health. Our test uses the ratio of downwind area to upwind area within the jurisdiction to identify strategic siting among major polluters.[7]

We are not the first to examine the challenges that pollution transport creates—e.g., the Clean Air Act of 1963 was understood to limit federal power to

[6]Consequently, coal EGUs are regulated and monitored closely by both federal (especially U.S. EPA) and local (state and county) authorities. In addition, coal EGUs are unique in their tendency to build tall smokestacks: there are 15 smokestacks in U.S. of at least 1,000 feet and nearly 300 smokestacks of at least 500 feet (CAMD, 2020; U.S. Government Accountability Office, 2011).

[7]Many other papers use wind variation (rather than a comparison of upwind to downwind areas) for causal identification, e.g., Barwick, Li, Rao, and Zahur (2018); Deryugina, Heutel, Miller, Molitor, and Reif (2019); Freeman, Liang, Song, and Timmins (2019); Holland, Mansur, Muller, and Yates (2019); Sullivan (2016).

cases where (1) "air pollution...originates in one state and adversely affects persons or property in another state" or (2) for "significant intrastate problems which state and local agencies are unwilling or unable to deal with" (Edelman, 1966). A host of "pollution transport" models have been developed to study the extent to which pollution travels, as well as the health and policy problems posed by pollution transit.[8] Sergi et al. (2020) find that despite national reductions in $PM_{2.5}$ from point sources since 2008, approximately 26% of counties have experienced worsening health damages from pollution—noting that "around 30% of all U.S. counties receive 90% of their health damages from emissions in other counties." Similarly, by decomposing pollution levels by each pollutant's distance from its source, Wang et al. (2020) find that "long-range" pollution is dominant in the U.S.[9] Nearly 60 years have passed since the CAA of 1963 recognized "the transport problem" in air pollution, but substantial gaps remain in our understanding of the origin and extent of the problem, or the damages that result.

More broadly, the evidence in this paper, in conjunction with the existing literature, highlights important policy challenges facing federalist systems. Local governments can export negative externalities 'abroad' when these externalities are physically separable from local benefits. We provide evidence of this behavior in an economically and historically important context: coal-fired power plants. First, we show significant evidence that U.S. counties and states sited coal power plants to reduce within-county and within-state downwind exposure. We then show that these plants—their locations, in combination with prevailing wind patterns and coal

[8] Another class of pollution transport models—reduced-complexity air transport models—make simplifying assumptions concerning meteorology and atmospheric chemistry equations in exchange for large computational benefits, e.g., the InMAP model (Tessum, Hill, & Marshall, 2017).

[9] Wang et al. (2020) define "long range" as farther than 100 km from the source—reasoning that this distance "likely represents regional background and long-range transport."

plants' tall smokestacks—export pollution quickly out of the source counties and states. By documenting this strategic behavior and illustrating the incentives that federalism's decentralization creates, our results identify areas where policymakers and regulators may reduce external costs and thereby increase net social benefits.

2.2 Institutions: Siting of Plants

Governments' and firms' decisions about where to site a new power plant depend upon a host of variables—proximity to water,[10] grid/transmission availability,[11] access to fuel[12] (e.g., rail lines, pipelines, wind/solar capacity), local regulatory oversight[13] (i.e., friendliness to industry), and local community characteristics as well[14]. In the rest of the paper, we will use "decision-makers" to refer to the joint government-firm decision process for siting a plant. A large literature considers how local environmental regulations and enforcement affect the locations of polluting firms across states and counties (see footnote 4). However, location decisions on a finer scale—i.e., within state or county—have received far less attention.

The logic of exporting negative externalities is simple. If a local decision-maker reduces the area downwind of polluters *within its administrative boundaries*,

[10] Steam-driven turbines and water-cooled plants mechanically require water. We document the distribution of plants' proximities to water in *Empirics* and Figure 1.

[11] In the Texas electricity market, Woerman (2020) demonstrates that grid congestion can induce market power—more than doubling firms' markups.

[12] Preonas (2019) documents markups driven by market power in coal-by-rail delivery to coal plants in the U.S.

[13] An abundant literature considers the effect of local pollution regulations on polluter locational choice—e.g., Becker and Henderson (2000); Gray (1997); Jeppesen and Folmer (2001); Jeppesen, List, and Folmer (2002); Levinson (1996); List, Millimet, Fredriksson, and McHone (2003); Mani, Pargal, and Huq (1997); McConnell and Schwab (1990); Millimet and List (2003); Shadbegian and Wolverton (2010).

[14] Wolverton (2009) finds a significant negative association between plant sitings and income.

then fewer of its own residents bear the costs of pollution.[15] As long as the polluters remain within administrative borders, the locality captures many of the plants' benefits—tax revenue, employment, economic activity/growth. In addition, by moving polluters farther downwind, the decision-maker may also complicate pollution attribution and regulation—reducing local regulatory costs associated with the emissions.[16] Broadly, this story follows a similar logic to NIMBY behavior: an actor (the local decision-maker) tries to enjoy the benefits associated with an economic activity without bearing the activity's costs. As we show, power plants can easily export their pollution (their main external costs) using wind and tall chimneys. Figure 2a shows an illustrative plant with limited downwind area (the dark purple shaded area) in its home county.

2.3 Data

Overview We combine several publicly available datasets that originate from a variety of federal agencies. The data fall into three broad categories: (1) electricity-generator data (i.e., power plants), (2) meteorological data, and (3) geographic data.

Electricity generators Our data on electricity generators (at both the generator and plant levels) come from two sources: (*i*) the Emissions & Generation Resource Integrated Database (Emissions & Generation Resource Integrated Database, 2018) and (*ii*) the EPA's EmPOWER Air Data Challenge,[17] which

[15] E.g., health costs and diminished local amenities like visibility.

[16] For an example of diminished regulatory cost, consider the Clean Air Act's National Ambient Air Quality Standards (NAAQS). By reducing the area downwind of a polluter, there is (mechanically) less space to site an air quality monitor.

[17] More details can be found at the EmPOWER website: https://www.epa.gov/airmarkets/empower-air-data-challenge.

provides data through the EPA's Clean Air Markets Division (CAMD, 2020). Specifically, we use the eGRID data to obtain each EGU's latitude, longitude, year of construction, fuel category (e.g., coal, gas, hydro), generation capacity, and operating status. These variables are available at the level of generator and plant. We employ eGRID data from 2010, 2012, 2014, 2016, and 2018 (the intermediate years are unavailable). The EmPOWER CAMD data supply each EGU's daily emissions of NO_x and SO_2 and the EGUs' associated stacks' heights—both of which are inputs to the particle-trajectory model HYSPLIT. Both datasets include useful data on EGU retirements and fuel conversions. Panel B of Figure 1 illustrates the distribution of generators' capacities across four broad fuel categories for units operating in 2018.

Notably, the CAMD and eGRID datasets jointly allow us to construct the *historical* distribution of power plants in the United States. Because we observe both retirements and fuel conversions, the resulting dataset reflects the spatial distribution and fuel types of power plants *at their time of construction*—the information most relevant to our question of strategic siting.[18]

Meteorology Our meteorological data come from NOAA's North American Regional Reanalysis (NARR) daily reanalysis data (Mesinger et al., 2006; North American Regional Reanalysis, 2006). We use the NARR meteorology data in two applications. First, we utilize NARR's long-term averages (1979–2000) for wind speed and direction to determine prevailing, historical wind patterns in our analysis of strategic plant sitings. Specifically, we use NARR's first three pressure levels (the

[18]The repeated cross-sections of eGRID provide further confidence in constructing this historical distribution. Further, the 2010 version of eGRID precedes the vast majority of coal EGU conversions and retirements.

levels nearest to the ground): 1000 hPa, 975 hPa, and 950 hPa.[19] Second, we feed the NARR data into HYSPLIT for the particle-trajectory model's meteorology. In both applications, we employ NARR's highest spatial resolution with horizontal and vertical spacing of approximately 32 km (at the lowest latitude) (North American Regional Reanalysis, 2006).

Geography For state borders, county borders, coastlines, and bodies of water, we rely upon the U.S. Census Bureau's *TIGER/Line* shapefiles and cartographic boundaries (US Census Bureau, 2016a, 2016b). The bodies of water are subdivided into area files (i.e., polygons that enclose areas) and linear files (i.e., line-based hydrology). Finally, we integrate data on counties' non-attainment histories using the U.S. EPA's *NAYRO* file in its *Green Book* collection (U.S. Environmental Protection Agency, 2017). In this paper, we focus exclusively on EGUs in the contiguous U.S.—omitting Alaska, Hawaii, and U.S. territories.

2.4 Empirics

We now turn to our empirical analysis. Recall that the hypothesized strategic negative-externality export requires (1) that decision-makers site large polluters to reduce within-jurisdiction exposure and (2) that polluters' emissions are, in fact, sufficiently exportable.

We begin in 2.4.1 by documenting the fact that state and county decision-makers sited many EGUs very close to county and state borders. There are non-strategic reasons EGUs might locate near borders—namely, many borders are defined by water, a critical input for electricity production. Next, in 2.4.2, we formulate a simple test for regulatory avoidance that implicitly accounts for non-

[19] Pressure levels (barometric pressure levels) represent the force exerted from the weight of the air. Pressure levels decrease non-linearly with height.

strategic reasons for locating on an administrative border. In 2.4.3 we apply this test for strategic siting and discuss its results.

Finally, in 2.4.4 and 2.4.5, using a particle-trajectory model (HYSPLIT), we demonstrate that coal power plants' emissions are indeed highly transportable. Together, our results show that decision-makers have tended to site units of an exportable externality strategically. These results jointly satisfy sufficiency in demonstrating our hypothesized behavior: local decision-makers attempt to capture local benefits and export their negative externalities.

2.4.1 Power Plants' Distances to Borders and Water.

Border distance We start by calculating each plant's distance to the nearest county and state border.[20] Figure 3 illustrates the result of this calculation—the distribution of EGUs' distances to their nearest state and county borders. We separate the distributions by the EGUs' fuel categories, as EGUs' fuel types drive differences in other inputs.[21]

Figure 3 demonstrates that many EGUs were sited very close to county borders (Panel A) and state borders (Panel B). Further, this tendency is particularly extreme in coal-fired and hydropower EGUs—though natural gas plants also exhibit this trend. Of the 605 operating coal units in 2018 with capacities of at least 25 MW, 30% are within 1 km of a county border, 57% are within 5 km of a county border, and 77% are with 10 km of a county border. For state borders, the corresponding percentages are 18% ($\leq$ 1 km), 25% ($\leq$ 5 km), and

[20] While plants are divided into generating units (e.g., boilers), latitude and longitude are constant at the plant level in the eGRID dataset—i.e., all EGUs *within a plant* (ORIS code) are specified as having the same location in eGRID. See appendix section A.0.1.1 for the details of this calculation.

[21] E.g., coal units require access to coal—generally via rail or barge—while natural gas units typically require access to the natural-gas pipeline.

29% ($\leq$ 10 km). Only hydropower EGUs skew more toward administrative borders than coal-fired EGUs. We formally test whether EGUs' placements are independent of borders using a Kolmogorov-Smirnov test. This test compares EGUs' distances to borders against a null distribution of distance-to-nearest-border for a uniform grid covering the entire contiguous US. If EGU placements are independent of borders, these distributions should be similar. All fuel types strongly reject this independence except for solar/wind's distance to county borders (see Table A1). As Figure 3 and these statistics suggest, a substantial (and disproportionate) share of U.S. coal-fired electricity generators sit near county and state borders.

Non-strategic explanations for EGUs' proximity to borders One explanation for coal EGUs' proximity to county and state borders is the strategic export of coal generation's negative externalities. However, plants may site near borders for other reasons. Most methods of electricity generation require water for steam, cooling, locomotion, or transportation (solar and wind are exceptions). If large bodies of water (rivers or lakes) form many state/county borders, then water as an input *could* explain plants' proximity to borders.[22]

We calculate the share of each county's and state's borders that intersect bodies of water by spatially joining administrative borders (both state and county borders) to the boundaries of bodies of water (using a 50-meter buffer to allow for *near misses*).[23]

We find that approximately 46.1% of state borders and 27.4% of county borders coincide with bodies of water. States differ greatly in the shares of their

[22]This explanation also requires that the interiors of counties (and states) do not contain other large bodies of water. Otherwise, EGUs could just as easily locate in counties' interiors rather than on borders.

[23]Appendix section A.0.1.3 describes this operation in detail.

borders (county and state) formed by water. Figure 4 illustrates this heterogeneity, and Figure 5 provides four examples of the county and state borders identified as coinciding with water (in dark blue lines). As demonstrated by Figure 4, states in the non-coastal, western U.S. make up the lower end of the distribution with very few county or state borders coinciding with water—e.g., in Colorado, Wyoming, and New Mexico, less than 1% of state borders coincide with water, and 2%–3% of county borders coincide with water. Many coastal states (including the Gulf Coast and Great Lakes) have relatively high shares of borders coinciding with water. However, some interior states also have high water shares—e.g., 65% of Kentucky's state border and 41% of its county borders coincide with large bodies of water. Thus, most states—and many counties—offer potential sites with water and proximity to the border.

Panel A of Figure 1 confirms that EGUs locate near bodies of water (again, except wind and solar): 99% of hydropower units and 62% of coal units are within 250 meters of a body of water.[24] For natural-gas units, 48% are within 250 meters of water. For wind and solar EGUs, only 30% of generators are within 250 meters of a body of water. Given that hydro and coal units require large amounts of water—and wind/solar units do not—these results validate the spatial calculations in the rest of the paper and confirm that water is, indeed, a binding locational constraint when siting plants. However, these results do not entirely explain the phenomenon of siting coal plants near borders. Many bodies of water exist

[24]Measurement error in the latitude and longitude of generators and the Census water files likely explains why hydropower does not hit 100%.

in the interiors of counties/states, yet coal EGUs tend instead to locate near administrative borders.[25]

2.4.2 Strategic Plant Siting: A Statistical Test. We now develop a simple, non-parametric test to detect whether plants have been strategically sited near borders to reduce their home counties' (or states') exposure to the plant's pollution—rather than being placed near borders due to plant's demand for water.

With this motivation in mind, it is clear that proximity to certain borders is more advantageous than proximity to other borders. If a plant locates on the downwind border of its county, then its emissions immediately will leave its county (for example, the plants depicted in Figures 2a and 2b). If a plant locates near the upwind border of its county, then its emissions will pass through a substantial portion of its county (e.g., Figure 2d). Thus, all else equal, local decision-makers wishing to reduce their county's pollution exposure will prefer to reduce the area within the county that is *downwind* of the plant.[26]

Now consider the possibility—our null hypothesis—that decision-makers do not try to export their coal pollution. Under this null, decision-makers search for a location that maximizes the plant's profit, independent of the share of emissions exported. Consequently, plants' locations should be independent of the downwind vs. upwind exposure of their emissions: this ratio is not an input to production, nor is it an input to electricity production. In the absence of emissions exports, it should be a 50-50 'coin-flip' whether the area downwind of the plant is larger or

[25]For example: The *interior* Catawba County in North Carolina contains the Marshall Steam Station, a 2.1-gigawatt coal plant located on Lake Norman.

[26]The same reasoning applies at the state level.

smaller than the area upwind (within the jurisdiction containing the EGU).[27] In other words, in the absence of strategic emissions exports, there is nothing special about *downwind* water.

Therefore, a simple, non-parametric test for strategic emissions exports in the siting of coal-fired power plants is to calculate the number of coal plants for which the *downwind area* (in the county or state that contains the plant) is less than the *upwind area*. We operationalize this test as an implementation of Fisher's Exact Test (Conover, 1971; Fisher, 1934, 1935; Imbens & Rubin, 2015). Under a sharp (one-sided) null hypothesis of *no strategic siting to reduce downwind area*, the test statistic n_s (the number of plants for which downwind area is less than upwind area) is distributed as a binomial distribution with size equal to the number of plants in the sample (N_T) and probability $p = 0.5$. Under this null, the expected share of plants with downwind area less than upwind area is 50%. Consequently the p-value for the corresponding test statistic is

$$p\text{-Value}\big(n_s\big) = \mathbf{P}\big(X \geq n_s;\, n = N_T,\, p = 0.5\big) = \sum_{x=n_s}^{N_T} \binom{N_T}{x} 0.5^{N_T}.$$

Given that county and state decision-makers both potentially face incentives to reduce pollution exposure within their administrative units, we implement our test for strategic siting at both administrative levels.[28]

Our test offers several attractive features. First, the identifying assumption is that a decision-maker will only minimize a plant's downwind area to avoid the

[27] Using a uniform grid covering the contiguous U.S.—effectively a higher resolution version of the raster depicted in Figure A1—we confirmed that the probability a point in this grid is more upwind than downwind is, indeed, almost identical to 50 percent.

[28] It may be helpful to note that there is an upwind side and a downwind side for nearly every border in the U.S. (at least for borders that run orthogonal to the wind). Our test simply asks whether decision-makers disproportionately placed coal EGUs on the upwind side of the border (reducing their downwind areas).

costs associated with the plant's pollution. This assumption is plausible because coal- and natural-gas-fired electricity generators do not use the *areas* upwind or downwind—or their ratio—as inputs into their production or transport of electricity. Put differently, because EGUs do not use the *ratio* of downwind-to-upwind area for production or transport, strategic pollution export is the only real explanation for locating plants in a manner that reduces the county's (or state's) exposure downwind. If a latent factor correlates with the ratio of the downwind-to-upwind area at the state or county level, then our test will falsely conclude strategic siting. However, very few social, political, or physical processes consider the areas downwind or upwind of a point in space—let alone their ratio. Further supporting this assumption: when we analyze a fine, uniform grid covering the contiguous U.S., we find no evidence of a relationship between this ratio (or its inputs) and population density or population demographics.[29]

Second, this test is simple, straightforward, and provides an exact p-value that do not rely upon parametric or asymptotic assumptions (Imbens & Rubin, 2015).[30] Third, natural-gas EGUs provide a convenient falsification test for our approach. Natural-gas plants produce substantially less pollution than coal-fired EGUs, so counties and states do not face the same incentives to reduce gas EGUs' downwind pollution exposure. However, natural-gas plants face similar transmission constraints to coal plants. Consequently, if a latent factor is biasing our test toward

[29]Figures A3 and A4 illustrate that there is no relationship between share of county (or state) upwind (or downwind) and population density or population demographic composition. To falsify our identifying assumption, population density (or population composition) would need to bunch near downwind borders and avoid upwind borders. The figures contain no evidence that this bunching occurs.

[30]One drawback of the test's simplicity is that it does not incorporate other dimensions of strategy, e.g., stack heights. This omission does not bias the test for our specific hypothesis. It simply means we are testing for a specific strategy.

detecting "strategic siting," we should detect strategic siting for coal EGUs and natural-gas EGUs. In short, this simple procedure generates an intuitive test for strategic siting with exact p-values, a plausible identifying assumption, and a convenient falsification test.

In addition, our approach easily extends to test whether decision-makers located plants jointly reduce county and state downwind areas. Under the null of no strategic siting, the expected percentage of plants whose downwind area is less than the upwind area *at the county and state levels* is 25%.[31] More generally, this non-parametric test provides simple and clear evidence of whether decision-makers sited coal plants to reduce the downwind area in the plants' home counties and states.

To implement this test, we calculate the areas upwind and downwind of each coal and natural-gas plant in our data within the plants' counties and states. For the *wind* component of upwind and downwind areas, we use NARR's long-term averages of wind direction. The *area* is defined by the county's (or state's) intersection with right triangles emanating upwind or downwind of the plant. Figure 2 provides four examples of this calculation—illustrating the direction of the prevailing wind (the dark purple triangle in the compass), the *downwind area* (shaded dark purple), and the *upwind area* (shaded light gray). The plants in Figures 2a and 2b are located near borders in a manner that substantially reduced the downwind area in the plant's home county. The plants in Figures 2c and 2d were sited in parts of their county in which the downwind area is larger than the

[31] The null of no strategic siting implies that decision-makers' siting decisions are independent of the area downwind (or upwind) at the state and county levels. Under this null, the probability a plant is more downwind in the *state* is independent of the probability the plant is more downwind in the *county*. Thus, the probability of being 'downwind' at both levels is $0.25 = 0.5 \times 0.5$.

upwind area. Using these downwind and upwind areas, we implement our test for strategic siting.

2.4.3 Strategic Plant Siting: Results. Table 1 contains the results of our test for strategically sited coal and natural-gas plants. Given that coal EGUs produce substantial amounts of pollution, decision-makers have strong incentives to strategically locate coal plants to reduce the amount of jurisdictional area downwind of the plants. Natural-gas plants produce considerably lower emissions, giving decision-makers much less incentive to site natural-gas plants strategically. We separately test *coal plants* (column 1) and *natural-gas plants* (column 2). Table 1 contains three panels that respectively test strategic siting (A) within counties, (B) within states, and (C) within *both* counties *and* states. Each of the three panels bears strong evidence of strategic siting of coal plants that reduced the downwind areas within plants' counties (Panel A), states (Panel B), and both (jointly) counties and states (Panel C). There is no evidence that natural-gas plants were strategically located to reduce their downwind areas at any level.

Panel A tests strategic siting at the county level. Among the 514 coal plants, 56.81% sit where the area downwind of the plant (in its county) is less than the area upwind. Under the null of no strategic siting, with 514 plants, one would observe a distribution at least this extreme (in the right tail) approximately 0.12% of the time (i.e., a p-value of 0.0012).[32] For the the 1,254 natural-gas plants, the corresponding share of *strategically located* plants at the county level is 49.44% with p-value of 0.6641.[33] At the county level, our test finds large and statistically

[32] We do not expect this number to be near 100%, as governments and firms face many constraints when siting coal plants (e.g., water, rail, regulation, and local opposition to some sites)—in addition to likely having heterogeneous preferences.

[33] Recall that under the null, the expected share of strategically located plants at the county or state level is 50%.

significant evidence within the group most incentivized to strategically site (coal plants) and no evidence within the group with few incentives to do so (natural gas plants).

The results at the state level (**Panel B**) paint a very similar picture as the county-level results. There is statistically significant evidence of strategic siting among coal-fired power plants (53.89% strategic with a p-value of 0.0426) and no evidence of strategic siting within natural-gas plants (45.77% with a p-value of 0.9987).[34]

In **Panel C** of Table 1, we test whether plants are strategically located both within their counties *and* within their states. Under the null hypothesis of *no strategic siting at either level*, the expected share of strategically sited plants is 25%. Across the 514 coal plants, 34.82% sit in locations consistent with strategic siting at both county and state levels (p-value less than 0.0001). With an expected value under the null of 25%, thr level of strategic siting in this case (34.82%) is economically significant: an additional 50 coal plants (10%) sit in locations where they can export their pollution. As before, natural-gas plants show no evidence of strategic siting to reduce the area downwind of plants (25.04%; p-value of 0.4978). Again, we find highly significant evidence that decision-makers sited coal plants to reduce downwind exposure within the jurisdiction where the plant is located (counties and states).[35]

[34] If anything, natural-gas plants appear to be sited in an anti-strategic manner at the state level—i.e., where the downwind area typically exceeds the upwind area. One explanation for this behavior is that natural-gas plants may share bodies of water with coal plants, but gas plants are willing to 'take' the downwind side of the resource (the gas plants do not need the strategic location). An alternative explanation is that, when converting coal units to natural gas, decision-makers may prefer to replace less strategically located coal units with natural-gas units.

[35] These results are robust to dropping coastal counties; see Table A2.

Whether we consider counties, states, or both levels simultaneously, we find substantial evidence that decision-makers sited coal plants to reduce the areas downwind in plants' counties and states. We apply the same test in the case of strategic siting to natural-gas plants—a class of plants for which local administrators should have few incentives to site strategically. We fail to detect any significant evidence of strategic siting in this natural-gas-plant placebo test. Therefore, we conclude that Table 1 provides strong and statistically significant evidence that local decision-makers strategically placed coal-fired power plants to reduce the area downwind of plants within plants' counties and states.

This result of strategic siting is necessary for our hypothesis of local strategic export of pollution from coal-fired power plants, but it is not sufficient. In the next section, we substantiate the second part of our hypothesis by documenting the extent of coal pollution's mobility.

2.4.4 Pollution Mobility: Methods. To estimate the extent to which coal-fired EGUs' emissions travel beyond the counties and states that house the EGUs, we employ a state-of-the-art particle-trajectory model known as HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) (R. Draxler, Stunder, Rolph, Stein, & Taylor, 2020; R. R. Draxler & Hess, 1998). Developed by NOAA's Air Resources Laboratory, HYSPLIT is a heavily vetted and frequently used tool for calculating the trajectory and dispersion of chemicals through the atmosphere (Stein et al., 2015). Over its 30 years of development, researchers have used HYSPLIT to model the transport and dispersion of emissions from coal-fired EGUs (Henneman, Choirat, Ivey, Cummiskey, & Zigler, 2019; Henneman, Choirat, & Zigler, 2019; Henneman, Mickley, & Zigler, 2019), facility-level pollution (Grainger & Ruangmas, 2017; Hernandez-Cortes & Meng, 2020), smoke plumes

from forest fires (Stein, Rolph, Draxler, Stunder, & Ruminski, 2009), volcanic ash (Stunder, Heffter, & Draxler, 2007), mercury (Ryaboshapko et al., 2007), and methane emissions from the Marcellus Shale play (Ren et al., 2019).

HYSPLIT requires pre-generated, gridded meteorological data, for which we use the 32-km resolution NARR (North American Regional Reanalysis) data from NOAA (North American Regional Reanalysis, 2006). We then model particle trajectories for the NO_x and SO_2 emissions of every coal-fired EGU above 25 MW[36] in the contiguous U.S. every day during January 2005 and July 2005. As described in *Data*, unit-level emissions releases and stack heights come from CAMD (2020).[37] Modeling emissions for January and July allows us to depict the differences in emissions and meteorology between winter and summer. We model particles' paths for 48 hours after their release.

We illustrate the output of HYSPLIT in Figure 6. The algorithm calculates particle paths for hundreds of particles emanating from a specific EGU's three-dimensional location (longitude-latitude-height) at a given date-time of release.

2.4.5 Pollution Mobility: Results. From Figure 6 it is clear that the two plants' emissions leave their source counties within hours—and a large amount of the plants' emissions leave their source states within 24 hours of being released. This quick departure from the source counties and states occurs in both January and July. Figure 6 also highlights the fact that distance and direction for pollution transport may vary significantly by season (even for a single plant).

[36] Our threshold of 25 MW is a common cutoff in regulation—e.g., the Acid Rain Program, the Mercury and Air Toxics Standards (MATS), and the Cross State Air Pollution Rule (CSAPR) each focused on EGUs of 25 megawatts or greater.

[37] One shortcoming of this HYSPLIT-driven approach is that it does not model chemical reactions in the atmosphere (e.g., formation of $PM_{2.5}$ or ozone).

2.4.5.1 Transporting Pollution Away from Sources.

We now formalize and generalize these insights concerning the export and transport of coal emissions. For each coal plant, we calculate the share of that plant's emissions that travel beyond the plant's county and state (for each hour after the initial release). We separately calculate these plant-hour shares by administrative unit (county vs. state), month (January vs. July), and pollutant (NO_x vs. SO_2). For instance, in January 2005, 32.9% of NO_x emissions from coal plant "3470" (depicted in Figure 6c) left the plant's *county* within one hour after the initial release. However, none of these NO_x emissions left the plant's *state* (Texas) within one hour after release. Four hours after release (still for plant 3470 in January 2005): 94.6% of NO_x emissions were outside the plant's county, and 11.0% were outside the plant's state. As Figure 2d illustrates, plant 3470 is located upwind of much of its county and even more of its state (Texas), so it is reasonable that it would take time for its emissions to leave both jurisdictions. For plants that have been more strategically located—e.g., plant 1378 in Figure 2b was ideally sited to reduce in-county emissions—most of the emissions immediately leave the county: by one hour after release, 69.5% of its emissions had already left the county.

Figure 7 displays our pollution-mobility results for all coal plants operating in 2005. The four subplots separate the results by administrative level (top panel (A): *county*; bottom panel (B): state) and pollutant (left pane: SO_2; right pane: NO_x). The x-axis shows the number of hours that have passed since the initial emissions release; the y-axis gives the share of particles that have left the source's administrative unit. The thin lines in each figure depict individual coal plants' monthly averages (black for January; light red for July). The heavy lines with dots

provide the average across all plants for each hour, weighted by each plant's mass of emissions.

The implications of Panel A of Figure 7 are clear. For most coal plants in the U.S., nearly all of the plants' pollution leaves these plants' home counties within six hours of the release. This fact holds in both seasons, but the departure is even faster in winter months (since these months have, on average, stronger winds). Panel B paints a similar picture for the timing of emissions' departures from source states: within 12 hours of release, 50%–85% of emissions leave the *state* of origin—and for many plants, this number is closer to 90% (again, particularly in the winter). Figure 7 demonstrates that pollution transport—a result of the geography of plant sitings, stack heights, and local meteorology—-creates a substantial wedge between the sources that export coal-based emissions and the downwind counties/states that receive the emissions.

2.4.5.2 Decomposing the Sources of Local Pollution. For a complementary perspective, we use HYSPLIT to decompose the sources of local, coal-based pollution. We separate the total *coal-EGU-generated pollution* within a county by the sources of the pollution. Specifically, we classify emissions sources by (1) whether the sources are in the same county, (2) whether the sources are in the same state, and (3) whether the sources' counties are in attainment (compliance) with national ambient air quality standards (NAAQS).[38] In 2005, 485 counties were out of attainment with the NAAQS (i.e., non-attainment) for at least one of the six criteria pollutants.[39]

[38]Note that we first sum all coal-generated emissions that HYSPLIT locates within a county. This sum ignores where the emissions originated—so long as HYSPLIT places the emissions in the given county. We then decompose this sum by the emissions' sources.

[39]Our HYSPLIT analysis focuses on 2005, so we only consider counties' 2005 attainment status. Counts of violations by standard: 8-hour O_3 (1997), 422; $PM_{2.5}$(1997), 208; PM_{10} (1987), 49; CO

Figure 8 illustrates the results of a source-based decomposition with pollution sources separated into five groups: (1) the county's own emissions, (2) *attainment* counties within the same state, (3) *non-attainment* counties within the same state, (4) *attainment* counties in a different state, and (5) *non-attainment* counties in a different state. Panel A shows the results of the decomposition for SO_2 emissions; Panel B for NO_x. In each panel, we repeat the exercise by 'receptor' county attainment status (attainment on left; non-attainment on right) and season.

Given our previous finding that nearly all emissions leave their origin county within six hours, it is unsurprising that a tiny share of a county's coal-EGU-based emissions comes from the county's own EGUs.[40] However, it is rather remarkable just how small the share of own-county emissions are relative to the contributions of other sources: the own-county shares (in black in Figure 8) range from 1% to 8%. While still small, it is notable that the share of own-county emissions is much larger for non-attainment counties than for attainment counties. This finding is consistent with coal plants' emissions (or existence) contributing to non-attainment designations. However, the vast majority of coal-EGU-based emissions in non-attainment counties appears to originate in other counties and states.

Across all counties, regardless of attainment state, the vast majority of emissions originate in other states—i.e., 65% to 85% (the sum of the yellow and orange segments in Figure 8). While this result may at first seem mechanical—each county only has one *own state* and 49 *other states*—it requires substantial transmission of *other states'* emissions. Without sizable cross-boundary

(1971) 100; SO_2(1971), 10; lead (1978), 2. A county can violate multiple standards (i.e., there were 702 violations in 485 counties).

[40] This result is also driven by the fact that many counties do not have any coal EGUs of their own.

transmission, counties and states would pollute themselves and not others. This result reiterates the importance of long-distance transport of coal pollution.[41]

Coal pollution is indeed highly exportable—even at the scale of states (and beyond). Along with our previous result of strategic downwind siting, this result closes the loop on our hypothesis of strategic export of negative externalities.

2.5 Discussion and Conclusion

In this paper, we empirically investigate the hypothesis that decision-makers have historically sited a major class of polluters—coal-fired power plants—to strategically export their negative externalities (pollution) downwind. After documenting coal EGUs' tendency to locate near borders, we formally test whether coal EGUs have been disproportionately sited nearer to the downwind borders of the counties and states. Our test finds large and significant evidence that decision-makers have located coal EGUs to reduce the area downwind of the plants within the counties and states that contain these plants. Our placebo test—using natural gas EGUs—does not exhibit this behavior.

Showing that local decision-makers have disproportionately located coal EGUs downwind within counties and states is a necessary condition for our strategic export hypothesis. For sufficiency, we must also show that coal-based pollution is exportable. Toward this goal, we use a particle-trajectory model (HYSPLIT) that illustrates the extreme mobility of coal-based emissions. Our results suggest that nearly all coal EGUs' pollution leaves the source counties

[41] Also potentially of interest in Figure 8: the difference between the emissions sources for attainment and non-attainment counties (the left and right halves of the figure). In *non-attainment* counties, the plurality (41%–50%) of coal-based emissions originates in *non-attainment* counties in other states. For *attainment* counties, a larger share comes from *attainment* counties in other states.

within six hours of the release. Within 12 hours of release, 50%–85% of emissions leave the *state* of origin—and for many plants, it is closer to 90%.

Jointly, these two pieces of evidence demonstrate that many local decision-makers historically located coal EGUs to enjoy their local benefits without facing their costs.

While these results focus on historical siting decisions, they have important implications for current environmental policy. In contemporary federalist regulations—e.g., the Clean Air Act and the Cross-State Air Pollution Rule (CSAPR)—dealing with cross-boundary pollution requires more coordination and resources than pollution that remains in (and mainly affects) its source county and state.[42] Strategically sited polluters emitting highly transportable pollution from tall smokestacks[43] create a complex and challenging regulatory situation.

The shapes of some non-attainment areas—areas deemed out of compliance with the Clean Air Act air-quality standards—reflect this complexity. Some non-attainment areas knit together whole counties with adjacent pieces of other counties and "islands" surrounding major point sources (often coal plants). For example, Figure 9 shows the Huntington-Ashland non-attainment area (which violated the 1997 $PM_{2.5}$ standard) in light orange. The Huntington-Ashland non-attainment area—a single non-attainment area—covers nine counties (5 whole counties; 4 partial counties) across three states (Kentucky, Ohio, and West Virginia). Six of the counties form a contiguous area. The remaining three counties (two in OH; one in WV) are islands—where each island circumscribes multiple coal plants (circled,

[42] To a degree, these challenges in federalist regulation of local pollutants mirror the international community's coordination failures for limiting greenhouse gases.

[43] In 2018, the average height of a smokestack attached to a coal-fired EGU in the U.S. was approximately 500 feet, and the maximum height was 1,038 feet (calculated from CAMD (2020) data).

red dots). This complex non-attainment area required substantial coordination across counties and states, source-attribution modeling, and federal oversight.[44] The Huntington-Ashland non-attainment area offers a single example of the complexities that can result from federalist environmental policy.

Figure 10 depicts a related challenge created by cross-boundary coal-based emissions (here, NO_x). In 2005, Shelby County (Tennessee) was designated as a non-attainment due to its violation of the 8-hour Ozone standard of NAAQS.[45] Panel A of Figure 10 shows all the coal-plant-generated NO_x emissions that eventually arrived in Shelby County during July 2005 (as estimated by HYSPLIT). We draw emissions' paths to Shelby County in grey; non-attainment counties (in 2005) are cross-hashed in red. We outline Shelby County in bright yellow. The figure illustrates that Shelby County's emissions originate throughout a broad geographic swath, stretching from Texas to Kansas to Indiana to Georgia, including attainment and non-attainment counties.[46] Overall, Panel A emphasizes the fact that large regions of the country affect one locality's air quality—a challenge for a federalist system with many small units.

Panel B of Figure 10 zooms in on the region surrounding Shelby County, Tennessee (the "zoomed" area is approximately 900 km east-west and 600 km north-south). Counties' fill color in Panel B matches their contribution (as a share) to Shelby County's coal-generated NO_xin July 2005. Panel C provides both the legend for the colors and the histogram for the distribution of counties' shares of

[44]Figure A5 provides an example of another "complex" non-attainment area contained within a single state (the non-attainment area containing Evansville, Indiana).

[45]NO_x, which we consider in Figure 10, is a precursor of both Ozone and $PM_{2.5}$.

[46]Notably, the emissions that eventually make their way to Shelby County come from a wide range of directions—emphasizing the importance of the temporal variation in prevailing wind directions embedded in HYSPLIT.

contribution to Shelby County's NO_x. Remarkably, although Shelby County had an operating coal plant in 2005 (and was out of attainment), the coal plant in Humphreys County, TN (a county which was *in attainment* in 2005) contributed more to Shelby County's NO_x than did Shelby County's own plant. Further, the coal plant in Independence County, Arkansas (also an *in-attainment* county in 2005) contributed approximately the same amount of NO_x emissions to Shelby County as did Shelby County's coal plant.[47] As illustrated in Figure 10, the vast majority of coal-based NO_x emissions in Shelby County, Tennessee—a non-attainment county—came from other states, and a majority of Shelby County's emissions originated from sources in attainment counties.

These two anecdotes highlight the challenges facing regulation and coordination within federalist systems. The results in our empirical section confirm that these cases and their challenges are not rare exceptions: Facing the spatial patchwork of jurisdictions created by the federalist structure of the U.S., local decision-makers have strategically sited polluters to export pollution. More broadly, our results point to the potential for local governments' actions to erode the efficiency of federalist systems—and potentially suggest a more prominent role for the federal government when externalities are exportable.

[47] Humphreys County, TN is home to the TVA's Johnsonville Fossil Plant, a 1.5-gigawatt coal power plant. Independence County, AR, houses Entergy Arkansas's 1.7-gigawatt "Independence" coal plant.

2.6 Figures

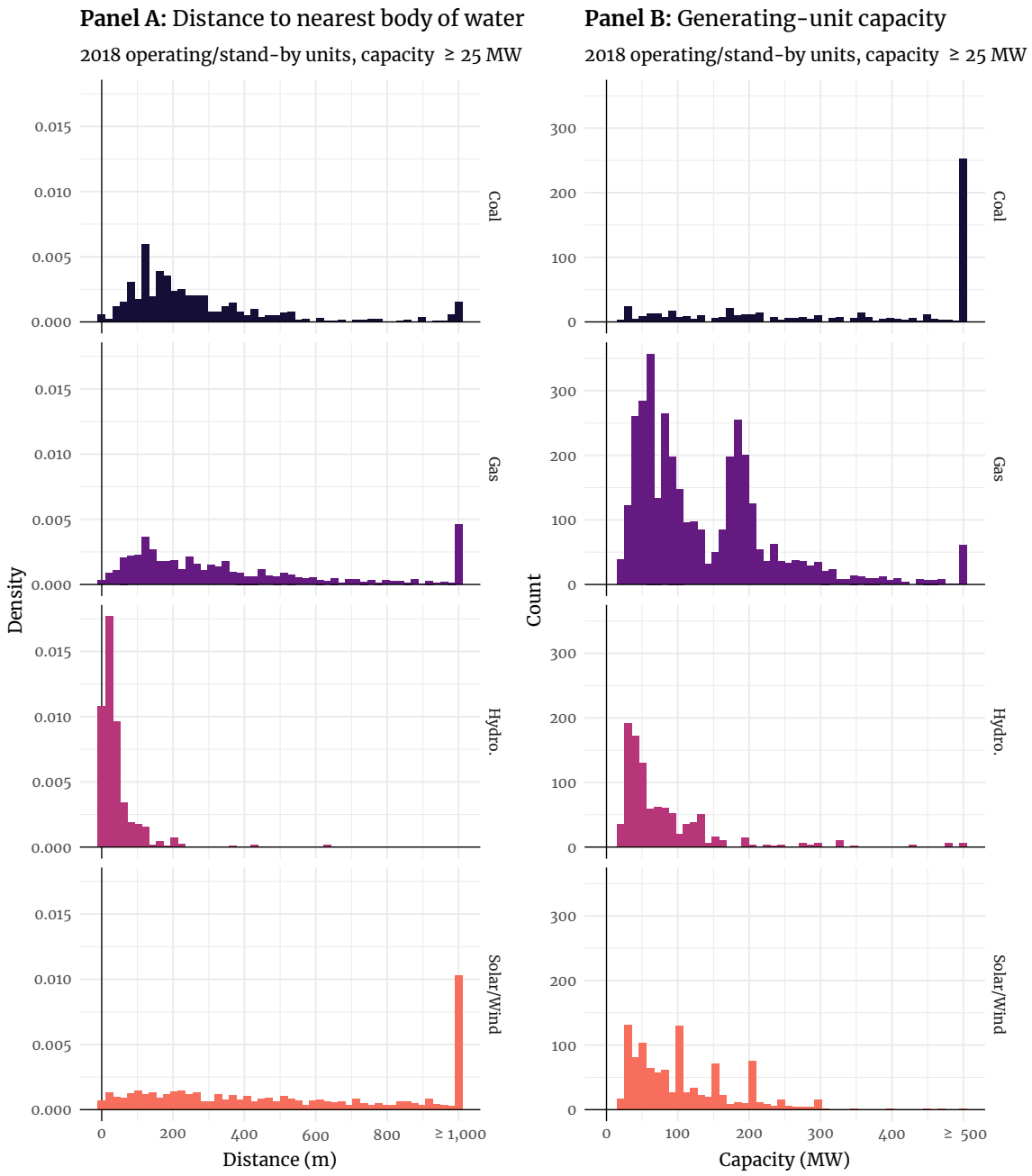


Figure 1. The distributions of EGUs' distances to their nearest body of water (**Panel A**, left) and EGUs' generation capacities (**Panel B**, right) by fuel category (row and color).

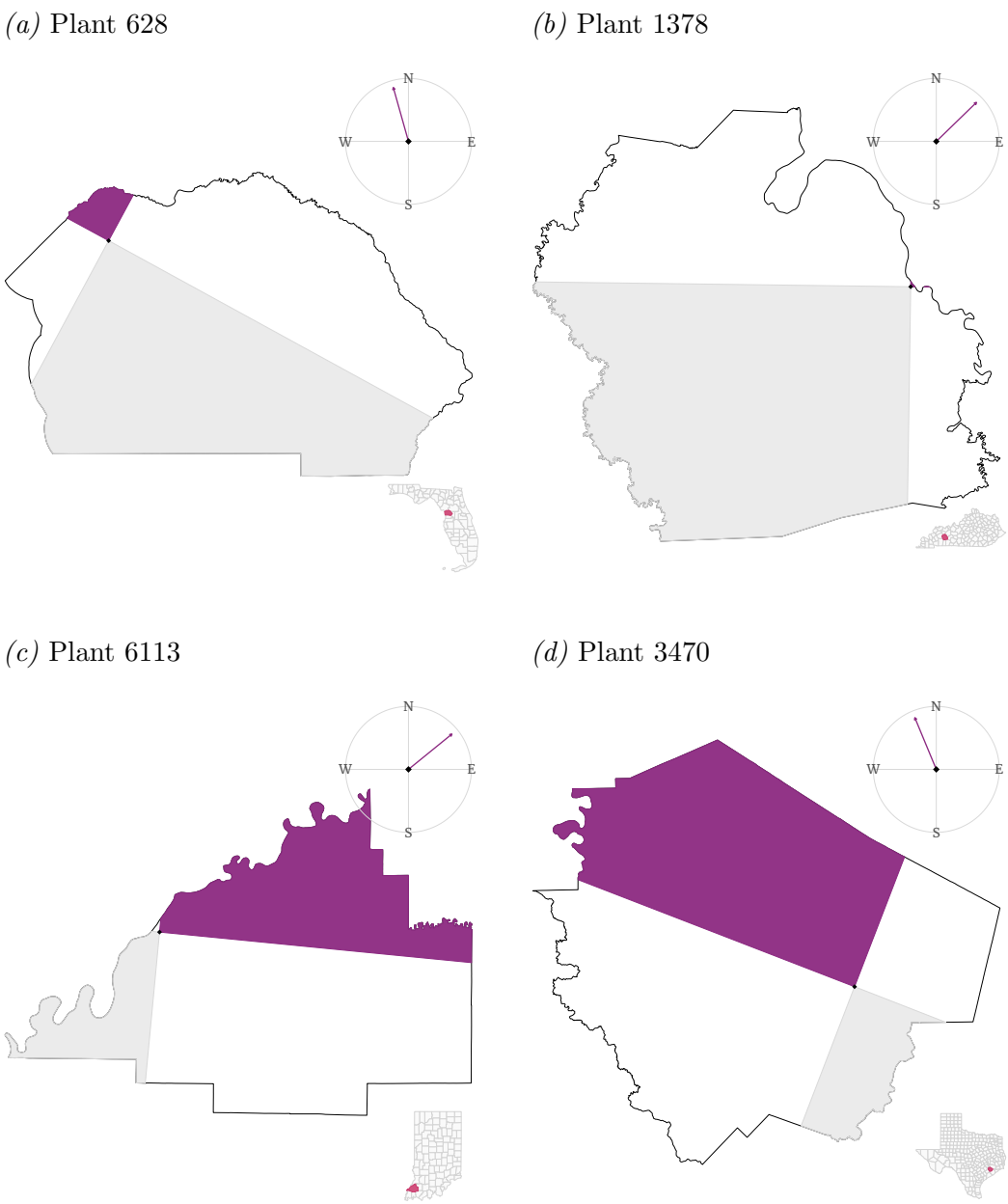


Figure 2. Upwind and downwind areas for four coal-fired generators. Dark, purple areas denote the 90-degree *downwind* area from the plant's location (the small, black diamond). Light gray refers to *upwind* areas. The outlined shape depicts the plant's county; the inset thumbnail highlights the plant within its state. The purple arrow within the compass points in the direction of the plant's prevailing wind direction (North American Regional Reanalysis, 2006).

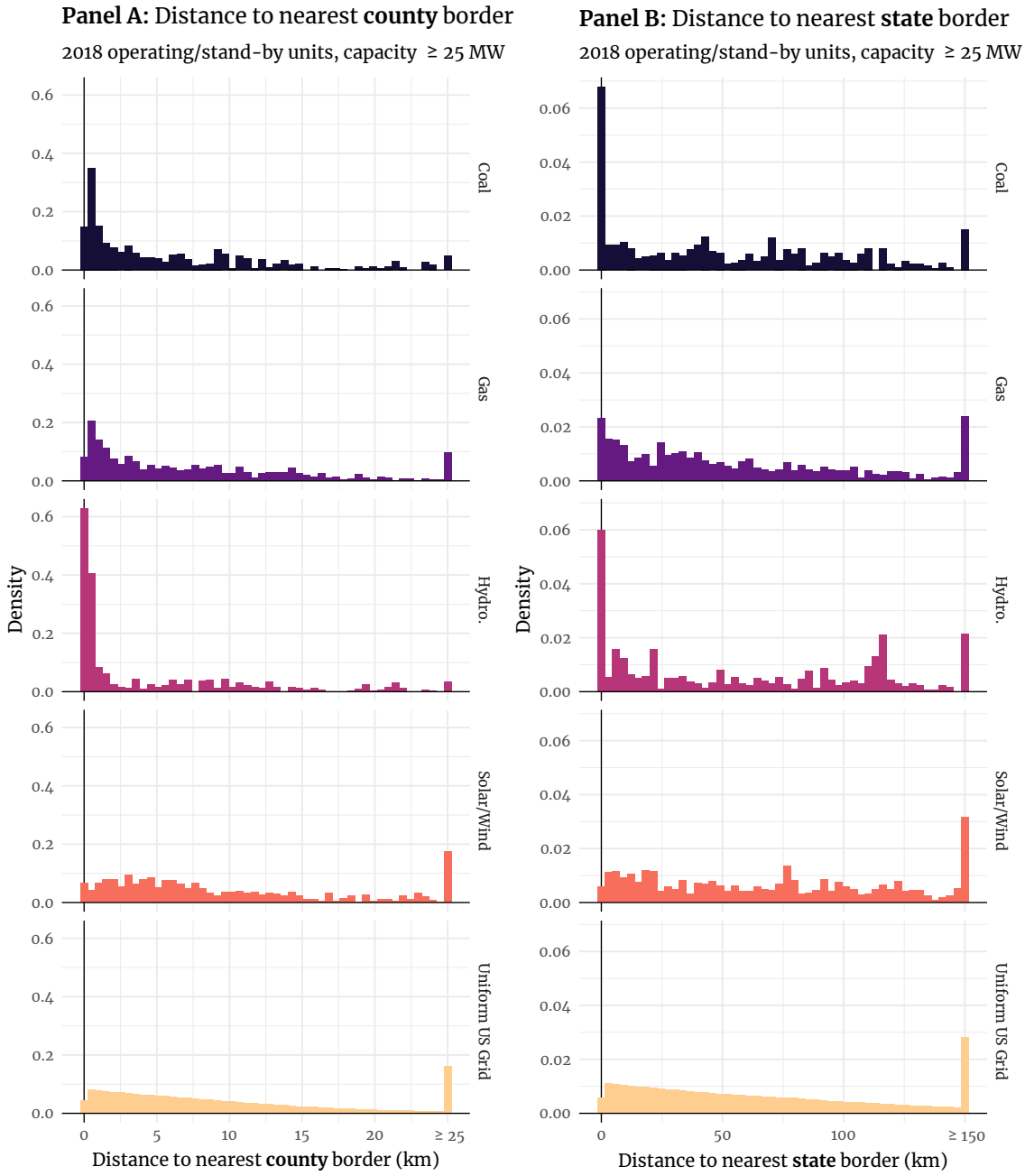


Figure 3. Empirical densities of the distributions of EGUs' distances to their nearest county (**Panel A**, left) or state (**Panel B**, right) border. The sample includes all operating and stand-by EGUs with capacities ≥25 MW within the contiguous U.S. in 2018. The first five rows of colored charts above separately produce the densities by fuel category. The final row reveals the density of *distance to the nearest border* from a uniform grid of points covering the contiguous U.S.

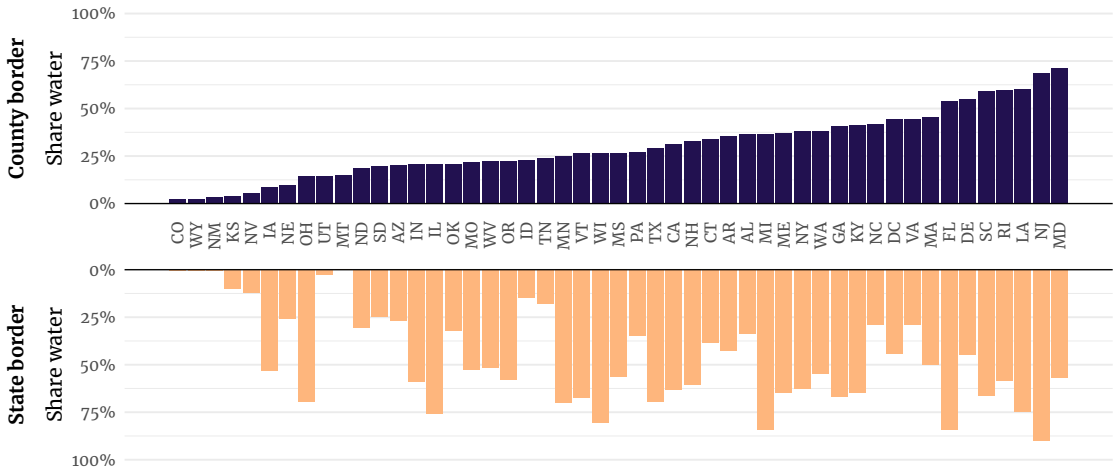


Figure 4. The share of county borders (top) and state borders (bottom) that coincide with bodies of water, by state. The states are sorted from smallest share of county-borders coinciding with water (Colorado) to largest share (Maryland). Alaska and Hawaii are excluded. Figure 5 provides four example states (LA, OR, SC, and SD) from these calculations.

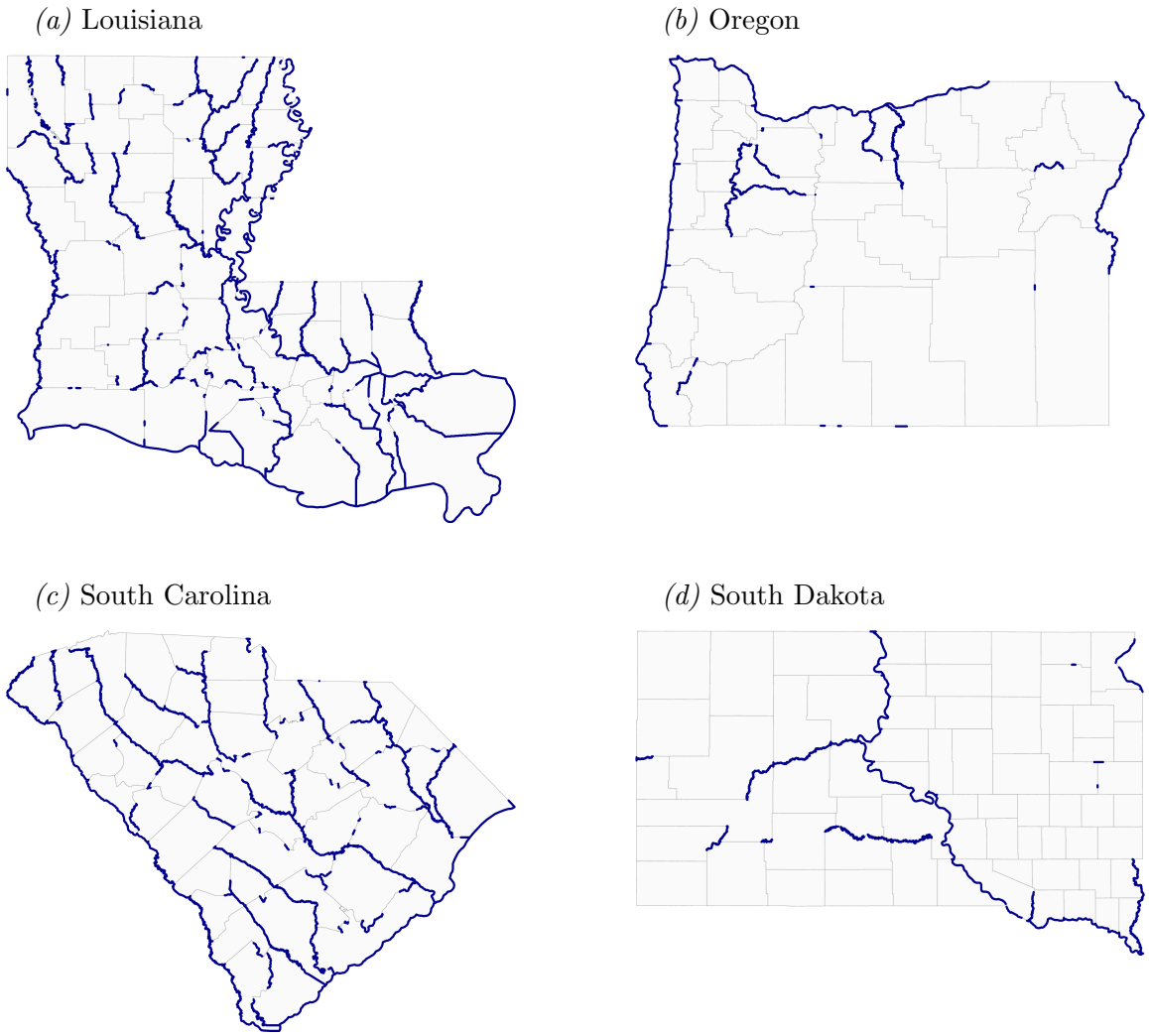


Figure 5. Examples of the output of our calculations of county and state borders that coincide with bodies of water. Darker blue lines denote administrative borders (state and/or county) that coincide with water; paler gray lines depict administrative borders that do not. Overall, our algorithm for detecting where jurisdictional borders coincide with water appears to be successful.

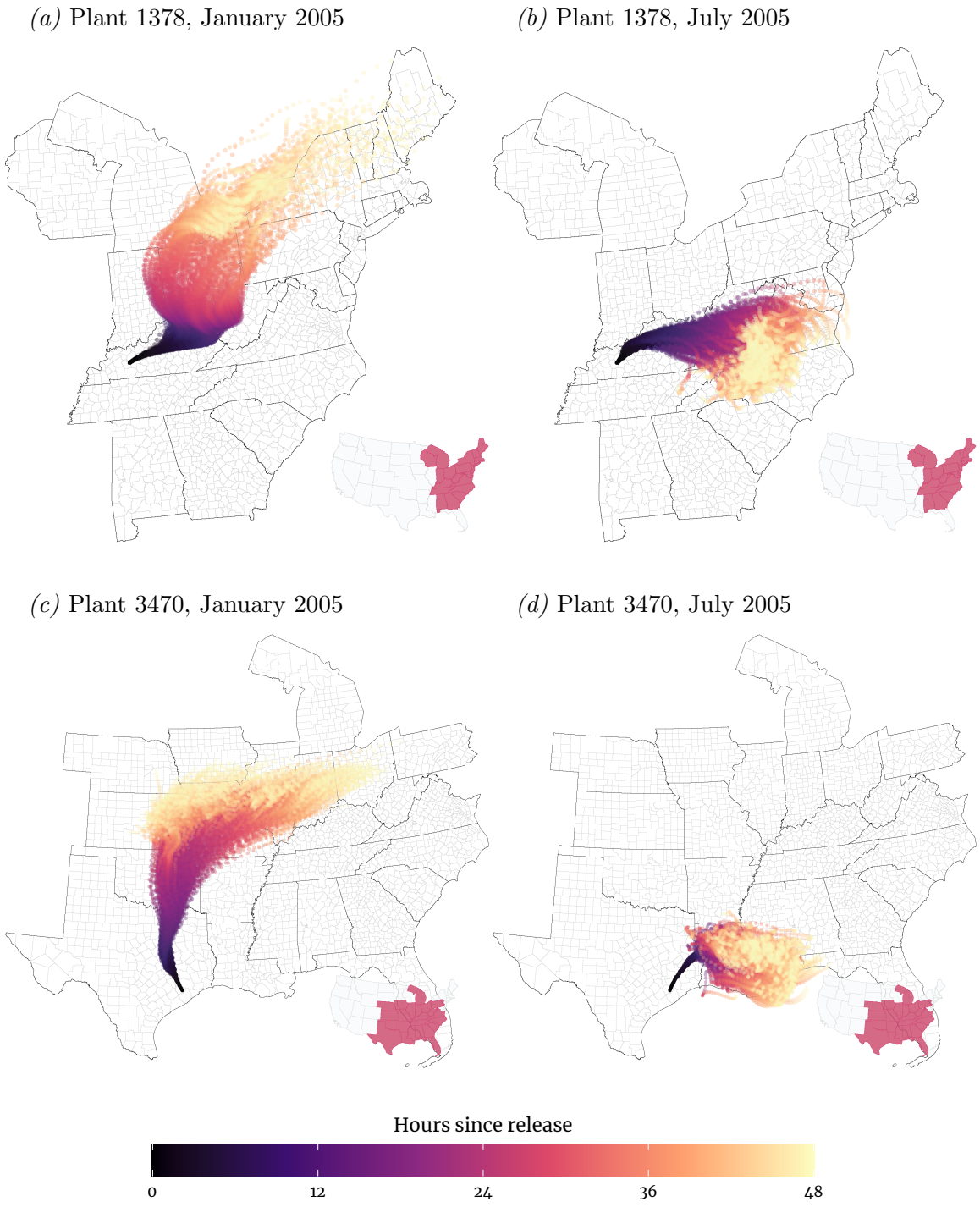


Figure 6. Particle trajectory and dispersion in HYSPLIT for two plants (ORIS codes 1378 and 3470) during January 2005 and July 2005. For each day of the month, HYSPLIT models 420 particles starting at the latitude, longitude, and altitude of the plants' chimneys. We track particles for 48 hours after their initial release; particles' colors denote the number of hours since their emission. The plants correspond to Figures 2b and 2d.

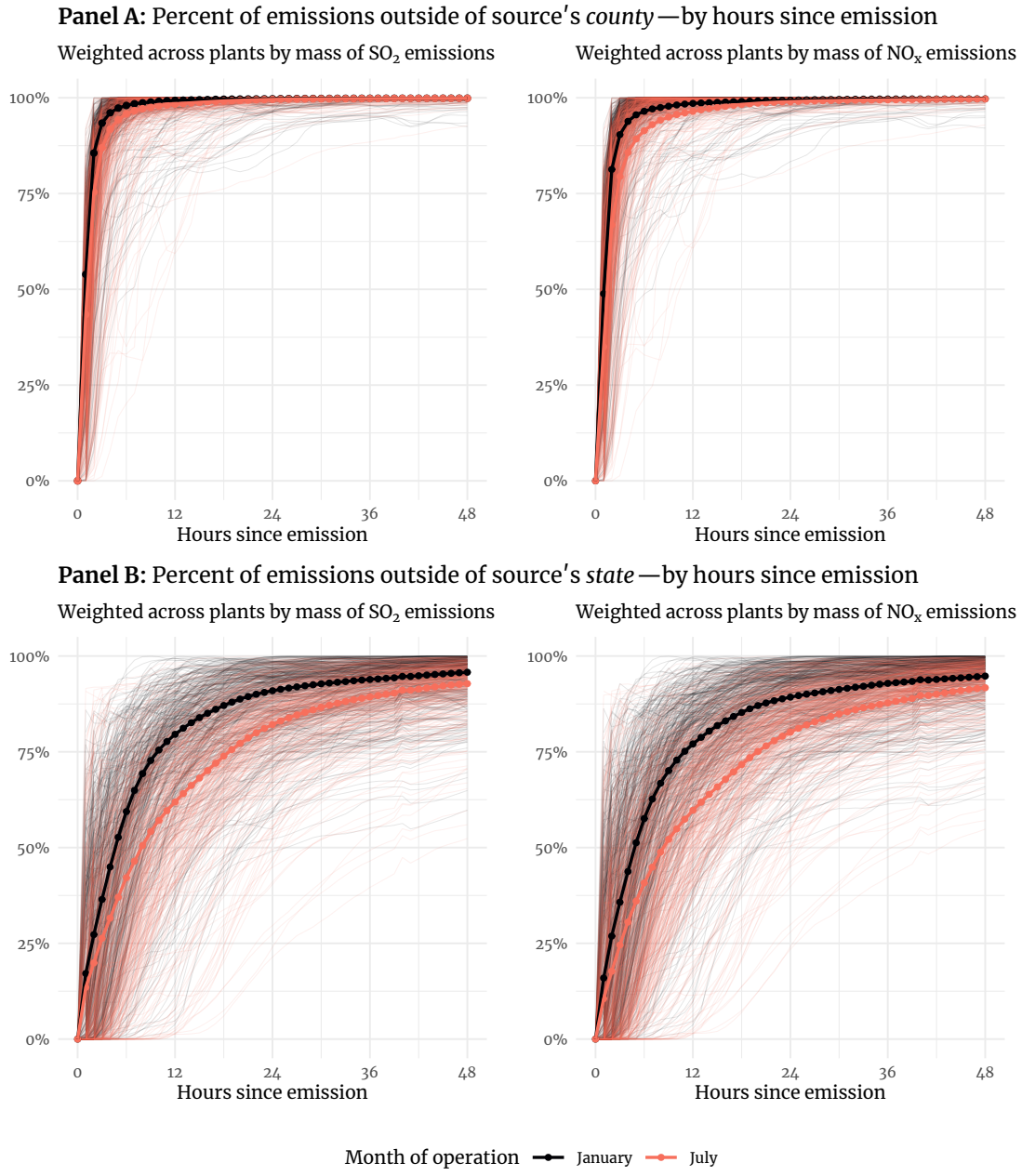


Figure 7. The share of coal plants' emissions that have left plants' origin counties (top, **Panel A**) or origin states (bottom, **Panel B**) by the number of hours that have passed since the particles were released (as modeled by HYSPLIT). Each of the four subfigures contains two months of emissions: January 2005 (black) and July 2005 (light, red). Thin lines depict individual plants in a given month. Thick lines (decorated with hourly points) denote the monthly average across plants (weighted by mass of emissions). The left column weights by SO_2; the right column by NO_x. Differences between the months capture seasonal differences in meteorology and in the distribution of generation. Sample: coal-fired generators $\geq$ 25 MW operating in Jan./July 2005.

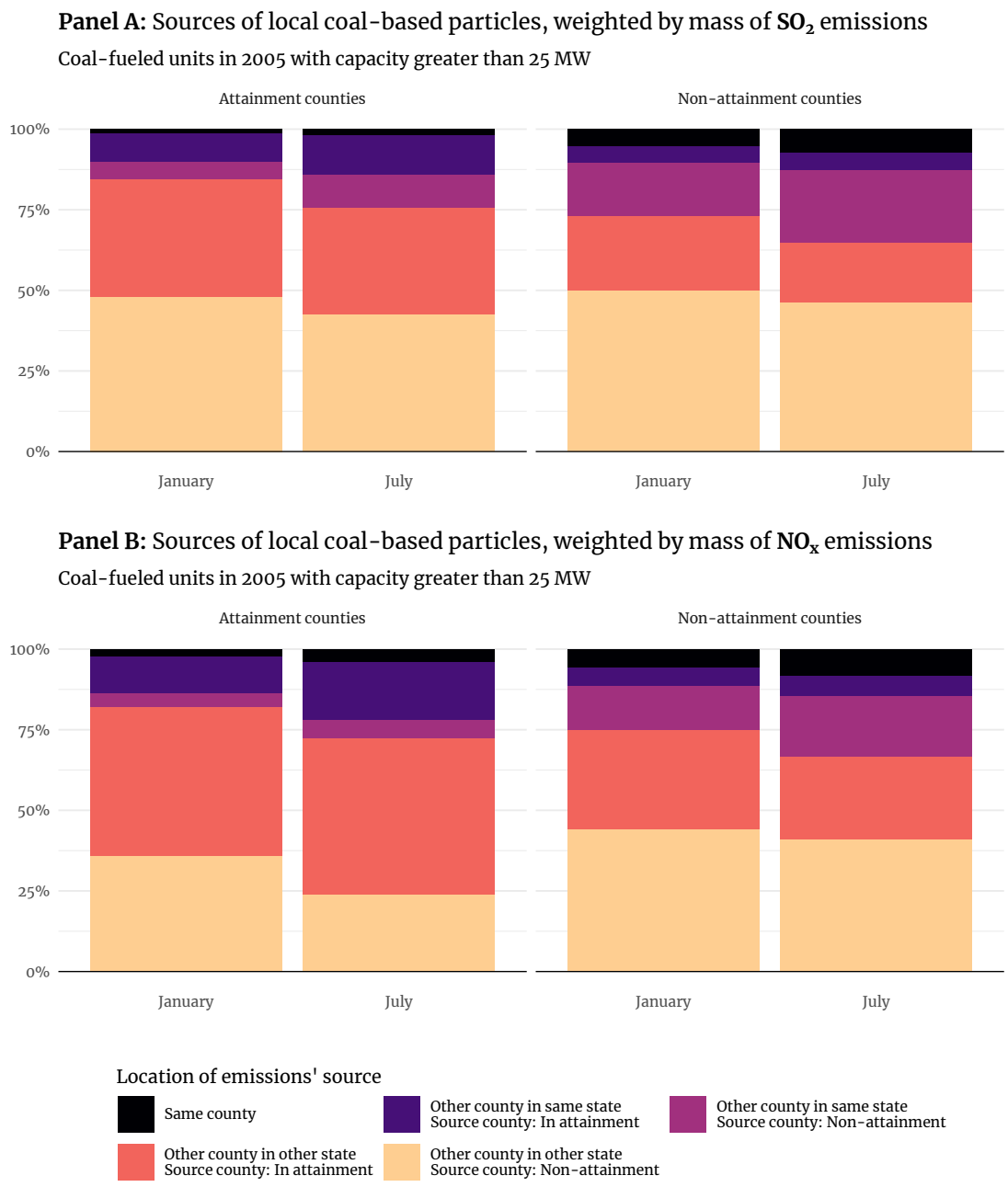


Figure 8. The Source-based decomposition of location, coal-based pollution. They described, on average, where a county's pollution *come from* based upon (1) the month (Jan. or July 2005), (2) the county's *attainment status*, and (3) the type of particle (SO_2 or NO_x). Particle trajectories come from HYSPLIT. The five colors refer to five categories of pollution sources by the EGU source's location (described in the legend). **Panel A** focuses on SO_2 emissions; **Panel B** on NO_x. Sample: coal-fired generators $\geq$ 25 MW operating in Jan./July 2005.

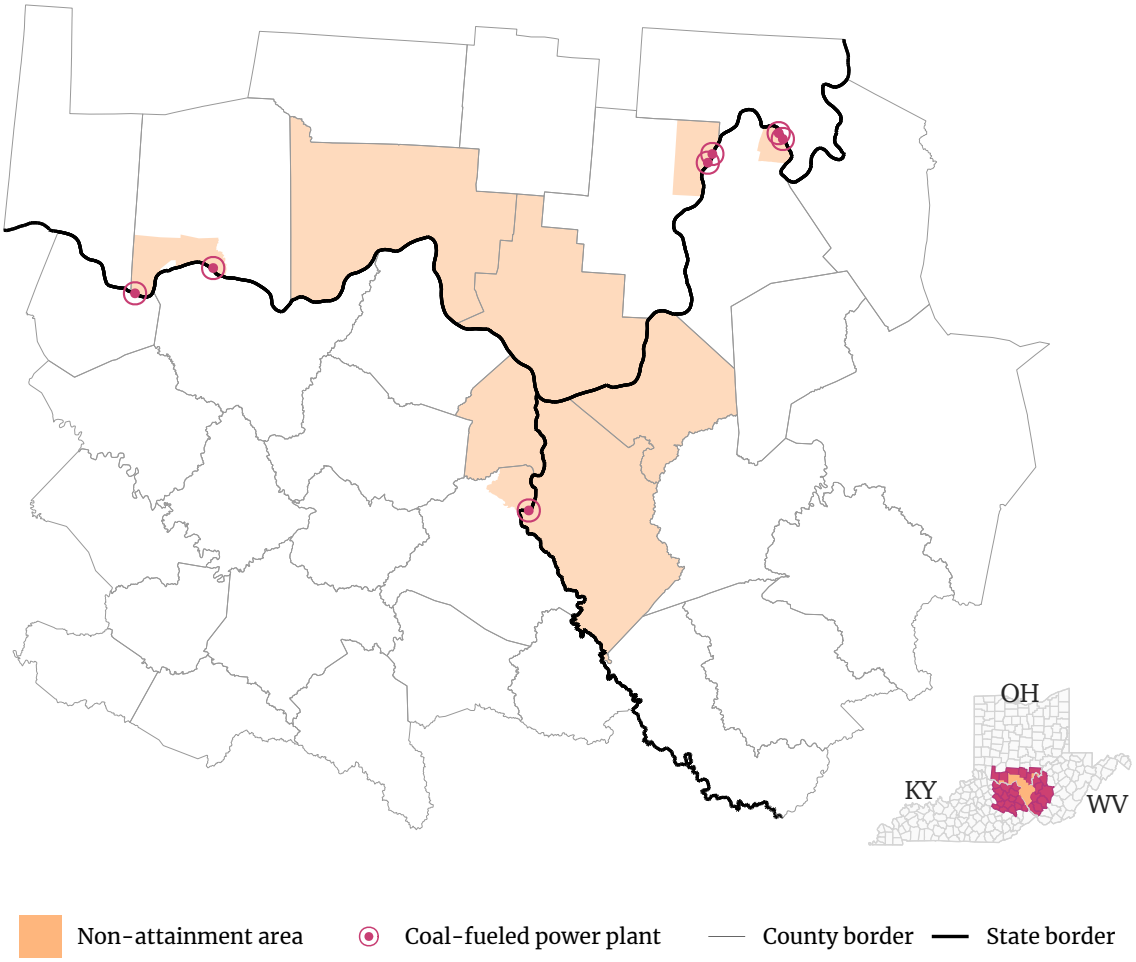


Figure 9. This map illustrates the complexity of the Huntington-Ashland non-attainment area (orange), which covers nine counties (5 whole; 4 partial) across three states. Six of the counties form a contiguous area. The remainder of the non-attainment area is comprised of "islands" that cover six coal plants (red-circled dots) in three different counties (two in OH; one in WV). This non-attainment area is for the 1997 $PM_{2.5}$ standard. Figure A5 depicts another example of a "complex" non-attainment area (Evansville, Indiana).

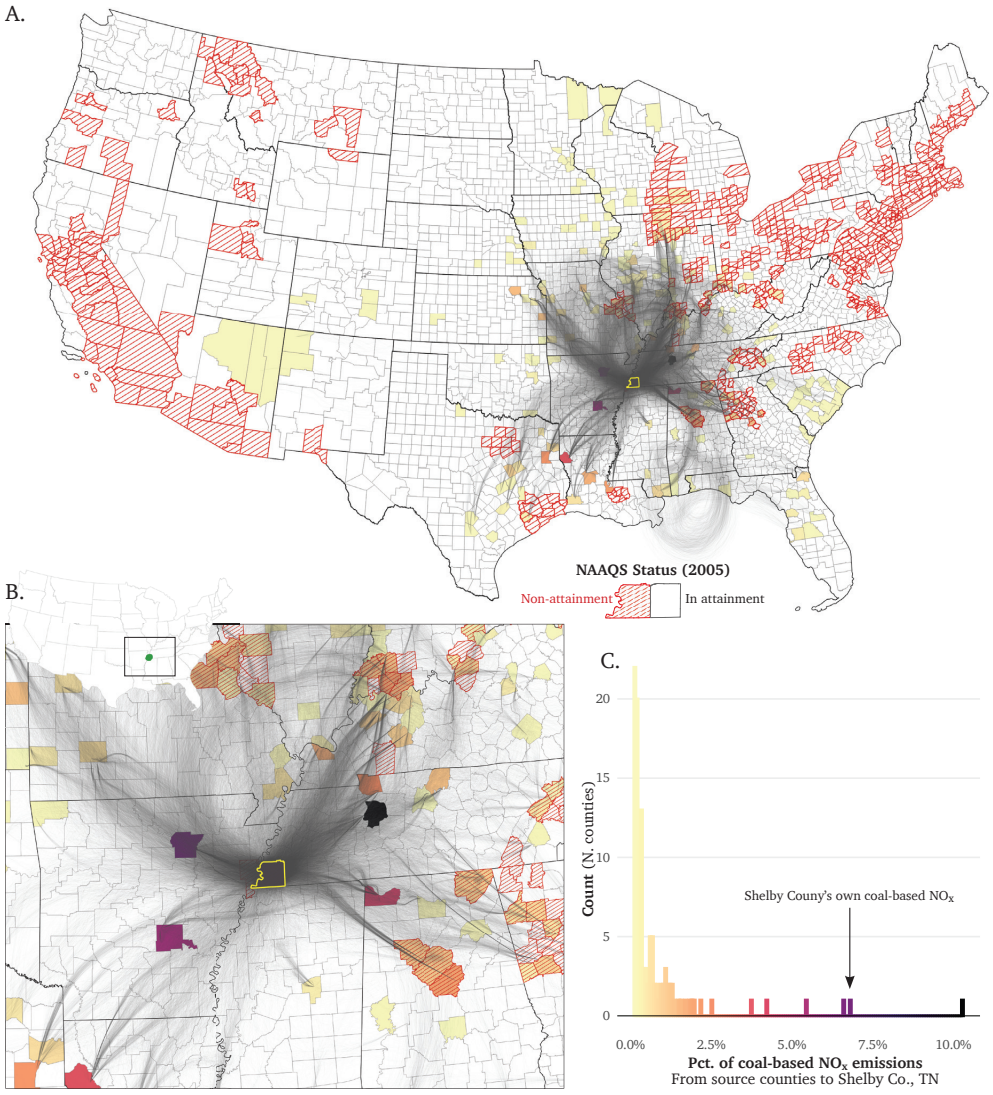


Figure 10. This figure shows the origins, paths, and shares of all coal-plant-based NO_x emissions that eventually enter Shelby County, TN during July 2005 (modeled by HYSPLIT). In 2005 Shelby County, TN was in violation of the 8-hour Ozone NAAQS (NO_x is an Ozone precursor). Subfigure **A**'s grey coal-based NO_x trajectories reveal that the sources of coal-based NO_x emissions in Shelby County include many states (from TX to GA to IL) both in attainment and non-attainment counties. Non-attainment (for any NAAQS) are hashed in red. **B** zooms in on the region surrounding Shelby County (∼900 km × 600 km). Counties are colored (filled) by the share of coal-based NO_x emissions that they contribute to Shelby County, TN. **C** provides the legend for **B**'s colored shares and plots the distribution of these shares—the x axis is the share of Shelby County's coal-generated NO_x emissions that each county contributes.

2.7 Tables

	Coal-fueled plants			Natural-gas-fueled plants		
	(1) **All**	(2) Post-CAA	(3) Pre-CAA	(4) **All**	(5) Post-CAA	(6) Pre-CAA
Panel a: Siting strategically within **county**						
Count	515	286	229	1,258	995	263
Count *strategic*	297	165	132	612	482	130
Percent *strategic*	57.67%	57.69%	57.64%	48.65%	48.44%	49.43%
Fisher's exact test of H_o: In-**county** downwind area $\geq$ upwind area						
Under H_o: E[Percent strategic: County] = 50%						
P-value	0.0003	0.0054	0.0122	0.8381	0.8448	0.5974
Panel b: Siting strategically within **state**						
Count	515	286	229	1,258	995	263
Count *strategic*	279	152	127	575	466	109
Percent *strategic*	54.17%	53.15%	55.46%	45.71%	46.83%	41.44%
Fisher's exact test of H_o: In-**county** downwind area $\geq$ upwind area						
Under H_o: E[Percent strategic: State] = 50%						
P-value	0.0321	0.1574	0.0563	0.9989	0.9788	0.9978
Panel c: Siting strategically within **both county *and* state**						
Count	515	286	229	1,258	995	263
Count *strategic*	182	98	84	310	249	61
Percent *strategic*	35.34%	34.27%	36.68%	24.64%	25.03%	23.19%
Fisher's exact test of H_o: Downwind area $\geq$ upwind area in **county and state**						
Under H_o: E[Percent strategic: County $\wedge$ State] = 25%						
P-value	<0.0001	0.0003	0.0001	0.6258	0.5049	0.7710

Table 1. Testing strategic location: Comparing up- and down-wind areas for coal and natural gas plants—before and after the Clean Air Act (CAA) of 1963. We define a plant's location as "strategic" if the downwind area *within its home county (or state)* is less than its upwind area *within its home county (or state)*. We calculate *downwind* and *upwind* areas based upon 90-degree right triangles with a vertex at the plant pointing up- or down-wind based upon the locally prevailing wind direction. Figure 2 illustrates this calculation. The columns that reference *Post-/Pre-CAA* refer to whether the plant's first year of operation was after or before the Clean Air Act of 1963. *Sources:* Emissions & Generation Resource Integrated Database (2018) and authors' calculations.

CHAPTER III

THE ENVIRONMENTAL COST OF LAND USE RESTRICTIONS

This chapter is co-authored with Mark Colas. It is published as: "The Environmental Cost of Land Use Restrictions." Quantitative Economics, 2022. 13(1): 179-223. I had a crucial role in developing the project, the statistical estimation, and writing the paper. Additionally, I formulated many of the (ultimately successful) responses to reviewers that resulted in the paper being accepted.

3.1 Introduction

Higher levels of carbon dioxide (CO_2, hereafter "carbon") are associated with a multitude of global environmental issues. The amount of energy a household uses, and therefore the amount of carbon emissions a household is responsible for, depends partially on where the household lives.[1] For example, Oklahoma City has high summer temperatures and relies heavily on coal-fired power plants while San Francisco has a moderate climate and uses electricity produced largely by hydroelectric plants. Households in Oklahoma City therefore consume large quantities of air conditioning using electricity generated by carbon-intensive power plants. Conversely, households in San Francisco use less electricity generated from more carbon-efficient power plants. As a result of these differences in electrical power plant technologies and climate, government policies that shape the distribution of households across cities may have important implications for national carbon emissions.

[1] For example, Glaeser and Kahn (2010) show that the median household in Memphis is responsible for nearly twice the amount of carbon emissions as the median household in San Francisco.

Local land use restrictions that limit urban population density are often employed to improve the "greenness" of a city. However, these restrictions potentially limit population growth in many of the most desirable cities (Glaeser, Gyourko, & Saks, 2005). Additionally, Glaeser and Kahn (2010) document that cities with lower average household carbon emissions also have tighter land use restrictions, suggesting that these restrictions may discourage people from living in lower carbon-emitting cities. The goal of this paper is to quantify the effect of local land use restrictions on national household carbon emissions.[2]

To this end, we specify and estimate a spatial equilibrium model wherein power plant technologies and energy demand vary across cities. Heterogeneous and imperfectly mobile households choose which city to live in and how much housing, electricity, natural gas, and fuel oil to consume. Rents and wages are determined in equilibrium by the location and consumption choices of these households. Furthermore, cities vary in the tightness of local land use restrictions. All else equal, tighter land use restrictions imply higher costs of living, which will reduce the incentives for households to live in these cities.

Our model allows a household's carbon emissions to vary across cities for two main reasons. First, to reflect climate differences, we allow the utility benefits derived from the use of electricity, natural gas, and fuel oil to vary by city. Second, due to spatial variation in the technologies employed in power plants, the carbon intensity of electricity production varies across cities. Land use restrictions are tighter in cities with more carbon-efficient power plants and lower energy demand. Thus, households will be incentivized to live in "dirtier" cities.

[2]We take the relationship between land-use restrictions and carbon emissions as given and examine the quantify the effect of land use regulations on national carbon emissions. We leave it to future work to determine *why* carbon-efficient cities generally have tighter land-use restrictions.

For our analysis, we combine data from three main sources. We use the US Census and American Community Survey (ACS) for data on households' city of residence, income, and rents, as well as expenditures on electricity, natural gas, and fuel oil for a large sample of households.[3] We combine these household expenditure data with state-level energy prices from the Energy Information Association (EIA) to calculate implied household usage of each of these three energy types. Next, we use data from the Emissions & Generation Resource Integrated Database (eGrid) for power plant locations, output, and CO_2 emissions.

Following Glaeser and Kahn (2010), we use these data to document how the amount of carbon a household emits depends on the city in which they reside. One issue is that differences in emissions across cities may, to some extent, reflect sorting of households with different propensities to use energy. Therefore, the observed differences in average emissions across cities may not reflect the direct effect of location on household emissions. In response to this issue, we employ the semi-parametric selection correction approach introduced by G. B. Dahl (2002) to estimate selection-corrected predicted energy usage and carbon emissions associated with living in various US cities.

We document substantial variation in carbon emissions across cities. For example, we estimate that if a household resides in Memphis, they would produce three times the annual carbon emissions had they resided in Honolulu. Additionally, we show that household carbon emissions are negatively correlated

[3]Throughout the paper we use the 70 largest Core Based Statistical Areas as our definition of cities. The rest of the United States is aggregated up to the census division level.

with a standard measure of land use restrictions, the Wharton Land-Use Regulation Index—greener cities tend to have tighter land use restrictions.[4]

Next, we combine these estimates of predicted energy usage with our other data sources to estimate our spatial equilibrium model. We estimate the parameters of household utility using the two-step estimator introduced by Berry, Levinsohn, and Pakes (1995) with data on household locations from the Census and ACS and our estimates of predicted energy usage.[5] We use data from eGrid to estimate the carbon emissions associated with electricity production across regions and estimates from Saiz (2010) to calibrate the parameters of housing supply curves as a function of land use restrictions.

California legislators recently voted down SB-50, a bill that would have relaxed local land use restrictions in California cities.[6] We use our estimated model to simulate the effects of such a policy. Specifically, we set land use restrictions in California cities to the level faced by the median urban household in the United States. Due to the moderate climate and carbon-efficient power plants, California cities are associated with remarkably low carbon emissions. However, land use restrictions are tight—the San Francisco CBSA is in the 86th percentile of the Wharton Index while Los Angeles CBSA is the 78th.[7]

[4]The Wharton Land-Use Regulation Index was created from a survey sent to 6,896 municipalities across the US, with questions that range from how many regulatory boards one must clear before construction to city-specific density and open space requirements.

[5]This approach has been utilized in a spatial setting by e.g. Diamond (2016), Piyapromdee (2019), and Colas and Hutchinson (2021).

[6]SB-50 was referred to as the "More HOMES Act" (Housing, Opportunity, Mobility, Equity, and Stability). The bill focused on the relaxation of density restrictions and reducing the number of areas zoned for single-family homes, particularly in areas near public transit and in commercial areas.

[7]Quigley and Raphael (2005) write, "California represents the most extreme example of autarky in land-use regulations of any U.S. state."

As a result of relaxing these restrictions, we find that the long-run population in California cities increases. As demands for natural gas and electricity are lower in California, national usage of natural gas and electricity drop by 0.3% and 0.5%, respectively. Overall, this leads to a 0.6% decrease in national household carbon emissions, associated with a decrease in the social cost of carbon of $310 million annually.[8] This change is driven by a decrease in energy usage and an increase in the proportion of total electricity consumption coming from cleaner power plants in California. Furthermore, given that California cities have high productivity levels, this leads to increases in average income for both unskilled and skilled workers.

Next, we entirely remove the existing negative correlation between land use restrictions and carbon emissions by setting the Wharton Index in all cities to the level faced by the median urban household in the US. Households respond by leaving the Midwest and South and moving to the West Coast and Northeast. Demand for natural gas is higher and electricity demand is lower in the cold Northeast. These population shifts therefore increase national gas usage and decrease total electricity usage. Overall, these changes in the spatial distribution of households and their energy consumption lead to a nearly 3.5% drop in national carbon emissions, implying a drop in the social cost of carbon of $1.7 billion annually.

The main goal of this paper is to better understand the relationship between differences in land use restrictions across cities and national household carbon emissions. However, household energy consumption also contributes to local pollution. Therefore, we also use the model to analyze the effects of relaxing land

[8]For the social cost of carbon calculations, we use the estimate of the social cost of carbon in the year 2020 from Nordhaus (2017) of $44.4 per metric ton of CO_2 in 2020 dollars.

use regulations on exposure to local pollutants. We focus on particulate matter of 2.5 micrometers or smaller ($PM_{2.5}$), a common measure of local air quality and that has been at the helm of US air quality regulation for the last two decades.[9] Similar to household carbon emissions, emissions of $PM_{2.5}$ vary across space due to differences in energy consumption and the spatial distribution of power plants. However, unlike carbon emissions, emissions of local pollutants differentially affect air quality across cities. We use the Intervention Model for Air Pollution (InMAP) source-receptor matrix (Tessum et al., 2017) to map electricity production to ambient air concentration of $PM_{2.5}$ across locations in the United States.

We find that the air quality in most US cities improves as a result of relaxing land-use restrictions in California. This result reflects that households use less electricity in California and that power plants used in California, in addition to being carbon-efficient, emit low levels of local pollutants.[10] However, $PM_{2.5}$ exposure for the average household increases slightly, as there is an increase in the number of households in the relatively polluted Southern California.

Related Literature Our paper is related to two recent papers, Hsieh and Moretti (2019) and Herkenhoff, Ohanian, and Prescott (2018), who find that relaxation of land use restrictions in high-productivity cities would lead to large

[9]Exposure to fine-grained particulate matter is associated with various detrimental health and economic outcomes such as (but not limited to): higher infant mortality (Chay & Greenstone, 2005), increased cognitive decline in seniors (Ailshire & Crimmins, 2014), and reduced property values (Chay & Greenstone, 2003).

[10]These reductions in $PM_{2.5}$ concentration in most cities can be thought of as a "co-benefit" of the reduction in carbon emissions. See Aldy et al. (2020) for a discussion of co-benefits in air quality regulation.

increases in GDP.[11] Our model focuses on an entirely different set of outcomes and incorporates energy demand, energy production, and emissions. Households in our model are differentiated in terms of education, family composition, age, race, and the state in which they were born. As a consequence, our model allows for rich substitution patterns across cities and allows us to analyze how changes in land use restrictions affect both the population and demographic composition across cities.

Our work builds on the descriptive findings in Glaeser and Kahn (2010) (GK). GK measures predicted carbon emissions for households in different cities across the country and documents a negative correlation between carbon emissions and land use restrictions. Relative to GK, the primary contribution of this paper is to build and utilize a structural model to quantify the effects of land use restrictions on national carbon emissions. National carbon emissions are determined in equilibrium; household locational sorting, energy demand, and housing supply/demand all determine the extent to which land use restrictions affect national carbon emissions. Estimating the effects of a counterfactual change in land use restrictions necessitates a structural equilibrium model.

This paper is also related to a large literature on how exposure to environmental externalities varies by location (See Chay and Greenstone (2005); Currie, Davis, Greenstone, and Walker (2015); Muehlenbachs, Spiller, and Timmins (2015); or Fowlie, Rubin, and Walker (2019), for example). In our setting, exposure to carbon emissions is independent of the household's location—the effects of carbon emissions and thus climate change are felt globally, not just locally.

[11] Albouy and Stuart (2014) also find that relaxation land use restrictions would lead to a large redistribution of households across cities, but are less concerned with the effect of these changes on national productivity.

However, the amount of carbon dioxide emitted by a household depends on where the household is located.

Another related literature analyses the effects of population density on the local carbon emissions (See Fragkias, Lobo, Strumsky, and Seto (2013); or Jones and Kammen (2014), for example). Other recent research uses simulation methods to analyze the effects of various land use restrictions and transportation policies on within-city locational sorting and local carbon emissions (See Larson and Yezer (2015) or Borck (2016), for example). Our work differs from these two themes in that we focus on the sorting of households *across* cities, rather than the determinants of emissions within a city.[12] This paper is also related to Fan, Fisher-Vanden, and Klaiber (2018), who use a spatial equilibrium model to analyze the effects of climate change on household location choices and welfare. Finally, in other complementary work, Mangum (2016) analyzes the effects of housing and land stock allocations on carbon emissions.[13]

3.2 Data

This paper utilizes individual data on household sorting and energy expenditures from the Census and ACS, detailed data on power plants from eGrid, and state level energy pricing data from the EIA. In what follows, we briefly describe each of the main data sources and how they are used in our analysis. Further details on the data can be found in Appendix B.0.1.

[12] Gaigné, Riou, and Thisse (2012) argue that analysis of the effects of density-increasing policies on carbon emissions must account for relocation of households and firms across cities.

[13] Compared to Mangum (2016), our paper focuses more on the households sorting across cities and energy usage. Mangum's focus on the housing construction process allows for a more nuanced understanding of how different land use restrictions affect the housing stock.

CBSA Level Data We utilize Core Based Statistical Areas (CBSAs) as our definition of a geographic area. CBSAs correspond to distinct labor markets and are the Office of Management and Budget's official definition of a metropolitan area. To measure land use restrictions in each CBSAs we utilize a standard metric developed by Gyourko, Saiz, and Summers (2008), the Wharton Land-Use Regulation Index (WLURI). This index was created from a survey sent to 6896 municipalities across the US, with questions that range from how many regulatory boards one must clear before construction to city-specific density and open space requirements. A higher value of the Wharton-Index implies more stringent restrictions and higher costs of developing land and is associated with more inelastic housing supply curves (e.g. Saiz (2010), Albouy and Ehrlich (2018), or Diamond (2016)).

Household Data We use household-level data from the US Integrated Public Use Microdata Series (IPUMS); we utilize the 1990 Census, the 2000 Census, 5% five-year American Community Survey (ACS) from 2006 - 2010, and 5% five-year American Community Survey (ACS) from 2013 - 2017 (Ruggles et al., 2010). Since our model contains a rich-level of household heterogeneity, a large data-set is imperative for our analysis. IPUMS provides information on yearly, household-level expenditures data on natural gas, electricity, and fuel oil in addition to information on demographics, location, and housing expenditures.[14]

Our model is concerned with emissions generated at home, therefore, we focus on three primary **energy types**: natural gas, electricity, and fuel oil.[15] We

[14]Renters and households living in multi-family homes may not pay for their own energy. We describe how we correct for this in Appendix B.0.1.4.

[15]GK also impute emissions produced by cars. They find that differential emissions from cars are less important than differences in emissions from electricity and natural gas in explaining total

combine data on expenditures on these three energy types with state level price data from the US Energy Information Administration (EIA) to impute household consumption of natural gas, electricity, and fuel oil.[16]

Power Plants and Emissions For each of the three energy types we consider, we use linear conversion factors to map usage of each energy type to carbon emissions. We assume 117 lbs of CO_2 are emitted per thousand cubic feet of natural gas consumed and 17 lbs of CO_2 are emitted per gallon of fuel oil consumed. Carbon emissions associated with electricity usage depend on where the electricity is consumed—electricity used in areas that generate electricity via coal plants will lead to more carbon emissions than in areas that rely more heavily on renewable sources.

We therefore utilize power plant-level data from the Emissions & Generation Resource Integrated Database (eGRID). These data provide information on the location, primary fuel input, emissions rate, and total megawatt hours of electricity generated for every power plant in the US. To assign households to power plants, we use the nine North American Electric Reliability Council (NERC) regions.[17] These regions can be thought of as closed electricity markets, as transmissions of electricity within a region is common but electricity is rarely transferred across regions (Glaeser & Kahn, 2010; Holland & Mansur, 2008).

differences in emissions across cities. Furthermore, emissions from driving calculated in GK are strongly correlated with total household emissions from other sources ($\rho = 0.56$) and a decrease in population density. If removing land use restrictions increases city density, this will also lead to decreases in emissions from driving as well. Therefore, including emissions from driving in our analysis would likely strengthen our main conclusions.

[16] https://www.eia.gov/state/seds/

[17] We omit the region for Alaska since no city in our model is in Alaska.

We calculate the emissions factor associated with each NERC region as the weighted average CO_2 emissions per megawatt hour of electricity of all plants in the NERC region. The emissions factors range from roughly 800 to 1550 lbs of CO_2 emitted per megawatt hour of electricity consumed.[18] All CBSAs within a NERC region have the same CO_2 conversion factor.

For information on local pollutants, we employ data from the National Emissions Inventory (NEI) and the EPA's Air Quality System (AQS) data. The NEI contains information about power plants such as stack height and emissions velocity, and emissions of various local pollutants such as $PM_{2.5}$. We use these data as inputs to construct a "pollution-transfer" matrix, which maps electricity usage in any given city to changes in ambient air quality in all other cities.[19] The AQS data provide hourly levels of total particulate matter by city. We average across all hours in 2017 within each CBSA to obtain our measure of local average $PM_{2.5}$ concentration.[20]

3.3 Descriptive Statistics

In this section, we calculate selection-corrected predicted household usage of electricity, natural gas, and fuel oil usage across cities and the associated carbon emissions. We focus on results calculated using data from 2017; earlier years are primarily used for estimation of parameters in the structural model (described in detail in Section 3.5).

[18] For the full distribution of emissions factors, see Section B.0.1.3.

[19] The pollution-transfer matrix and its construction are described in detail in Sections 3.4.2 and B.0.1.10, respectively. The pollution-transfer matrix maps total energy production in each NERC region to ambient air concentration in all cities in the model.

[20] To obtain census region average $PM_{2.5}$ concentrations, we average over readings for all counties that are not part of the CBSAs in the given region.

Our goal is to isolate the role of a household's location on their energy usage and therefore carbon emissions. This will allow us to understand how the distribution of households – and therefore policies that affect the distribution of households – interact with national carbon emissions. We therefore construct a measure of predicted energy usage per household in each CBSA, controlling for differences in household composition, demographics, and unobserved differences in propensity to consume energy. First, we calculate each household's imputed energy usage in natural gas, electricity, and fuel oil as their reported expenditure on each of these energy types divided by the state-level price of each energy type. We then employ the selection-correction method developed by G. B. Dahl (2002) to compute the selection-corrected predicted usage of each energy type in each city.[21]

With the predicted per-household energy use in hand, we can calculate predicted carbon emissions for each CBSA.[22] We multiply the selection-corrected predicted usage for each fuel type with the respective emissions factor. As discussed in Section 3.2, we assume a constant emissions factor for fuel oil and natural gas. The emissions factor for electricity use varies across NERC regions.[23]

3.3.1 Predicting Energy Usage Across Cities. Consider the following equation for household i's usage of energy type m conditional on living in location j:

$$E_{ij}^m = \alpha_j^m + \beta_j^m X_i + u_i^m, \tag{3.1}$$

[21] The results with no selection correction are included in Appendix B.0.2.2.

[22] We repeat this analysis with methane emissions in Appendix B.0.2.9.

[23] The emissions factor for each NERC region is displayed in Appendix B.0.1.3.

where E_{ij}^m is household i's usage of energy type m, conditional on living in location j, α_j^m is a CBSA-specific intercept term, X_i is a vector of household i's observable characteristics, β_j^m is a vector of parameters which varies by location j, and u_i^m represents household i's idiosyncratic propensity to use energy type m, representing, for example, household i's unobservable preferences for using air conditioning.[24]

One difficulty with estimating (3.1) is that households may sort across locations based on their idiosyncratic propensity to use energy such that u_i^m is not mean-zero conditional on households' chosen locations. For example, households with a low tolerance for cold temperatures might avoid living in cities with cold weather, and may also have a greater propensity to use heating and therefore natural gas. This would induce a correlation between the unobserved propensity to use natural gas and the probability of living in cold cities. Concretely, let $V_{ij} = \bar{V}(X_i, B_i) + \varepsilon_{ij}$ represent household's i's return for living in city j, where $\bar{V}(\cdot)$ is a component of utility which depends on household i's observable characteristics X_i and the birth state of the household head, B_i, and ε_{ij} represents household i's idiosyncratic preference for living in location j.[25] A household chooses to live in location j if it provides the highest return, that is if $j = \text{argmax}_{j'}(V_{ij'})$. If $\mathbb{E}[u_i^m \varepsilon_{ij}] \neq 0$, then this will induce selection on unobservables: $\mathbb{E}\left[u_j^m | j = \text{argmax}_{j'}(V_{ij'})\right] \neq 0$. This would occur, for example, if people who prefer

[24] This term could also vary by location j and be written as u_{ij}^m. We have written it as a household level term rather than a household by location level term for expositional purposes.

[25] Importantly, the household head's birth state is assumed to not differently affect the households energy usage. As such, birth state serves as an exclusion restriction which helps to identify selection on unobservables in the outcome equation. One concern with using birth state as an exclusion restriction is that the climate in which an individual grew up might influence their preferences for energy usage as an adult. In Appendix B.0.2.1, we consider alternative specifications with controls for average temperature in the state where the household head was born. The results are qualitatively similar.

to live in Houston also have stronger preferences for using air conditioning. We will refer to $\mathbb{E}\left[u_j^m | j = \text{argmax}\left(V_{ij'}\right)\right]$ as the "selection bias" term.

To deal with this selection issue, we employ a semi-parameteric selection correction based on the method proposed by G. B. Dahl (2002) (henceforth "Dahl"). Dahl shows that there exists a function that maps the household's choice probabilities to the selection bias term. Concretely, let $\mathbf{P}_{iJ}$ give the vector of the household's choice probabilities for all cities in the set J. Then there exists a function $M_j^m(\cdot)$, such that $M_j^m\left(\mathbf{P}_{iJ}\right) = \mathbb{E}\left[u_j^m | j = \text{argmax}_{j'}\left(V_{ij'}\right)\right]$.[26] Therefore, the selection bias can be controlled for if the econometrician controls for the function $M_j^m\left(\cdot\right)$ such that the estimating equation becomes

$$E_{ij}^m = \alpha_j^m + \beta_j^m X_i + M_j^m\left(\mathbf{P}_{iJ}\right) + \hat{u}_i^m. \tag{3.2}$$

Dahl notes that full estimation of (3.2) is generally infeasible as $M_j^m\left(\cdot\right)$ is an unknown function of the choice probabilities for all J cities. Therefore, we introduce two additional assumptions. First, following Dahl, we make an "index sufficiency assumption": that the function $M_j^m\left(\cdot\right)$ can be replaced with an alternative function which only takes a subset of the choice probabilities as arguments. Second, we assume that the amount of selection on unobservables is constant for all cities within the same state such that $\mathbb{E}\left[u_i^m | j = \text{argmax}_{j'}\left(V_{ij'}\right)\right] = \mathbb{E}\left[u_i^m | \hat{j} = \text{argmax}_{j'}\left(V_{ij'}\right)\right]$ for any j, $\hat{j}$ in the same state. For example, this implies that, conditional on the vector of observables X_i, the expectation of the idiosyncratic term u_i^m is the same in Dallas as it is in Houston.[27]

[26] This is subject to an invertibility condition: that these choice probabilities contain the same information as differences in subutility terms across cities.

[27] This second assumption is useful in generating predicted values of energy usage. As we explain below, to separately identify α_0^m from the intercept of the selection correction function,

Taken together, these two assumptions imply that the control function can be written as a function of a subset of the *state* choice probabilities. Let $\mathbf{P}^{\text{st}}_{i\hat{S}}$ give the vector of the household's choice probabilities for all states in the set $\hat{S}$, where $\hat{S}$ is a subset of the full set of states. We can then estimate the parameters of (3.1) using the equation

$$E^m_{ij} = \alpha^m_j + \beta^m_j X_i + M^m_j\left(\mathbf{P}^{\text{st}}_{i\hat{S}}\right) + \hat{u}^m_i, \tag{3.3}$$

where $M^m_j\left(\mathbf{P}^{\text{st}}_{i\hat{S}}\right)$ is a correction function which depends on $\mathbf{P}^{\text{st}}_{i\hat{S}}$.

In practice, we specify the function $M^m_j(\cdot)$ as a function of the probabilities of choosing the three largest states by population, the probability of choosing the state containing city j, and the interactions between the probability of choosing the state containing city j and the probabilities of choosing the three largest states.[28] For estimates of the state choice probabilities, we use the same approach as Dahl. Specifically, we divide households into cells which vary in their demographic characteristics and their state of birth and calculate state choice probabilities as the proportion of households within each cell that chooses a given state.

Finally, one complication arises because the intercept term α^m_j is not separately identified from the intercept of the control function. We overcome this identification issue by using the intuition of "identification at infinity" (Chamberlain, 1986; Heckman, 1990): suppose the econometrician observes households with demographics $\hat{X}$ and birth state $\hat{B}$ such that $\text{Prob}\left(j = \text{argmax}_{j'}\left(V_{ij'}\right)|\hat{X}, \hat{B}\right) = 1$. Since households with these characteristics

we need to extrapolate the control function to households for which the probability of choosing a given location is equal to one. As state choice probabilities are closer to one than city choice probabilities, using state choice probabilities reduces the range over which we extrapolate the control function.

[28] We show the sensitivity of our estimates to different choices of the correction function in Appendix B.0.2.1.

choose location j with certainty, there is no selection on unobservables for these households: $\mathbb{E}\left[u_i^m | j = \text{argmax}_{j'}\left(V_{ij'}\right), \hat{X}, \hat{B}\right] = 0$. In terms of the selection correction function, this implies that $M_j^m\left(\hat{\mathbf{P}}_{i\hat{S}}^{\text{st}}\right) = 0$, where $\hat{\mathbf{P}}_{i\hat{S}}^{\text{st}}$ is the vector of choice probabilities where the probability of choosing the state containing city j is equal to one and the probability of choosing all other states is equal to zero. We first estimate equation (3.3). Then, we use the restriction that $M_j^m\left(\hat{\mathbf{P}}_{i\hat{S}}^{\text{st}}\right) = 0$ to back out the intercept of the selection correction equation and thus α_j^m.[29]

Finally, after estimating the parameters of the energy usage functions, we calculate predicted usage of energy type m as

$$\hat{E}_j^m = \hat{\alpha}_j^m + \hat{\beta}_j^m \bar{X},$$

where $\hat{\alpha}_j^m$ and $\hat{\beta}_j^m$ denote parameter estimates and $\bar{X}$ gives a vector of the mean values of each demographic characteristic.

[29] As a simple example, consider the case where M_j^m is specified as a first-order polynomial of choosing the state in question: $M_j^m\left(\mathbf{P}_{i\hat{S}}^{\text{st}}\right) = M_0 + M_1 P_{is(j)}^{\text{st}}$, where $P_{is(j)}^{\text{st}}$ is the probability of choosing the state containing city j. Then the intercept of this correction function M_0 can simply be calculated as $M_0 = -M_1$. Note that the identification at infinity argument relies on the econometrican observing households for whom the probability of choosing the state in question is close to 1. If we allow selection to occur at the city level within each state, we would need to observe households for whom the probability of choosing a given city is close to 1. This is a very strong assumption in the case of many cities, given that we only observe data on an individuals state of birth, not their city of birth.

CBSA	Rank	Emissions (1000 lbs)	Gas Emissions (1000 lbs)	Fuel Emissions (1000 lbs)	Electricity Use (MwH)	Electricity Conversion (1000 lbs/MwH)	Electricity Emissions (1000 lbs)
Lowest							
Honolulu, HI	1	14.24	0.00	0.00	9.36	1.52	14.24
Oxnard, CA	2	14.67	6.19	0.27	10.26	0.80	8.21
Riverside, CA	3	16.37	6.33	0.27	12.21	0.80	9.76
San Diego, CA	4	16.71	6.75	0.33	12.04	0.80	9.63
Los Angeles, CA	5	17.14	6.73	0.20	12.76	0.80	10.21
Sacramento, CA	6	17.96	7.72	0.47	12.23	0.80	9.78
Middle							
Baton Rouge, LA	33	26.56	4.41	0.43	20.98	1.04	21.72
Birmingham, AL	34	26.86	5.79	0.21	20.15	1.04	20.86
Jacksonville, FL	35	26.90	0.62	0.06	25.92	1.01	26.22
New Orleans, LA	36	27.15	4.61	0.41	21.38	1.04	22.13
Pittsburgh, PA	37	27.41	12.02	2.43	11.73	1.11	12.97
Houston, TX	38	27.51	4.12	0.13	22.92	1.01	23.25
Highest							
Tulsa, OK	65	40.21	12.47	0.28	21.60	1.27	27.46
Oklahoma City, OK	66	41.59	11.81	0.27	23.21	1.27	29.50
Indianapolis, IN	67	43.67	23.20	0.30	18.26	1.11	20.18
Memphis, TN	68	43.81	10.56	0.23	31.89	1.04	33.02
Omaha, NE	69	45.49	17.31	0.31	22.84	1.22	27.87
Milwaukee, WI	70	46.19	21.84	0.34	21.72	1.11	24.01

Table 2. Predicted CBSA level CO_2 emissions by fuel type for the six lowest emissions cities, the six median cities, and the six highest emissions cities in 2017. The third column ("Emissions") shows the sum of predicted CO_2 emissions from natural gas, fuel oil, and electricity for the CBSA. The next two columns show emissions from gas and fuel oil respectively, which are equal to predicted usage multiplied by the appropriate emissions factor. The last three columns show predicted electricity usage, the electricity emissions factor, and predicted electricity emissions, equal to predicted electricity usage multiplied by the emissions factor.

3.3.2 Selection-Corrected Predicted Usage and Emissions.

The predicted yearly household usage and emissions from the 2017 aggregated ACS for selected cities are shown in Table 2. To calculate these predicted emissions, we multiply selection-correction predicted usage by the appropriate emissions factors. We show results for the six lowest emissions cities, the six highest emissions cities, and the six median cities. The third column ("Emissions") shows the sum of predicted CO_2 emissions from natural gas, fuel oil, and electricity for the CBSA. Predicted household emissions vary considerably across cities. In Honolulu,

predicted emissions are slightly over 14 tons per year, whereas in Milwaukee they are over 46 tons per year.

The next two columns show emissions from gas and fuel oil respectively, which are equal to predicted usage multiplied by the appropriate emissions factor. Natural gas emissions are generally the largest in colder regions.[30] Emissions from fuel oil are generally quite small in magnitude compared to emissions from the other two energy types. The last three columns show predicted electricity usage, the electricity emissions factor, and predicted electricity emissions, which is equal to predicted electricity usage multiplied by the emissions factor.

Spatial variation in household carbon emissions comes from multiple sources. For example, power plants utilized in Memphis emit less CO_2 than Oklahoma City (1.04 lbs per MWh in Memphis compared to 1.27 in Oklahoma City). However, electricity usage in Memphis is so much higher than in Oklahoma City that overall household emissions are higher in Memphis, despite greater consumption of fuel oil and natural gas in Oklahoma City.[31] Conversely, consider emissions resulting from electricity in Houston compared to Tulsa. Households in Houston use more electricity than those in Tulsa. However, power plants near Tulsa are less carbon-efficient than those near Houston. Therefore, emissions from electricity use are higher in Tulsa. This underscores an important feature of the data: spatial variation in household electricity emissions is driven by both differences in energy usage and heterogeneity in power plants across regions.

[30] In Section 3.3.3, as in GK, we show that colder winter temperatures are highly predictive of natural gas usage.

[31] Since consumption of natural gas and fuel oil have the same conversion factor nationally, a higher level of emissions in one city necessarily means a higher level of consumption.

3.3.3 Energy Usage and Climate.

To further understand the differences in energy usage across cities, we now examine the relationship between energy usage and climate. Figure 11 shows the CBSA level relationship between average August temperature and predicted electricity usage and between average January temperature and predicted natural gas usage. Similar to Glaeser and Kahn (2010), we find strong relationships between temperature and consumption of different fuel sources. Electricity usage has a strong positive relationship with August temperature. Similarly, as January temperature increases, natural gas use decreases. Taken together, these results suggest that differences in energy usage across cities are largely driven by differences in climate.

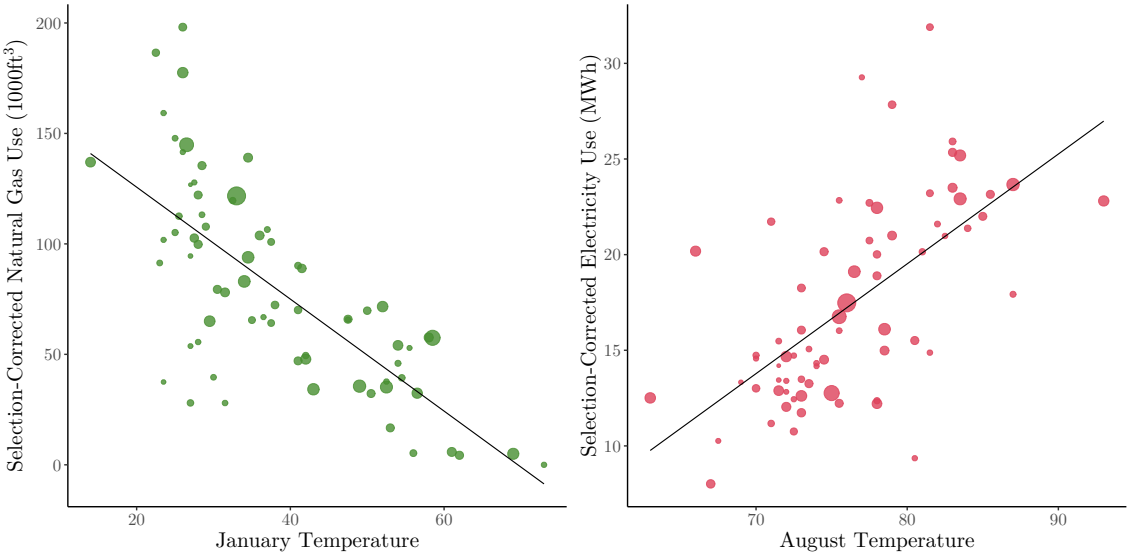


Figure 11. Temperature and household energy use by CBSA. The left panel displays predicted natural gas usage and average January temperature. The right panel displays predicted electricity usage and average August temperature. Temperature data was obtained from weather.com. Each point is a CBSA. Temperature corresponds to the midpoint of the average minimum and maximum daily temperature recorded in the month of interest. The size of each point reflects the population of the CBSA. Electricity usage is measured in MWh and natural gas usage is measured in 1000 ft^3.

3.3.4 Policy and Emissions. Spatial variation in household carbon emissions implies that any policy that affects the spatial distribution of households will also impact national carbon emissions. The primary policy we are interested in is land use restrictions. Figure 12 shows a scatterplot between CBSA-level predicted emissions and the Wharton Land-Use Regulation Index. The Wharton Index is displayed on the horizontal axis; higher values of this index correspond to tighter land use restrictions. The vertical axis displays predicted per-household CO_2 emissions, measured in pounds.[32]

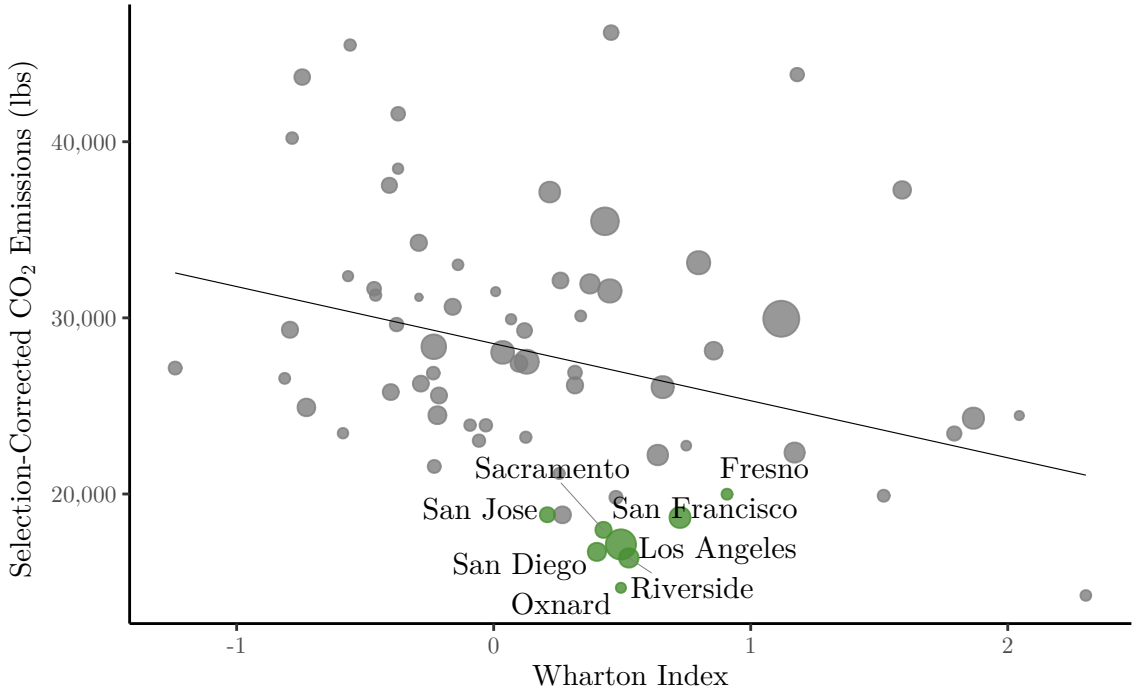


Figure 12. Predicted per-household CO_2 emission measured in pounds and Wharton Lane Use Index. Each circle is a CBSA; California cities are highlighted. The size of each circle reflects CBSA population.

[32]This statistic is the same as that displayed in the "Emissions" column of Table 2.

Generally, tighter land use restrictions are associated with lower predicted carbon emissions. In particular, California cities have very low predicted carbon emissions due to a combination of temperate climate and clean power plants. Cities in California also have very tight land use restrictions, as measured by the Wharton Index.

To be clear, the goal of this paper is not to explain what generates this relationship. Instead, our goal is to study the implication of this correlation, in the sense that tight land use restrictions inflate housing prices and incentivize individuals to live away from California and other cities with low carbon emissions. We proceed by building a spatial equilibrium model to quantify the effects of land use restrictions on national carbon emissions.

3.4 Model

We employ a static spatial equilibrium featuring heterogeneous households, with endogenous wages and rents, similar to those used in Diamond (2016), Piyapromdee (2019), and Colas and Hutchinson (2021). We extend this class of model by allowing locations to vary by carbon and local pollutant output of regional power plants and the marginal benefits of energy usage. Therefore, our model is able to map changes in the distribution of households across locations to carbon and local pollutant emissions.

Household sorting is a static, discrete choice—households choose the city which provides the highest utility in terms of wages, rents, amenities, and energy prices. Households purchase three **energy types**—electricity, natural gas and fuel oil—which they use to produce **energy services** (e.g. air conditioning, home heating). The utility benefit of energy services varies by city. For example, the benefit of air conditioning (and thus electricity) is high in Memphis while the

benefit of home heating (and thus natural gas) is high in Minneapolis. Each location has an upward sloping housing supply curve whose elasticity depends on local land use restrictions and the amount of land available for development. Cities with tighter land use restrictions will have more inelastic housing supply curves and higher rents, all else equal. Competitive firms combine skilled and unskilled workers using a CES production function to produce a numéraire consumption good. Thus, local wages and rents are endogenous to the distribution of households across cities. Changes in land use restrictions across cities will change housing supply curves across cities and impact the equilibrium distribution of households.[33]

As we show in Section B.0.1.1, emissions vary across household types. We allow households in the model to vary in their education level, race, age group, marital status, and whether or not the household has children. These households types vary in their preferences over locations, energy services, and housing. Within this demographic group, households also receive a premium for living in cities close to their birth state. This allows for rich substitution patterns in response to policy changes—a decrease in rents in San Francisco, for example, will lead to larger inflows of households who are born in California. Finally, households receive an idiosyncratic preference draw over each location, where the variance of the draw depends on household demographics. Therefore, households are imperfectly mobile across cities.

The amount of carbon emissions generated by a household varies for two reasons in the model. First, the marginal benefit of each energy type varies by city. Cities with a higher marginal benefit of energy usage will have higher levels of energy usage, all else equal. Second, the production technology and carbon

[33]We define an equilibrium in this setting in Appendix B.0.1.6.

efficiency of power plants vary across NERC regions. Electricity used at a given city must be produced by a power plant in the associated NERC region. Therefore, electricity usage in cities located in NERC regions with more carbon-efficient power plants will lead to lower emissions.

In addition to carbon emissions, household electricity usage also leads to emissions of local pollutants and therefore deterioration in air quality – where we explicitly focus on $PM_{2.5}$ concentration as our measure of local air quality. We construct a pollution-transfer matrix using a state-of-the-art source-receptor matrix, the InMAP source-receptor matrix (ISRM) (Goodkind, Tessum, Coggins, Hill, & Marshall, 2019).[34] We use this pollution-transfer matrix to map household electricity usage, and therefore $PM_{2.5}$ emissions, in any given city to ambient air quality in all cities in the model. In our baseline model, we do not directly account for $PM_{2.5}$ concentration in household utility. In Section 3.7.3 we show that our results are very similar when we include $PM_{2.5}$ concentration in the utility function.[35]

3.4.1 Households. Let j index cities and i index households.[36] Households are endowed with a demographic type d, which includes the household

[34]InMAP is a "reduced-complexity" air transport model. Other popular reduced complexity models include COBRA (EPA, 2020) and APEEP (Muller & Mendelsohn, 2007). All of these models make simplifiying assumptions around atmospheric chemistry equations to ease the computational burden of estimating pollution transfer. More complex atmospheric dispersion models such as HYSPLIT (Stein et al., 2015) and WRF-Chem generally perform better for predicting pollution transfer over long-distances. However, the added complexity makes them far more computationally expensive. InMAP performs similarly to many other reduced-complexity models (Tessum et al., 2017) and is easy to implement, making it our preferred method for modeling pollution transfer. See Hernandez-Cortes and Meng (2020) for a discussion of the limitations of InMAP relative to HYSPLIT.

[35]Our model only endogenizes $PM_{2.5}$ concentration arising from household electricity usage. The changes in $PM_{2.5}$ arising from household electricity usage are small relative to differences in total $PM_{2.5}$ concentrations across cities.

[36]In Section 3.5, we introduce t subscripts to indicate which variables and parameters vary over time. We omit these t subscripts for clarity as we present the model.

head's education, marital status, race, whether or not they have children, and age group.[37] Locations vary by amenities, which we denote as λ_{ij}. To solve their decision problem, the household chooses a location j that yields maximal utility from consumption of the numeraire c, housing H, and energy services $\hat{E}^m$ (such as heating or air-conditioning), and from amenities.

We parameterize the household's utility function as:

$$u_{ij} = \alpha_d^c \log c + \alpha_d^H \log H + \sum_m \alpha_{jd}^m \log \hat{E}^m + \lambda_{ij}, \tag{3.4}$$

where α_d^c, α_d^H, and α_{jd}^m are parameters which scale the marginal benefit of consumption, housing and energy services. We set $\alpha_d^c = 1$ for each demographic group d to normalize for scale. α_{jd}^m varies across cities to reflect differences in the marginal benefit of energy services across cities (perhaps due to differences in weather).[38]

We assume energy services are produced by the household using a fixed proportions energy production function which maps energy types into energy services. Let E^m denote usage of energy type m, where $m \in \{Elec, Gas, Fuel\}$. The energy service production function takes the form:

$$\hat{E}^m = f\left(x^m\right) = a^m E^m, \tag{3.5}$$

where a^m is a parameter that maps units of energy into energy services.

[37] We divide households into two age groups based on the age of the household head. Households with heads over 35 years old are defined as "older" households.

[38] In Section 3.3.3, we show that local temperature is highly predictive of energy usage. We assume that these parameters are not a function of the local population or population density. In Appendix B.0.1.5, we provide evidence that the population is not a significant predictor of energy demand. Fragkias et al. (2013) come to a similar conclusion. Further, we assume that land use restrictions do not directly affect energy use. Some targeted land use restrictions, such as urban growth boundaries or endangered species habitats, may have positive environmental effects. See Lawler et al. (2014), for example, for a discussion.

Substitution of 3.5 into 3.4 yields:

$$u_{ijt} = \log c + \alpha_d^H \log H + \sum_m \alpha_{jd}^m \log E^m + \lambda_{ij} + C, \tag{3.6}$$

where $C = \sum_m \alpha_{jd}^m \log a^m$ is an additive constant that does not depend on the household's choices.

The households' budget constraint is given by:

$$I_{jd} = c + R_j H + \sum_m P_j^m E^m,$$

where I_{jd} is the income level of a household of demographic d living in city j and E^m is usage of energy type m. R_j and P_j^m represent the prices of housing and the price of energy of type m in city j. We normalize the price of consumption of c to one.

We decompose the amenity term, λ_{ij}, into five distinct components. In particular, we let:

$$\lambda_{ij} = \gamma_d^{hp} \mathbb{I}(j \in B_i) + \gamma_d^{\text{dist}} \phi(j, B_i) + \gamma_d^{\text{dist2}} \phi^2(j, B_i) + \xi_{jd} + \sigma_d \epsilon_{ij}, \tag{3.7}$$

where $\mathbb{I}(j \in B_i)$ is an indicator for location j being in the head of household i's birth state, $\phi(j, B_i)$ and $\phi^2(j, B_i)$ are the distance and squared distance, respectively, between the household head's birth state and location j. ξ_{jd} is a shared unobservable component of amenities and ϵ_{ij} is an idiosyncratic preference shock with dispersion parameter σ. Differences in ϵ_{ij} across individuals and cities reflect unobservable variation in attachment to a location that an individual might have. We assume that ϵ_{ij} follows a Type 1 Extreme Value distribution.[39] We make

[39] We do not allow for the possibility of endogenous amenities, as in Diamond (2016). We also refrain from directly modeling the effects of carbon emissions on household's utility. Our model only includes US households, and therefore cannot reliably capture the full social cost of carbon. In Section 3.7.3, we consider an extension in which local pollutants enter the household utility function.

an important assumption that unobserved amenities, ξ_{jd}, are taken as exogenous and are not a function of land use restrictions, as is relatively standard.[40]

Solving the household's maximization problem yields constant income shares on housing and energy of all types. We write a household of demographic group d's optimal choice of housing, conditional on living in a city j as,

$$H^{\star}_{jd} = \frac{\alpha^{H}_{d} I_{jd}}{\boldsymbol{\alpha}_{jd} R_j}$$

where, to simplify notation, we define $\boldsymbol{\alpha}_{jd} \equiv 1 + \alpha^{H}_{d} + \sum_m \alpha^{m}_{jd}$. Optimal usage of energy type m is also a constant fraction of income:

$$E^{m\star}_{jd} = \frac{\alpha^{m}_{jd} I_{jd}}{\boldsymbol{\alpha}_{jd} P^{m}_{j}}. \tag{3.8}$$

We then solve for the indirect utility function associated with location j:

$$V_{ij} = (\boldsymbol{\alpha}_{jd}) \log I_{jd} - \alpha^{H}_{d} \log R_j - \sum_m \alpha^{m}_{jd} \log P^{m}_{j} + \lambda_{ij}, \tag{3.9}$$

where we drop additive constants which have no affect on household decisions. The household's problem can be thought of as a discrete choice over all the locations, conditional on optimal housing and energy consumption. Given that the idiosyncratic preference draws are distributed as Extreme-Value Type I, we can write the probability that a household i chooses a location j as

$$P_{ij} = \frac{\exp\left(\bar{V}_{ij}/\sigma_d\right)}{\sum_{j' \in J} \exp\left(\bar{V}_{ij'}/\sigma_d\right)}, \tag{3.10}$$

[40] See Hsieh and Moretti (2019), Piyapromdee (2019), or Colas and Hutchinson (2021), for example. Diamond (2016) and Herkenhoff et al. (2018) allow amenities to be endogenous to household composition, but not land use restrictions directly. This assumption is supported by the findings in Albouy and Ehrlich (2018), who find that land use restrictions increase the cost of housing production without improving local quality of life. Clearly, some land use restrictions, such as park systems, have a direct positive effect on amenities. We can think of our counterfactuals here as the relaxation restrictions that do not have a large direct effect on local amenities.

where $\bar{V}_{ij} = V_{ij} - \sigma_d \epsilon_{ij}$ is the household's indirect utility of choosing location j minus the idiosyncratic preference draw. We write the total number of households of demographic d who choose to live in location j as N_{jd}.

3.4.2 Energy Production and Emissions.

Carbon Emissions We allow for three types of energy in our analysis: natural gas, fuel oil, and electricity. We assume fuel oil and natural gas are purchased on an international market and treat the supply for these types of energy as perfectly elastic. We assume that the carbon byproduct of fuel oil and natural gas are constant regardless of where energy is consumed. Total household emissions of CO_2 from natural gas and fuel oil in city j is the sum of usage of the energy type multiplied by the appropriate conversion factor:

$$\mathrm{CO}_2{}_j^m = \hat{\delta}^m \sum_d N_{jd} E_{jd}^m, \quad m \in \{Gas, Fuel\},$$

where $\sum_d N_{jd} E_{jd}^m$ is the total amount of fuel of type m consumed by people living in city j and $\hat{\delta}^m$ is the amount of CO_2 emissions per unit of fuel of type m.

We assume electricity is generated across NERC regions in the United States and then is transmitted to local labor markets within those regions. Within each NERC region, perfectly competitive power plants produce electricity.[41] In our baseline specification, we assume that the marginal cost of energy production is constant.[42] In Section 3.7.2, we consider a model extension with

[41] Electricity is a homogeneous good with a large number of with many producers. However, when transmission constraints bind, generation companies may have local market power. See Joskow and Tirole (2000) for a discussion.

[42] The short run supply of electricity is often modeled as a dispatch curve with constant marginal or linear marginal cost curves. However, as we are considering a long-run equilibrium, the supply curve is given by the long run marginal cost curve, allowing for the construction of new reactors or the entry of new plants.

increasing marginal cost. The results are qualitatively similar in both cases.

We allow the conversion factor for electricity to vary by NERC region to reflect geographic variation in the carbon intensity of power plants.[43] For example, a larger percentage of power in the Western NERC region (WECC) comes from hydroelectric dams, whereas the Southern NERC region (SERC) relies more heavily on coal power.

Let $\hat{\delta}^m_{\mathcal{R}}$ represent the conversion factor of electricity to CO_2 emissions in NERC region $\mathcal{R}$ and let $\mathcal{R}(j)$ map cities to their corresponding NERC regions. We write CO_2 emission resulting from electricity usage in CBSA j as

$$CO_2{}^m_j = \hat{\delta}^m_{\mathcal{R}(j)} \sum_d N_{jd} E^m_{jd}, \quad m \in \{Elec\}.$$

For simplicity, we use the following notation for emissions factors:

$$\delta^m_j = \begin{cases} \hat{\delta}^m & m \in \{Gas, Fuel\} \\ \hat{\delta}^m_{\mathcal{R}(j)} & m \in \{Elec\} \end{cases}.$$

Local CO_2 emission of each energy type m can then be written as

$$CO_2{}^m_j = \delta^m_j \sum_d N_{jd} E^m_{jd}.$$

We can then write national emissions from each energy type m as: $CO_2^m = \sum_j CO_2{}^m_j$, and national emissions across all energy types is simply the sum of the energy specific emissions levels, $CO_2 = \sum_m CO_2^m$.

Next we can examine how the distribution of households across cities affects the level of national carbon emissions.[44] We rewrite national emissions in terms of

[43] Note that $\hat{\delta}^m$ for fuel oil and natural do not vary by location as conversion factors for these types of fuel are independent of location.

[44] The main cost of carbon emissions are felt globally and not modeled directly here.

the covariance between the distribution of households and the efficiency of local electricity usage multiplied by the local energy usage:

$$\text{CO}_2 = \sum_m \sum_d \left(J \cdot \text{Cov}\left(N_{jd}, E^m_{jd}\delta^m_j\right) + N_d \mathbb{E}\left[E^m_{jd}\delta^m_j\right]\right),$$

where N_d, the total number of households of group d, and J, the total number of cities, are both model primitives. The expectation $\mathbb{E}\left[E^m_{jd}\delta^m_j\right]$ is taken over cities j. National emissions are increasing in the covariance of population and the product of energy usage and energy conversion factors. Therefore, policies that lead households to live in cities that are associated with higher energy usage and less carbon-efficient power plants will lead to increases in national carbon emissions.

As demonstrated in Section 3.3.4, the tightness of land use restrictions is negatively correlated with local predicted CO_2 emissions levels. Furthermore, tighter land use restrictions increase local rents and, in equilibrium, lead to lower population levels in these cities. In Section 3.6, we examine the quantitative implications of this relationship between land use restrictions, energy demand, and power plant technology on national carbon output.

Local Pollutants In addition to carbon emissions, the model features local pollutants. We focus on Particulate Matter of 2.5 micrometers or smaller ($PM_{2.5}$) as our primary measure of local pollution. We focus on particulate matter emissions from electricity only, as natural gas emits a negligible amount of $PM_{2.5}$ (EPA, 2019).[45] In the model, we distinguish between $PM_{2.5}$ *emissions* (measured in tons/year) and $PM_{2.5}$ *concentration* (measured in micrograms per cubic meter (measured in μ/m^3).

[45]Furthermore, fuel-oil is used by households in relatively few states and will not have first order consequences for overall $PM_{2.5}$ concentration.

While $PM_{2.5}$ is considered a local pollutant because it has negative consequences for those directly exposed, emissions of $PM_{2.5}$ from a given location can impact air quality locally, regionally, and nationally. $PM_{2.5}$ emissions are highly transportable; Morehouse and Rubin (2021) estimate that roughly 90% of particulate matter emissions from coal-power plants leave the state in which they were emitted within 48 hours. Electricity demand in California, for example, will lead to an increase in $PM_{2.5}$ emitted from power plants in the associated WECC NERC region. These additional emissions affect air quality not only in California but potentially all western states and—to a lesser extent—the rest of the United States.

To map emissions of $PM_{2.5}$ from a given NERC region to concentration of $PM_{2.5}$ for each city in the model, we employ a state-of-the-art "source-receptor" (SR) matrix derived from a recent integrated assessment model, the Intervention Model for Air Pollution (InMAP) (Goodkind et al., 2019). The entries of the SR matrix provided by InMAP (henceforth ISRM) are "transfer coefficients" – which give the marginal impact of particulate matter emissions (measured in tons/year) in any given location on the ambient concentration (measured in μ/m^3) in any other location. ISRM accounts for power-plant stack height, the velocity at which the particles were emitted, and local atmospheric conditions. We use ISRM to construct a "pollution-transfer" matrix, where each entry gives the conversion factor between electricity produced in NERC region R (measured in MWh) to pollution in CBSA j (again, measured in μ/m^3). With this matrix in hand, we can estimate the extent to which changes in household energy demand lead to changes in ambient air quality.[46] Let `P` denote this pollution-transfer matrix, and

[46]For details on this procedure, see Appendix B.0.1.10.

let $\mathbf{E}^{\text{elec}}$ give the vector of household electricity produced in each NERC region. The contribution of household electricity usage to $PM_{2.5}$ concentration in each CBSA is given by:

$$\mathbf{PM}^{\text{endog}} = \mathbf{E}^{\text{elec}} \times \mathtt{P}. \tag{3.11}$$

A given element of this vector, PM_j^{endog}, gives the concentration of $PM_{2.5}$ in a city that arises from (national) household energy usage.

In addition to emissions from household electricity usage, $PM_{2.5}$ can originate from many sources (EPA, 2019).[47] To account for this, we assume that $PM_{2.5}$ concentration in a given city is the sum of $PM_{2.5}$ concentration resulting from household electricity usage and $PM_{2.5}$ produced by other sources, which we assume to be invariant to household location choices. Concretely, letting PM_j denote the overall level of $PM_{2.5}$ in city j, we assume

$$PM_j = \overline{PM}_j + PM_j^{\text{endog}}, \tag{3.12}$$

where $\overline{PM}_j$ is the fixed, city-level pollution.

3.4.3 Housing Supply. Each city has an upward sloping housing supply curve. The elasticity of the housing supply curve is allowed to vary by city as a function of the amount of available land and the strictness of land use restrictions. Specifically, we follow Kline and Moretti (2014) and parameterize the inverse housing supply curve in city j as:

$$R_j = z_j H_j^{k_j}, \tag{3.13}$$

[47] Furthermore, household energy consumption form a relatively small proportion of total $PM_{2.5}$ emissions—for details see Appendix B.0.2.3.

where H_j is quantity of housing supplied, z_j is a scale parameter, and k_j is a parameter equal to the inverse elasticity of the housing supply curve $\left(\text{i.e., } \frac{\partial \log R_j}{\partial \log H_{jt}} = k_j\right)$. Taking logs of (3.13), we obtain

$$\log(R_j) = k_j \log(H_j) + \log(z_j). \tag{3.14}$$

The term k_j plays a crucial role in our analysis. Higher values of k_j imply more inelastic housing supply curves and higher rent levels. Therefore, cities with higher values of k_j will have lower equilibrium population levels, all else equal.

As shown by Saiz (2010), local land use restrictions, as measured by the Wharton Land Use Index, and the fraction of land that is unavailable for development due to geographic constraints are strong determinants of more inelastic housing supply curves. We follow Saiz (2010) and parameterize k_j as a function of land use restrictions and geographic constraints:

$$k_j = \nu_1 + \nu_2 \psi_j^{WRI} + \nu_3 \psi_j^{GEO}$$

where ψ_j^{WRI} is the Wharton Land Use Index and ψ_j^{GEO} measures the amount of land that is unavailable for development due to geographic restrictions.[48] A higher value of ν_2 implies that cities with tighter land use restrictions will have a more inelastic housing supply. As shown in Section 3.3.4, cities with higher values of ψ_j^{WRI} generally have lower carbon emissions per household. In the model, this disincentivizes households from living in cities with low carbon emissions.

[48] Another option would be to use a more dissagregated measure of land use restrictions. This would allow us to decompose the effects of various types of land use restriction on carbon emissions.

Specifically, given that the idiosyncratic preferences draws are distributed as Extreme-Value Type I, the partial equilibrium elasticity of location choice with respect to rents is approximately equal to:[49]

$$\frac{\partial \log P_{ij}}{\partial \log R_j} \approx -\frac{\alpha_d^H}{\sigma_d}.$$

We can solve for the partial equilibrium effect of a household's choice probability with respect to land use restrictions as

$$\frac{\partial \log P_{ij}}{\partial \psi_j^{WRI}} \approx -\nu_2 \frac{\alpha_d^H}{\sigma_d} \log(H_{jt})$$

The partial equilibrium effect of land use restrictions is proportional to the expenditure share on housing and the importance of land use restrictions in dictating the housing supply elasticity ν_2, and inversely proportional to σ_d, the dispersion in the idiosyncratic preference draw. Higher values of σ_d imply household location choices are less responsive to changes in rents; thus, variation in land use restrictions will have smaller effects on household sorting.

3.4.4 Wages. Perfectly competitive firms in each city combine skilled and unskilled labor in a CES production function to produce the numéraire consumption good, where we define household heads with a college degree as skilled and household heads with less than a college degree as unskilled.[50] Therefore, wages for skilled and unskilled workers in each city are determined endogenously by the ratio of skilled to unskilled workers. Specifically, firms use a combination of skilled (S) and unskilled labor (U), as inputs in the following production function:

[49] Differentiating P_{ij} with respect to rents yields $\frac{\partial \log P_{ij}}{\partial \log R_j} = -\frac{\alpha_d^H}{\sigma_d}(1 - P_{ij}) \approx -\frac{\alpha_d^H}{\sigma_d}$ for small values of P_{ij}.

[50] Data on energy usage by firms are generally less readily available than data on household energy usage. As such, we choose to focus on household emissions. Glaeser and Kahn (2010) argue that commercial energy use and household energy use are likely to be highly correlated.

$$Y_j = A_j[(1-\theta_j)U_j^{\frac{\varsigma-1}{\varsigma}} + \theta_j S_j^{\frac{\varsigma-1}{\varsigma}}]^{\frac{\varsigma}{\varsigma-1}} \tag{3.15}$$

where U_j and S_j are defined as the total efficiency units of labor supplied by unskilled and skilled workers in city j, respectively. A_j is the total factor productivity in city j and θ_j is the relative factor intensity of skilled workers. The elasticity of substitution between skilled and unskilled workers is given by ς.[51]

Firms take wages as given and choose skilled and unskilled labor quantities to maximize profits. We derive labor demand curves as a result of the firms skilled and unskilled labor first order conditions for profit maximization:

$$\begin{aligned} W_{js} =& A_j \left(\frac{Y_j}{A_j}\right)^{\frac{1}{\varsigma}} \theta_j S_j^{-\frac{1}{\varsigma}} \\ W_{ju} =& A_j \left(\frac{Y_j}{A_j}\right)^{\frac{1}{\varsigma}} (1-\theta_j)\, U_j^{-\frac{1}{\varsigma}}, \end{aligned} \tag{3.16}$$

where W_{js} and W_{ju} are the wage rates for skilled and unskilled labor, respectively.

Within education groups, demographic groups are perfectly substitutable in production but vary in their productivity and therefore supply different amounts of efficiency units of labor. Income levels for an individual household are given by the number of efficiency units of labor supplied by the household multiplied by the appropriate wage rate. Income for a household of demographic group d living in city j is given by $I_{jd} = W_{ju}\ell_d$ for unskilled workers and $I_{jd} = W_{js}\ell_d$ for skilled

[51] One straightforward way to introduce capital into the model is the assume that production is Cobb-Douglas in capital and a CES labor supply such that $Y_{jt} = A_{jt}K_{jt}^{\eta}\left([(1-\theta_j)U_j^{\frac{\varsigma-1}{\varsigma}} + \theta_j S_j^{\frac{\varsigma-1}{\varsigma}}]^{\frac{\varsigma}{\varsigma-1}}\right)^{1-\eta}$ where η is a parameter. If capital supply is perfectly elastic, this production function implies wage equations that are equivalent to those here. See Colas (2019) for details.

workers, where ℓ_d represents the number of efficiency units supplied by agents of demographic group d.

3.5 Data Inference

In this section we describe our estimation procedure. We focus most of our exposition on the estimation of household location choice and energy use parameters. Estimation of the housing supply and production are relatively standard and details are therefore relegated to appendices B.0.1.8 and B.0.1.9. The carbon emissions factors are calculated as in Section 3.2.[52] We choose to use the 70 largest CBSAs, as defined by population in 1980. These 70 locations make up approximately 55% of the entire US population in 2017. We map individuals that do not live in one of these 70 areas into their corresponding census division, creating nine additional choices.

Note that by defining our locations as CBSAs, we are abstracting from household location choices across municipalities or neighborhoods within CBSAs. While neighborhoods or municipalities within a CBSA may differ in many dimensions, they are unlikely to differ substantially in their associated carbon emissions because climate and the set of power plants where electricity is produced are relatively constant within a given CBSA.

3.5.1 Households. We estimate household preferences using the two-step "BLP" procedure using repeat cross-sectional data from the 1990 Census, 2000 Census, 2010 aggregated ACS, and 2017 aggregated ACS (Berry, Levinsohn, &

[52] That is, we assume 117 lbs of CO_2 emitted per thousand cubic feet of natural gas consumed and 17 lbs of CO_2 emitted per gallon of fuel oil consumed. We calculate the weighted average CO_2 emissions of all plants in a NERC region. We then assign each of the CBSAs to a NERC region, thus assigning all individuals in our sample a carbon emissions factor for electricity.

Pakes, 2004).[53] As we estimate household preferences using multiple cross sections of data, we introduce t subscripts to indicate which variables and parameters vary over time and which parameters are assumed to be constant over time. We rewrite the household's indirect utility function as

$$V_{ijt} = (\boldsymbol{\alpha}_{jd}) \log I_{jdt} - \alpha_d^H \log R_{jt} - \sum_m \alpha_{jd}^m \log P_{jt}^m + \gamma_{dt}^{hp} \mathbb{I}(j \in B_i) + \gamma_{dt}^{\text{dist}} \phi(j, B_i) + \gamma_{dt}^{\text{dist2}} \phi^2(j, B_i) + \xi_{jdt} + \sigma_d \epsilon_{ijt}. \tag{3.17}$$

Therefore, the set of parameters to be estimated are α_d^H and α_{jd}^m, the parameters governing the budget shares of consumption, housing and energy spending, respectively; γ_{dt}^{hp}, $\gamma_{dt}^{\text{dist}}$ and $\gamma_{dt}^{\text{dist2}}$, the parameters governing the strength of home premium and the disutility of living further away from one's birth state; ξ_{jdt}, the unobserved city-level amenities; and σ_d, the parameters that govern the variance of idiosyncratic preference draws.

It will be useful to express indirect utility as the sum of the component of utility that varies by household and the "mean utility" that is constant for households of a given demographic group. Dividing (3.17) by σ_d, we can write indirect utility as

$$\hat{V}_{ijt} = \mu_{jdt} + \hat{\gamma}_{dt}^{hp} \mathbb{I}(j \in B_i) + \hat{\gamma}_{dt}^{\text{dist}} \phi(j, B_i) + \hat{\gamma}_{dt}^{\text{dist2}} \phi^2(j, B_i) + \epsilon_{ijt}, \tag{3.18}$$

where

$$\mu_{jdt} = \frac{(\boldsymbol{\alpha}_{jd})}{\sigma_d} \log I_{jdt} - \frac{\alpha_d^H}{\sigma_d} \log R_{jt} - \sum_m \frac{\alpha_{jd}^m}{\sigma_d} \log P_{jt}^m + \hat{\xi}_{jdt} \tag{3.19}$$

and where "hatted" parameters represent a given parameter divided by σ_d. Using (3.19), we write the probability that an household i chooses a location j as

[53] This estimation procedure has been utilized extensively in the urban economics literature, for example by Diamond (2016), Piyapromdee (2019), and Colas and Hutchinson (2021)

$$P_{ijt} = \frac{\exp(\mu_{jdt} + \hat{\gamma}_{dt}^{hp}\mathbb{I}\left(j \in B_i\right) + \hat{\gamma}_{dt}^{\text{dist}}\phi\left(j, B_i\right) + \hat{\gamma}_{dt}^{\text{dist2}}\phi^2\left(j, B_i\right))}{\sum_{j' \in J}\exp\left(\mu_{j'dt} + \hat{\gamma}_{dt}^{hp}\mathbb{I}\left(j' \in B_i\right) + \hat{\gamma}_{dt}^{\text{dist}}\phi\left(j', B_i\right) + \hat{\gamma}_{dt}^{\text{dist2}}\phi^2\left(j', B_i\right)\right)}. \quad (3.20)$$

In the first step of estimation, we estimate the birth state premium parameters and the mean utility via maximum likelihood. The log-likelihood function is given by

$$\mathcal{L}(\hat{\gamma}_{dt}^{hp}, \hat{\gamma}_{dt}^{\text{dist}}, \hat{\gamma}_{dt}^{\text{dist2}}, \mu_{jdt}) = \sum_{i=1}^{N_d}\sum_{j=1}^{J}\mathbb{I}_{ij}\log(P_{ijd}), \quad (3.21)$$

where $\mathbb{I}_{ij}$ is an indicator equal to one if individual i lives in location j and zero otherwise.[54]

In the second step of estimation, we decompose the mean utility terms, μ_{jdt}. First, we define $\tilde{\alpha}_{jd}^m = \frac{\alpha_{jd}^m}{\boldsymbol{\alpha}_{jd}}$. Given the Cobb-Douglas utility function, the expenditure share on fuel type m of demographic group d in city j is given by $\frac{E_{jd}^m P_j^m}{I_{jd}} = \tilde{\alpha}_{jd}^m$. We first choose the $\tilde{\alpha}_{jd}^m$ parameters to match the expenditure share on each fuel type by each demographic group in each city. Specifically, we calculate the expenditure share on each type of fuel by city and demographic group using our selection-corrected energy usage estimates from Section 3.3.1.[55]

As we show in Appendix B.0.1.11, we can rewrite mean utility as

$$\mu_{jdt} = \beta_d^w \tilde{I}_{jdt} + \beta_d^r \log R_{jt} + \hat{\xi}_{jdt}. \quad (3.22)$$

where $\tilde{I}_{jdt} = \frac{\log I_{jdt} - \sum_m\left(\tilde{\alpha}_{jd}^m \log P_{jt}^m\right)}{1 - \sum_m \tilde{\alpha}_{jd}^m}$, $\beta_d^w = \frac{\left(1+\alpha_d^H\right)}{\sigma_d}$, and $\beta_d^r = -\frac{\left(\alpha_d^H\right)}{\sigma_d}$. We refer to $\tilde{I}_{jdt}$ as "energy-budget adjusted income". This is a household's log income after adjusting

[54] Computationally, we invert the choice probabilities using the contraction mapping in Berry (1994) to obtain the unique mean utility associated with every guess of the parameter vector $[\hat{\gamma}_d^{hp} \;\; \hat{\gamma}_d^{\text{dist}} \;\; \hat{\gamma}_d^{\text{dist2}}]$.

[55] That is, we use the estimates of selection corrected-emissions from Section 3.3.2 and calculate demographic specific energy usage as $E_j^m = \tilde{\alpha}_j^m + \hat{\beta}_j^m \bar{X}_d$, where $\bar{X}_d$ gives the vector of demographic characteristics of households in group d.

for the fact that 1) the fraction of income that is spent on energy depends on local energy prices, and 2) income is more valuable in locations with high marginal utility of energy.

To limit the number of parameters to be estimated, we place additional restrictions on α_d^H and σ_d. As we show in Appendix B.0.1.1, conditional on location, a household's marital status, presence of children, and age of the household head are the most important determinants of emissions. Conditional on these characteristics, the education level and race of the household head play only a minor role in determining emissions. As such, we allow the σ_d and α_d^H parameters to vary by household marital status and the presence of children.[56] Therefore, we can write these parameters as $\alpha^H_{Marr,Child}$ and $\sigma_{Marr,Child}$ and let $\beta^w_{Marr,Child} = \frac{\left(1+\alpha^H_{Marr,Child}\right)}{\sigma_{Marr,Child}}$ and $\beta^r_{Marr,Child} = -\frac{\left(\alpha^H_{Marr,Child}\right)}{\sigma_{Marr,Child}}$.

Plugging in these parameter restriction and taking first differences of (3.22) over our four datasets yields our estimation equation:

$$\Delta\mu_{jdt} = \beta^w_{Marr,Child}\Delta\tilde{I}_{jdt} + \beta^r_{Marr,Child}\Delta\log R_{jt} + \Delta\xi_{jdt}. \tag{3.23}$$

In general, we expect that changes in unobserved amenities, $\Delta\xi_{jdt}$, will be correlated with changes in rents, $\Delta\log R_{jt}$, and adjusted incomes $\Delta\tilde{I}_{jdt}$. For example, consider an increase in unobservable amenities in city j: this will increase utility directly and induce households into city j. Mechanically, this leads to an

[56] Many papers in the literature focus on differences in mobility by education group, rather than by marital status and the presence of children (e.g. Bound and Holzer (2000) or Diamond (2016)). However, since education is not a strong predictor of household emissions, we found it much more important to focus on marital status and the presence of children, which play a large role in determining household emissions. We also considered specifications in which we allowed these to vary by marital status, the presence of children, and age of the household head. These results are included in Appendix B.0.2.4.

increase in housing demand, and as a result equilibrium rents rise, thus causing a change in $\Delta \log R_{jt}$. A similar argument can be made for adjusted income. As such, we estimate (3.23) via two-step GMM, using instrumental variables to deal with these endogeneity issues.[57]

First, we use the measure of labor-demand shifts introduced by Katz and Murphy (1992) to generate variation in income across cities.[58] The instrument interacts historical industry concentration patterns at the city level with national changes in hours worked across industry. Formally, letting ι index industries, and letting $e(d)$ denote the education group associated with demographic group d, the Katz-Murphy index for city j from the previous period t' to the period in question t can be written as

$$\Delta Z_{jdt} = \sum_{\iota} \omega^{1980}_{\iota j e(d)} \left(\text{Hours}_{\iota, e(d), -j, t} - \text{Hours}_{\iota, e(d), -j, t'} \right),$$

where $\omega^{1980}_{\iota j e(d)}$ is the share of total hours worked in industry ι in city j by education group $e(d)$ in 1980 as a share of total hours worked in city j by education group $e(d)$ in 1980. $\text{Hours}_{\iota, e(d), -j, t}$ is the national hours worked in industry ι, education $e(d)$, for all cities besides city j. Therefore, the term in parentheses gives the change in national hours worked in industry ι between the current time period and the previous time period. Cities with historical concentrations of growing industries will generally experience increases in income while cities with declining industries will experience decreases in income. The income changes generated by

[57] Estimates via two-stage least squares, continuously updating GMM and limited information maximum likelihood are very similar.

[58] The instrument has been used as an instrument for cross city wage changes in Piyapromdee (2019) and Notowidigdo (2013).

this instrument are assumed to be uncorrelated with $\Delta\xi_{jdt}$, changes in city-level unobservable amenities.[59]

To generate variation in rents, we also include the ψ_j^{WRI}, our measure of land-use restrictions, and the interaction between the Katz-Muphy index and ψ_j^{WRI} as instruments. In essence, cities with tighter housing supply restrictions and therefore more inelastic housing supply curves will experience larger changes in rents, especially in response to changes in population. As an example, if two cities experience a positive labor-demand shock, captured by a positive value of the Katz-Murphy index, the city with the more inelastic housing supply curve will experience a larger rent increase. This variation in rents is assumed to be uncorrelated with changes in unobservable amenities.[60]

3.5.2 Particulate Matter Concentration. Next, we estimate the fixed level of $PM_{2.5}$, $\overline{PM}_j$, in each city. Recall from (3.12) that total particulate matter concentration in a city, PM_j, is equal to the sum of particulate matter that is endogenous to household location choices and $\overline{PM}_j$, the level of particulate matter arising from other sources. We measure PM_j using data on the overall level of ambient particulate matter concentration for each CBSA from the EPA's Air Quality Systems data. With the total level and endogenous component in hand, the exogenous component of particulate matter concentration can be calculated as

$$\overline{PM}_j = PM_j - PM_j^{\text{endog}}.$$

3.5.3 Parameter Estimates and Model Validation.

[59] See Goldsmith-Pinkham, Sorkin, and Swift (2020) for a discussion of identification with "Bartik"-style instruments.

[60] We consider alternative instruments and parameterizations in Appendix B.0.2.4.

Location Choice Parameters Table 3 shows our estimates of β_d^w and β_d^r, the parameters which determine the location choice elasticities with respect to adjusted income and rents. The first column provides estimates for single households, the second for married households without children, and the third column provides estimates for married households with children. We estimate that β_d^w and β_d^r are largest in magnitude for single households, and lowest for married households with children, implying that single households will be the most responsive in their location decisions to changes in policy. This heterogeneity in location choice elasticities has important implications for the effects of relaxing land-use restrictions in California. In particular, this implies that single households will be the most likely to move to California in response to this policy change. However, as we demonstrate in Appendix B.0.1.1, single households have the lowest average carbon emissions of these three groups while married households with children have the largest. Households who are the most mobile also have the smallest impact on carbon output.

		Married	
	Single	No Children	With Children
β^w: Adjusted Income	15.09	11.72	7.33
	(2.80)	(2.19)	(1.47)
β^r: Rent	-9.03	-6.90	-4.82
	(2.40)	(1.89)	(1.29)
σ: Idiosyncratic Component	0.17	0.21	0.40
	(0.03)	(0.04)	(0.11)
α^H: Housing Parameter	1.49	1.44	1.92
	(0.48)	(0.48)	(0.73)
Cragg-Donald Wald F Statistic	8.09	8.28	8.57

Table 3. Parameter Estimates. Standard errors in parentheses.

To the best of our knowledge, we are the first paper to estimate these parameters by marital status and by the presence of children. Therefore, it is difficult to directly compare our estimates to those in the literature. However, it is reassuring that the magnitude of our location choice elasticities are similar to those in Colas and Hutchinson (2021) and only slightly larger than those in Diamond (2016), who estimate parameters which vary by education but not by marital status or presence of children.[61]

The next rows of Table 3 translate these estimates to estimates of α_d^H and σ_d.[62] We can use these parameters to calculate the budget share of housing for each demographic group in each city as $\frac{\alpha_d^H}{\boldsymbol{\alpha_{jd}}}$. We find an average budget share of housing across household groups of .49, which is consistent with what has been previously estimated in the literature.[63]

Birth State Premium Appendix B.0.2.7 gives the estimates of γ_{dt}^{hp}, $\gamma_{dt}^{\text{dist}}$ and $\gamma_{dt}^{\text{dist2}}$, the parameters governing the strength of home premium and the disutility of living further away from one's birth state, for each year. For all years and demographic groups, households receive a large utility premium for choosing a location in their birth state. The utility value of a location is decreasing and convex in distance from birth state for all demographic groups.

[61] One concern is that we might suffer from a slight weak instruments problem as is relatively common in this literature. We assess the robustness of our key results to these parameter estimates in Section 3.7.1.

[62] This parameters cannot be directly compared with similar parameters in Colas and Hutchinson (2021), Piyapromdee (2019), or Diamond (2016) as α^H here does not equal the budget share of housing. Here, the budget share of housing is given by $\frac{\alpha_d^H}{\boldsymbol{\alpha_{jd}}}$.

[63] For example, Suárez Serrato and Zidar (2016) estimate a budget share of housing of 0.3, using data from the Consumer Expenditure Survey, Moretti (2011) estimates a budget share of housing of .41 using Census data and Diamond (2016) calibrates an expenditure share of local goods of 0.62.

Panel (a): White	College or more		Less than College	
Rank	Single (no kids)	Married (with kids)	Single (no kids)	Married (with kids)
1	Portland, OR	Portland, OR	San Diego, CA	Seattle, WA
2	Miami, FL	Miami, FL	Miami, FL	Portland, OR
3	Los Angeles, CA	Seattle, WA	Portland, OR	Los Angeles, CA
4	San Diego, CA	Los Angeles, CA	Seattle, WA	Honolulu, HI
5	Seattle, WA	San Diego, CA	Oxnard, CA	San Diego, CA
66	Memphis, TN	Memphis, TN	Springfield, MA	Memphis, TN
67	Youngstown, OH	Worcester, MA	Worcester, MA	Springfield, MA
68	Syracuse, NY	Springfield, MA	Albany, NY	Worcester, MA
69	Springfield, MA	Syracuse, NY	Rochester, NY	Albany, NY
70	Worcester, MA	Youngstown, OH	Syracuse, NY	Syracuse, NY
Panel (b): Non-white	College or more		Less than College	
Rank	Single (no kids)	Married (with kids)	Single (no kids)	Married (with kids)
1	Honolulu, HI	Los Angeles, CA	Los Angeles, CA	Los Angeles, CA
2	Los Angeles, CA	Honolulu, HI	Miami, FL	Honolulu, HI
3	Miami, FL	Seattle, WA	San Francisco, CA	Seattle, WA
4	Portland, OR	Miami, FL	San Diego, CA	San Francisco, CA
5	San Diego, CA	San Francisco, CA	Seattle, WA	Portland, OR
66	Rochester, NY	Knoxville, TN	Springfield, MA	Springfield, MA
67	Scranton, PA	Milwaukee, WI	Syracuse, NY	Albany, NY
68	Milwaukee, WI	Syracuse, NY	Albany, NY	Syracuse, NY
69	Youngstown, OH	Springfield, MA	Milwaukee, WI	Rochester, NY
70	Springfield, MA	Youngstown, OH	Rochester, NY	Milwaukee, WI

Table 4. Demographic group city ranks according to the shared, unobservable component of amenities for households with younger household heads.

Amenities Table 4 provides selected estimates of ξ_{jdt}, the shared unobservable component of amenities, for the year 2017. Recall that this parameter is allowed to vary by demographic group d and location j, meaning we estimate a separate value of ξ_{jdt} for each of our 24 demographic groups in each of our 79 locations for each year of our data. Table 4 displays the five cities with the highest and lowest values of ξ_{jdt} for households in the younger age group that vary in their race, education level, marital status, and the presence of children. The estimates for households with older household heads are similar and are included in Appendix B.0.2.8.

Across demographic groups, Miami, Los Angeles, and Seattle consistently rank among the highest amenity cities while upstate New York cities generally have low amenities. There is also interesting heterogeneity across demographic groups—

Portland is especially popular among educated white households, while Honolulu is more popular among minorities. Our estimates of ξ_{jdt} can be compared to estimates of "Quality of Life" from the urban economics literature.[64] Compared to Albouy (2012), our estimates assign slightly higher amenities to higher population cities relative to lower population cities. Consistent with Kahn (1995), we find that Los Angeles and San Francisco have higher amenities than Chicago and Houston in each year.

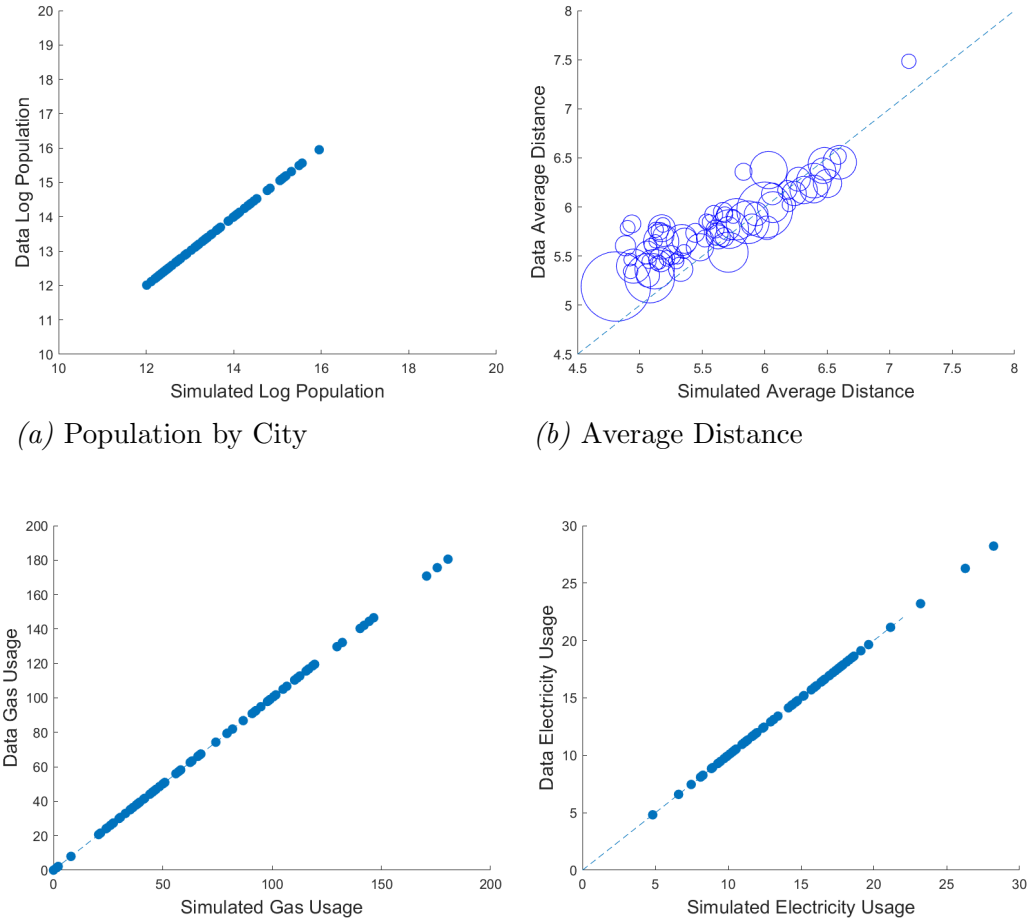
Model Fit Next, we assess how well our model fits the data. The results from 2017 are summarized in Figure 13. Panel (a) shows the log number of households in each city in the data and the baseline simulation. Each circle represents a CBSA. Given that we estimate a separate unobserved amenity value for each demographic group and each city (ξ_{jdt}), we can match these moments exactly. Next, we plot the simulated and observed log average distance between an agent's birth state and chosen city for each CBSA. The results are displayed in panel (b) of Figure 13. Each circle represents a CBSA, and the size of the circle is proportional to its population. The model fits this aspect of the data fairly well.

Panels (c) and (d) of Figure 13 show the predicted and actual average usage of natural gas and electricity in each city. As we allow the benefit of energy usage (α^m_{jdt}) to vary by city and demographic group, we can match these moments exactly.

3.6 Counterfactuals

In this section, we use the estimated model to simulate changes in land use restrictions. The results of the counterfactuals are summarized in Table 5. The first

[64] E.g. Blomquist, Berger, and Hoehn (1988), Kahn (1995) Albouy (2012). See Lambiri, Biagi, and Royuela (2007) for a review.



(a) Population by City

(b) Average Distance

(c) Natural Gas

(d) Electricity

Figure 13. Model fit results. Each circle represents a CBSA. Panel (a) shows the log number of households in each city in the data and the baseline simulation. Panel (b) plots the simulated and observed log average distance between an agent's birth state and chosen city for each city. The size of the circle is proportional to a city's population. Panels (c) and (d) show the predicted and actual average usage of natural gas and electricity in each city.

column shows the population distribution, fuel usage, emissions, and income in the baseline specification, with all parameters set at their baseline levels. The following columns present these statistics in each counterfactual.

	(1) Baseline	(2) Relax Cali	(3) Relax All
I. Percent Total Population			
California Cities	9.1	11.0	7.2
Other West	13.6	13.1	17.8
Midwest	22.2	21.7	9.3
South	37.3	36.6	23.1
Northeast	17.9	17.6	42.6
II. Mean Usage			
Gas (1000 cubic feet)	74.4	74.2	74.9
Electricity (MW h)	17.1	17.0	15.8
Fuel Oil (gallons)	60.4	59.5	138.6
III. Mean Emissions (lbs of CO_2)			
Gas	8711	8686	8772
Electricity	16331	16211	13242
Fuel Oil	1622	1598	3723
Total	26664	26495	25737
(%)	100	99.4	96.5
IV. Income (Relative to Baseline)			
Skilled	100.0	100.5	113.0
Unskilled	100.0	100.0	100.4
All	100.0	100.2	104.8

Table 5. Counterfactual results when California land-use regulations are relaxed. Each panel shows the simulated percent of the total population living in various geographic areas, mean energy usage, mean emissions, and mean income in each specification. See text for details on each simulation.

3.6.1 Relaxation of Land Use Restrictions in California.

California Senate Bill 50—which recently failed in the California legislature—would have overridden tight land use restrictions in California cities. In this section, we examine the effects of California adopting such a policy and relaxing local land use restrictions. As shown in Section 4.3, California cities are among the most carbon efficient in the country. However, they also have very tight land use

restrictions—San Francisco and Los Angeles are in the 86th and 78th percentiles in the strictness of land use restrictions, respectively. Intuitively, the relaxation of land use restrictions in California will lead to increases in California's population and decreases in overall carbon emissions. However, the magnitude of the decrease is an empirical and quantitative question.

Specifically, we simulate setting land use restrictions, ψ^{WRI}, in California cities to the level faced by the median urban household.[65] We display the main results in the second column of Tables 5. Setting land use restrictions in California to the level of the median urban household leads to a 20.5% increase in the total population in California cities, a 3.1% drop in the population of other locations in the West, and 1% to 2% drops in the Midwest, South, and Northeast.

	California Cities	Other West	Midwest	South	Northeast
I. Household Distribution					
% Change Population	20.5	-3.1	-1.9	-2	-1.7
II. Composition					
Change in Single Share	2.4	-0.5	-0.2	-0.2	-0.2
Change Share without Children	2.3	-0.4	-0.2	-0.1	-0.1
Change in College Share	-0.4	0.0	-0.1	-0.1	-0.1
Change Minority Share	0.1	-0.4	-0.3	-0.4	-0.3
III. Prices					
% Change Skilled Income	-0.1	0.2	0.2	0.2	0.2
% Change Unskilled Income	-1.0	0.2	0.0	0.0	0.0
% Change Average Rents	-4.8	-1.2	-0.7	-0.7	-0.8

Table 6. Changes in the composition of population in response to reduction in California land use restrictions.

Panels II and III of Table 5 show how these changes in the distribution of households translate to average usage and emissions. The relaxation of land use restrictions leads to decreases in usage of all three types of fuel, as households

[65] In a previous version of the paper, we simulated relaxing restrictions to the level of the median city.

move to the temperate California climate. Specifically, natural gas usage drops by 0.3%, electricity by 0.5%, and fuel oil usage drops by 1.5%. Panel III of Table 5 displays average emissions resulting from each type of fuel. Electricity emissions drop by over 0.7% despite only a 0.5% decrease in usage. As power plants utilized in California are relatively carbon-efficient, the drop in emissions from electricity is larger than the drop in electricity usage. All together, this implies a drop in national household carbon emissions of 0.6% or a $310 million dollar drop in the social cost of carbon annually.[66]

In addition to low emissions, cities in California are very productive. Panel IV of Table 5 shows the effects on average income, relative to the average income in the baseline. The average income of skilled workers increases by roughly 0.5% while the average income of unskilled workers increases slightly. This leads to an increase in income of 0.2% across all workers. Overall, the shift towards more productive and lower-emitting cities increases the output to emissions ratio by 0.7%.

Regional Effects To better understand the regional impacts of the policy change, Table 6 gives the change in population distribution and prices across regions. Panel II shows the change in regional demographic composition. The change in land use restrictions leads to increases in the share of unmarried households and households without children in California. As these groups are relatively lower usage groups, this composition effect leads to slightly smaller decreases in carbon emissions than the population change alone.

Panel III of Table 6 displays the change in average incomes and rents across regions. Interestingly, within California cities, average income decreases slightly, as

[66] As mentioned in the introduction, we use the estimate of the social cost of carbon in 2020 from Nordhaus (2017).

the proportion of households in low-income Fresno increases while the proportion in high-income San Jose decreases. Equilibrium rents in Californian cities drop by roughly 5%, as a result of both the change the land-use regulations and the resulting increase in population in California. Average rents decrease by roughly 1% in the other regions, reflecting drops in regional housing demand as households move to Californian cities.

	Percentage Change from Baseline			
	Income	Rents	Utility	$PM_{2.5}$ Exposure
I. Education				
College Education	0.5	-0.7	0.9	0.3
Less Than College	0.0	-0.9	0.6	0.3
II. Family Size				
Single	0.2	-0.8	0.7	0.4
Married w/o Children	0.2	-0.7	0.6	0.2
Married w/ Children	0.1	-1.0	0.3	0.1
III. Race				
White	0.2	-0.8	0.6	0.2
Nonwhite	0.2	-0.9	0.8	0.4

Table 7. Changes in average income, rents, utility, and pollution exposure by demographic group.

Distributional Effects In Table 7, we explore the distributional implications of the relaxation of land use regulations in California. The four columns give the percentage change in average income, rents, utility, and exposure to $PM_{2.5}$ for different demographic groups. Utility is measured in log dollar equivalents.[67]

[67]Given that the idiosyncratic preference draws are distributed as Extreme-Value Type 1, household i's expected utility is given by $\log\left(\sum_{j' \in J} \exp\left(\bar{V}_{ij'}/\sigma_d\right)\right)$ plus a constant. To translate this into log income equivalent, we first divide expected utility by $\boldsymbol{\alpha}_{jd}$ in each city j. Note that $\frac{\log\left(\sum_{j' \in J} \exp\left(\bar{V}_{ij'}/\sigma_d\right)\right)}{\boldsymbol{\alpha}_{jd}}$ gives expected utility measured in log income equivalent for a household who

Highly educated households benefit more than less educated households because the income premium in cities is generally larger for educated workers (Baum-Snow & Pavan, 2013). Further, single households experience larger utility gains than married households, as single households are more mobile and therefore better able to benefit from the drop in rents in Californian cities. Finally, all demographic groups see a small increase in their average $PM_{2.5}$ exposure as they increase their concentration in large cities and in particular, the relatively polluted cities in southern California.

Local Pollutants Figure 14 plots the changes in particulate matter concentrations by CBSA when land-use restrictions in California are relaxed. Specifically, the x-axis gives the change in $PM_{2.5}$ concentration when we relax land-use regulations in California compared to the baseline. The y-axis gives the number of cities that fall into a given range of changes. The colors indicate the four Census regions.

Similar to carbon emissions, there is a national reduction in $PM_{2.5}$ concentration from relaxing land-use regulations in California. The mechanism is quite similar to carbon emissions. Households are induced to live in California where they use less electricity—due to California's temperate climate—and the power plants they use are less $PM_{2.5}$-intensive. This leads to a reduction in $PM_{2.5}$ in most CBSAs.

However, unlike CO_2, the spatial distribution of $PM_{2.5}$ emissions is important as $PM_{2.5}$ is a local, and not global pollutant. When households move

lives in city j—an increase in 0.01 in this object for example, provides a change in expected utility equivalent to a 1% increase in income for household who lives in city j. We then take the average across cities weighted by the household's choice probabilities in the baseline counterfactual.

to California, this increases electricity demand in the WECC NERC region—which overlaps closely with the Western Census region. This leads to an increase in the level of $PM_{2.5}$ emissions in WECC and therefore a slight increase in $PM_{2.5}$ concentrations in cities in the Western region. In all other regions, CBSAs experience decreases in $PM_{2.5}$. This is because energy demand falls in these regions as households move away, and the offset in $PM_{2.5}$ from local energy demand is greater than the increase in $PM_{2.5}$ from far away sources (namely, California). This decreases both emissions of $PM_{2.5}$ and concentration of $PM_{2.5}$ in these regions.

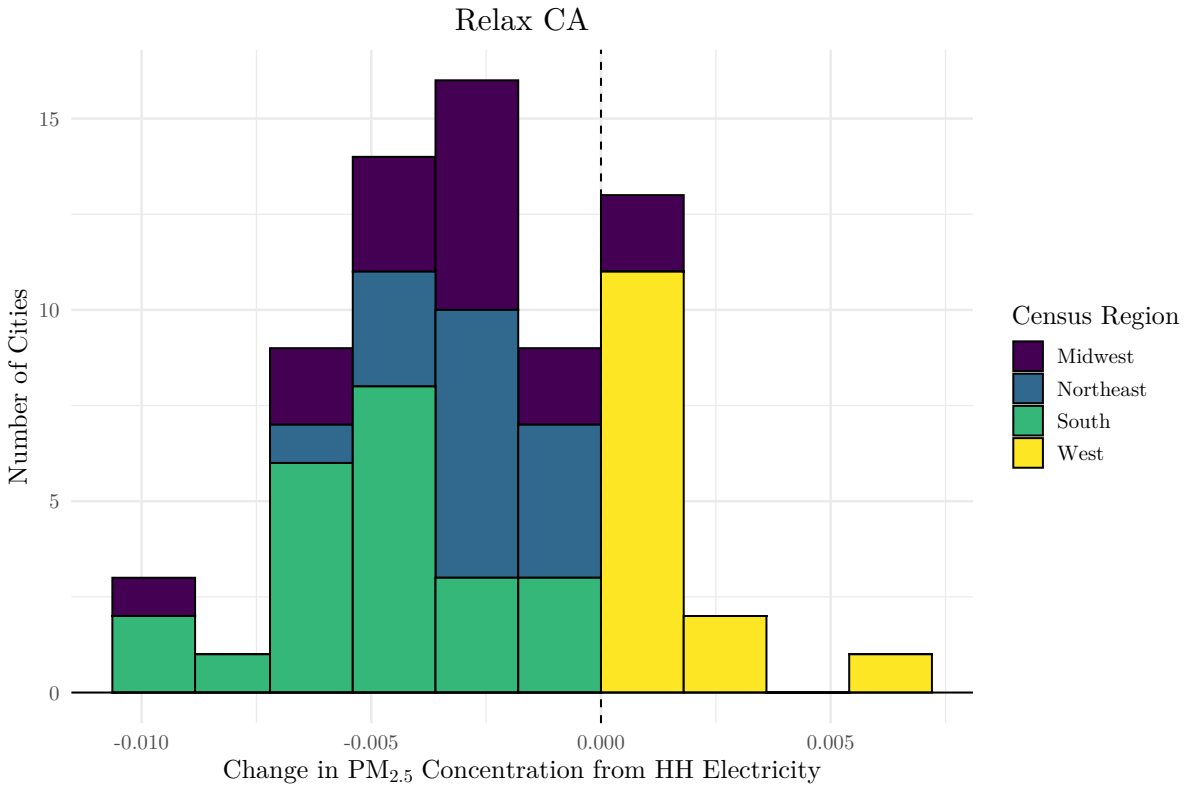


Figure 14. Histogram of CBSA level differences in particulate matter concentration from electricity relative when land-use restrictions in California are relaxed relative to the baseline.

3.6.2 Removing the Correlation Between Land Use Restrictions and Emissions.

The negative correlation between land use restrictions and city

level emissions has important implications for national carbon emissions. To further explore the implications of land use restrictions on carbon output, we simulate setting land use restrictions to the level faced by the median urban household in all cities.

The results are displayed in the third column of Table 5.[68] The results in panel I indicate that changing land use restriction in all cities leads to a dramatic relocation from the South and Midwest to the West and Northeast. Specifically, the population in the Northeast region increases from 18% of the total population to 43% while the population in the Midwest and the South decrease by roughly one third to one half.

Panel II and III show usage and emissions from each energy type. Demand for natural gas and fuel oil are high while demand for electricity is low in the cold Northeast. As a result, natural gas usage increases by 0.7%, while electricity usage decreases by 7.8%. As a result of this decrease in electricity usage and relocation towards cities with more efficient power plants, emissions from electricity decrease by 18.9%. Overall, this leads to a 3.5% decrease in national carbon output and over an 8.5% increase in the national carbon efficiency of output. This implies a drop in the social cost of carbon of $1.7 billion annually.

3.7 Robustness and Extensions

3.7.1 Sensitivity to Alternative Parameters. In this section, we examine the robustness of our main results to alternative values of key parameters. In particular, we recalculate the reduction in national carbon output resulting from the relaxation of land use restrictions in California cities for a range of parameter values. First, we examine the model's sensitivity to the scale parameter of the

[68] In Appendix B.0.2.3, we show how the distribution of local pollutants change in this counterfactual.

idiosyncratic preference draw, σ_d. Lower values of σ_d imply that household location choice is more elastic with respect to wages and rents.[69] Therefore, households will be more likely to change their location decisions in response to changes in land use restrictions.

The percentage reduction in carbon output relative to the baseline for a range of values of σ_d for single and married households is shown in Panel (a) of Figure 15. Recall in our baseline specification that we estimate $\sigma_d = 0.17$ for single households, $\sigma_d = 0.21$ for married households without children, and $\sigma_d = 0.40$ for married households with children and we found a decrease in carbon emissions of 0.6%. In the figure, σ_d for single households is displayed on the vertical axis and the average of σ_d for married households is displayed on the horizontal axis. We vary σ_d for married households such that the ratio of σ_d for married households with children compared to married households without children is held constant at the baseline level. Darker colors imply smaller changes in carbon emissions while lighter colors imply larger changes. Figure 15 illustrates that the change in carbon emissions is decreasing in σ_d for both single and married households. In the extreme case when $\sigma_d = 0.1$ for both types of households, households are very responsive to changes in rents. As a result, carbon emissions drop by 0.8% when we relax land use restrictions in California. When $\sigma_d = 0.8$ for both types of households, carbon emissions drop by roughly 0.3%.

Next, we examine the model's sensitivity to the budget share of housing parameter, α_d^H. Recall that we estimated $\alpha_d^H = 1.49$ for single households, $\alpha_d^H = 1.44$ for married households without children, and $\alpha_d^H = 1.92$ for married

[69] For each counterfactual in which we change σ_d, we recalculate the amenity values ξ_{jd} such as to keep the mean utility of each demographic group in each city equal to its baseline level. Therefore, the distribution of households of each demographic group given the baseline levels of land use restrictions will be equal to the baseline distribution with the original values of σ_d.

households with children. Higher values of α_d^H imply households spend a larger fraction of their income on housing and therefore will be more sensitive to housing prices in their choice of where to live. The results are displayed in Panel (b) of Figure 15. The vertical axis shows values of α_d^H for single and the horizontal axis shows α_d^H for married households. We change α_d^H for married households such that the ratio of the parameter for married households with children to married households without children is held at the baseline level. Larger values of α_d^H of both types of households imply larger decreases in carbon emissions. When $\alpha_d^H = 0.4$ for both types of households, carbon emissions drop by 0.3% when we relax land use restrictions in California. If we set $\alpha_d^H = 3.2$ for both types of households, carbon emissions drop by roughly 0.8%.

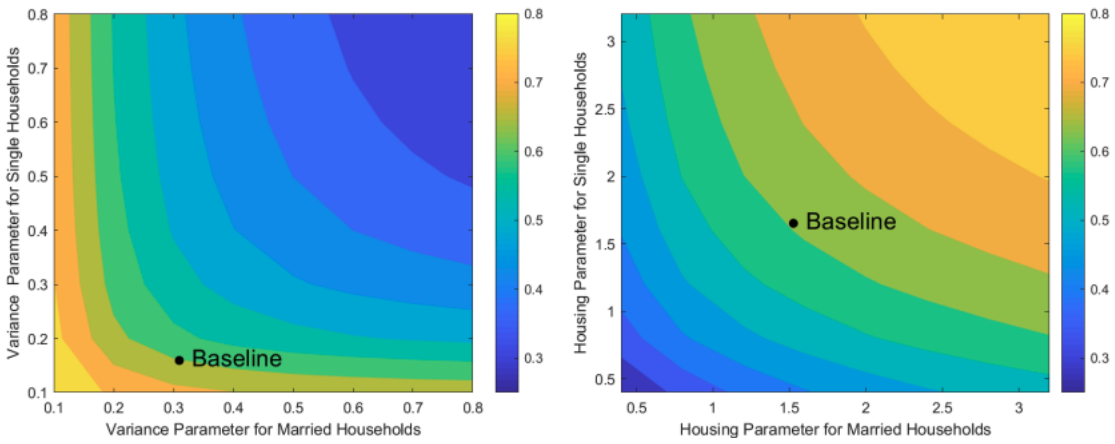


Figure 15. Percentage change in national emissions from relaxing land use restrictions in California for various parameter values. In panel (a), we display σ_d for single on the Y (vertical) axis, and σ_d for married households on the X (horizontal) axis. Panel (b) shows the percent reduction in carbon emissions as a function of α_d^H for both types of households.

3.7.2 Endogenous Electricity Pricing. In our baseline specification, we assume that electricity is produced at a constant marginal cost; therefore, the supply curve of electricity is perfectly elastic. In this section, we consider an extension in which the price of electricity is determined endogenously.

Specifically, we assume that electricity producers in each NERC region form an upward sloping long-run electricity supply curve, reflecting differences in costs or productivity of potential electricity production opportunities within a region. For low quantities, electricity can be produced at a low cost. As electricity production increases, increasingly less productive resources must be utilized, which therefore implies higher costs of production.

A number of papers examine the short run supply curves of electricity. The short run electricity supply curve is often modeled as a "dispatch curve" with constant or linear marginal costs, to reflect the unique way in which electricity is allocated in the very short term.[70] Essentially, electricity generators are ranked in terms of their marginal cost of producing electricity. As demand increases, plants are dispatched to produce power in increasing order of marginal cost. However, this type of modeling approach is likely not a good representation of the long run energy supply curve which we consider here. In the long run, energy producers may respond to changes in energy demand by opening new reactors and new plants. Therefore, we posit a more parsimonious long run electricity supply curve as:

$$C_{\mathcal{R}} = v_{\mathcal{R}} X_{\mathcal{R}}^{\kappa}$$

where $X_{\mathcal{R}}$ is the total quantity of energy produced in region $\mathcal{R}$, κ is a parameter equal to the inverse elasticity of the energy supply curve, and $v_{\mathcal{R}}$ is a region specific cost shifter.

Electricity is then transmitted to a specific local labor market at an additive transmission cost, ϕ_j.[71] Given the assumption of perfectly competitive generation companies, we can write the inverse energy supply curve to city j as

[70] For an example, see Ma, Sun, and Cheung (1999).

[71] This cost may directly reflect costs of transmissions or network congestion costs.

$$P_j^{\text{elec}} = b_j + \kappa \log\left(X_{\mathcal{R}(j)}\right)$$

where $b_j = \phi_j + \log\left(v_{\mathcal{R}(j)}\right)$.

To calibrate this model extension, we first calibrate the inverse elasticity of the electricity supply curve as $\kappa = \frac{1}{1.27}$, based on the estimates in C. Dahl and Duggan (1996). We then choose the parameters b_j to match state level electricity prices.

The main counterfactual results with endogenous electricity pricing are summarized in Appendix B.0.2.6, Table B8. Overall, the population distributions across all counterfactuals are quite similar to the counterfactuals with perfectly elastic electricity supply. Households spend a relatively small fraction of their income on electricity and therefore changes in electricity prices have little impact on their location choices. Natural gas and fuel oil emissions are also nearly identical to the case with a perfectly elastic electricity supply. However, the reductions in electricity usage and therefore overall carbon emissions are smaller in the case of endogenous electricity prices. Overall this leads to a 0.4% reduction in carbon emissions from the relaxation of land use restrictions in California.

3.7.3 Local Pollutants in Utility Function. In this section, we consider an extension of our model in which local pollutants enter the utility function. It is worth emphasizing that only local pollutants produced by household electricity usage are endogenous in our model, while pollutants produced by all other sources are held constant. A richer model would also include how the distribution of households in space would lead to an increase in pollution from changes in the spatial distribution of manufacturing firms, for example.

With this caveat in mind, let the households' utility function be given by:

$$u_{ij} = \alpha_d^c \log c + \alpha_d^H \log H + \sum_m \alpha_{jd}^m \log \hat{E}^m + \alpha^{PM} \log PM_j + \lambda_{ij}$$

where PM_j is the concentration of $PM_{2.5}$ is city j. We set $\frac{\alpha^{PM}}{\sigma_d} = -.255$ for all demographics groups based on the estimates from Bayer, Keohane, and Timmins (2009).[72] We recalculate the unobserved amenity parameters, ξ_{jd}, such as to keep the mean utility of each demographic group in each city equal to its baseline level. All other parameters are kept at their baseline levels. Mechanically, the inclusion of $PM_{2.5}$ in the utility function will lead to higher estimated values of ξ_{jd}, so the population shares in the model match the data shares.

The results are displayed in Appendix B.0.2.6, Table B9. The results are very similar to the baseline simulations. This is expected, given that $PM_{2.5}$ emissions from power plants only constitute a small fraction of total $PM_{2.5}$ emissions in a given city.[73]

3.7.4 Power Plant Substitution. One potential issue with our counterfactuals is that new power plants built in order to accommodate increases in demand for electricity may be cleaner or dirtier than the current stock of power plants in that region. Therefore, the carbon emissions factors we use in our analysis will change in response to increases in electricity demand. For example, our main counterfactual of the relaxation of land use restrictions in California led to a substantial increase in population and energy usage in California. As a result, new power plants may be constructed in the corresponding WECC NERC region which may be cleaner or dirtier than the current power plants in the region. If these new

[72] We use the instrumental variables results with no additional controls (Table 5, Column 3). The results with other IV estimates are similar.

[73] We display the contribution of household electricity to total $PM_{2.5}$ concentration across various cities in Appendix B.0.2.3.

power plants are cleaner (dirtier) than the current stock of power plants, we will underestimate (overestimate) the reduction in carbon emissions.

To investigate how endogenous changes in the composition of power plants might affect our results, we compare power plants built before and after 2000. We find power plants built after 2000 emit considerably less CO_2 per MWh than plants built prior. Specifically, for the WECC NERC region, we find that power plants built prior to 2000 emit 858 lbs of CO_2 per MWh of electricity, whereas plants built after 2000 emit only 597 lbs of CO_2 per MWh.[74] These results suggest that if new power plants were built in response to increases in California's population, these new plants would be more carbon-efficient than the current stock of plants.

3.8 Conclusion

Household carbon emissions vary considerably across cities. Land use restrictions, which are set by local governments, tend to be tighter in cities with low carbon emissions and therefore encourage households to live in cities with less moderate climates and higher greenhouse gas-emitting power plants.

We began by following Glaeser and Kahn (2010) and documented large spatial variation in both the carbon efficiency of power plants and energy consumption. Cities with more temperate climates (such as San Francisco) tend to emit substantially less carbon than other cities. Furthermore, these cities also tend to have very tight land use restrictions. To examine the effects of land use restrictions on national carbon emissions, we then estimated a model of household sorting, energy demand, and locations that vary by power plant technology. We found that the relaxation of land use restrictions in California leads to a decrease in national carbon output of 0.6% and a decrease in the social cost of carbon of

[74] Section B.0.2.5 in the data appendix has further information on the full distribution of emissions from power plants built before and after 2000.

$310 million annually. Our main conclusion is that the positive correlation between tighter land use restrictions and the greenness of cities has large implications for national carbon output.

Relaxing land use restrictions would likely impact the spatial distribution of firms in addition to households. In our model, firms only use labor in production. In reality, many firms use the same energy inputs as households – such as electricity and natural gas. Additionally, firms also use land in production. Since local factor prices (such as land) are first-order to firm location decisions (Suárez Serrato & Zidar, 2016), relaxing land-use restrictions in California would lead to firms sorting into California to take advantage of the lower land prices. Furthermore, when energy is an input to production and more production shifts towards California—where the electricity is carbon-efficient—carbon emissions would fall. Future research could incorporate firm sorting and energy demand into our framework to estimate the effects of land use regulations on industrial carbon emissions.

Additional work could use our model to analyze the spatial implications of the Clean Air Act. The model could also be used to analyze changes in carbon emissions as a result of improved energy infrastructure and therefore easier electricity transmissions across regions. Future research could extend the model to analyze the effects of improving insulation or policies that change the composition of power plants – such as renewable energy subsidies.

CHAPTER IV

CARBON TAXES IN SPATIAL EQUILIBRIUM

4.1 Introduction

Carbon emissions create well-recognized negative externalities. The Intergovernmental Panel on Climate Change (IPCC) inventories rising sea levels, temperatures, and changes in the pattern, frequency, and intensity of extreme weather events (storms, heat waves, and droughts) as a few of the likely consequences of climate change. Among professional economists, carbon taxes receive widespread support as a tool for reducing emissions due to the economic efficiency offered by the tax (Climate Leadership Council, 2021). Despite this support from economists, global policy efforts to implement carbon prices have been fairly limited.[1] Sallee (2019) argues that the distributional concerns arising from carbon pricing—and more specifically, the ability to precisely predict lump-sum transfers for those who bear the greatest burden of the tax—account for disparities between the policy preferences of economists and voters.

Several factors lead to heterogeneity across households in the burden of carbon taxes. Carbon is a byproduct of energy production, so differences in the carbon intensities of power plants across regions will lead to spatial differences in the effects of a carbon tax on electricity prices. These increases in electricity prices (and other fuel prices)—and their differences across space—create heterogeneous initial impacts on households through two distinct channels: labor demand and household energy expenditures. On the labor-demand side, a carbon tax will reduce output from emissions-intensive industries and lead to a reallocation of input demand away from energy and towards other inputs. Furthermore, industries tend to cluster in particular cities; this clustering implies that a carbon tax

[1] As of 2020, some 40 countries had enacted some form of carbon pricing (World Bank, 2020).

will have differential impacts on labor demand (and hence wages) across both cities and sectors. On the energy expenditure side, household demand for energy (derived from the demand for heating and cooling) varies across space, primarily due to differences in climate.[2] Thus, a nationally uniform carbon price creates heterogeneous impacts on households stemming from differences across cities in the energy price increases and the disutility delivered by these price increases as a result of putting a price on carbon. The labor demand channel and the household energy expenditure channel imply that the burden of a carbon tax is a function of the joint spatial distributions of sectors (and their respective production technologies), households, and the mix of fuels used regionally to generate electricity. This paper contributes to the carbon-pricing policy debate by estimating the spatial and sectoral distribution of incidence of a carbon tax across education groups in the United States.

To measure the burden of a carbon tax, I develop and estimate a quantitative spatial equilibrium model. At the core of the model is a discrete-choice problem for households, where households must select both a sector and a location. Wages are endogenous, and production is a city- and sector-specific function of imperfectly substitutable labor and energy inputs. Locations vary in terms of their amenities, their housing supply curves, the emissions intensity of their electricity generation and power-plant technologies, and their input-use intensities in producing goods and services. Output markets are assumed to be perfectly competitive.

Within the model, the welfare effects of a national carbon tax will vary geographically *and* sectorally for four reasons. First, locations vary in terms of

[2] Lyubich (2022) demonstrates that location explains over half of the variation in household carbon emissions—15-25% of the overall variation in carbon emissions.

the marginal benefit of energy consumption and the carbon content of the fuel mix used by local power plants (Glaeser & Kahn, 2010). Cities with warmer climates, such as Houston, tend to have a higher marginal benefit of electricity consumption due to greater demand for air-conditioning, resulting in higher household energy demand and thus greater carbon emissions. Second, within the U.S., the carbon intensity of power plants varies significantly across regions; some areas, such as the Midwest and South, are much more dependent on coal-fired electricity than other areas, such as the West.[3] Third, sectors vary in the amount of energy they use for production processes and the degree to which labor and energy are technologically substitutable. For example, the manufacturing sector is significantly more energy-intensive than the services sector, so a carbon tax will have a considerably larger effect on manufacturing wages and employment than for services. Lastly, in the context of the model, cities vary in their industrial employment composition.
Due to the differences in the fuel mix used by regional power plants, a carbon tax will differentially impact energy prices across cities, which in turn will have heterogeneous impacts on worker's wages. In response to the energy price and wage changes, workers can relocate across cities and sectors as the relative attractiveness of each city-sector changes. This re-sorting of workers across space and sectors results in equilibrium adjustments in rental markets that further affect utility. Therefore, accurate recovery of the overall tax incidence across different types of households necessitates a general equilibrium analysis that permits households and firms to respond in a variety of ways to the changes induced by the carbon tax.

I discipline the model's structural parameters using a variety of publicly available data. I use data from the American Community Survey (ACS) and Census to estimate household preference parameters and household carbon

[3]Figure C2 in Section C.0.5 provides a map of US regional emissions rates from electricity.

emissions. More specifically, I estimate the model's labor supply parameters using a two-step procedure involving maximum likelihood and instrumental variables, similar to the approach taken by Berry et al. (2004). I use a combination of calibration and estimation to obtain production functions that are both city-specific and sector-specific.

To monetize the incidence of a carbon tax, I calculate the average compensating variation, which is the average dollar amount of additional income that would be required for a household in each city-sector to maintain utility levels in spite of the carbon tax. I find significant unequal distributional consequences across cities and sectors from a uniform carbon tax. I find that a carbon tax of $31 per ton would result in a mean decrease of 926 vs. 1,417 utility-equivalent dollars per year for college vs. non-college households, respectively – with significant heterogeneity across cities and sectors. Furthermore, I find that this carbon tax would reduce emissions by 19.8% and lead to a reallocation of workers away from manufacturing (by 11.1%) and into less-carbon intensive jobs in the services sector. Furthermore, to differences in wages across cities, the compensating variation measure (monetized incidence) masks important underlying heterogeneity in the carbon tax burden. Thus, I separately examine compensating variation in *percentage* terms rather than dollars— which I refer to simply as "incidence," (as opposed to *monetized* incidence or compensating variation). I find that cities on the West Coast and New England experience lower incidence than cities elsewhere due to their relatively carbon-efficient power plants and services-oriented economies.

The model predicts migration patterns across cities consistent with the spatial variation in tax incidence. Generally, cities in the Western and New England Census Divisions experience population increases, while cities in the South and Midwest experience population decreases. For example, California

experiences an overall population increase of roughly 2%. Cities in California have mild climates, source their electricity from carbon-efficient power plants, and have services-oriented economies. Thus, a carbon tax induces migration to California, as households benefit from California's economy being relatively resilient to climate policy. Despite being less mobile, non-college-educated households move at greater rates in response to the carbon tax—underscoring the more-significant tax burden that these households bear.

Motivated by the remarkable decline in the share of electricity generated from coal over the last 15 years, I use the model to decompose the incidence of a carbon tax into two distinct components: coal and non-coal.[4] I re-calculate regional emissions rates from electricity generation in the absence of coal and re-simulate the $31 carbon tax in the context of this hypothetical alternative (and cleaner) grid.[5] I find that the compensating variation falls by nearly 40%, and this decline exhibits regional heterogeneity. Coal-generated electricity is highly carbon-intensive and thus is responsible for a larger share of the total incidence in coal-dependent regions. This finding suggests that as the electricity grid decarbonizes, the payment required to compensate households for a carbon tax will decline by a significant amount.

I use my model to simulate a carbon tax with various compensation schemes. Due to the regressive nature of flat carbon taxes, some proposed carbon-pricing legislation includes progressive transfer payments.[6] In simulations with

[4]The decline in coal's share is well documented— for example, see Mendelevitch, Hauenstein, and Holz (2019). I summarize the national and regional decline in the share of coal-generated electricity in Figure 18.

[5]To be clear, this exercise is purely decompositional. Carbon pricing policies impact the rate of adoption for carbon-efficient electricity generation since coal is carbon-intensive. For an example, see Scott (2021a).

[6]For example, the recently introduced Stemming Warming and Augmenting Pay (SWAP) act specifically calls for "...20% (of revenues) to establish a carbon trust fund for block grants to

carbon taxes and transfers, I focus on varying the progressivity of the payments to households. The simulations reveal a novel relationship between the progressivity of transfers and *aggregate* carbon emissions: aggregate emissions are lower with progressive transfers than with lump-sum transfers. The mechanism behind this relationship is straightforward. Wages exhibit spatial and sectoral correlation, so progressive transfers are also correlated across space and sectors. In an equilibrium with progressive transfers, lower-wage cities receive larger transfers and thus attract more workers (relative to a policy with lump-sum transfers), all else equal. The progressivity of the transfers will impact aggregate carbon emissions if wages are correlated with carbon emissions (at the city-sector level) due to the reallocation of workers into areas with higher transfers.

Indeed, there is a positive correlation between wages and emissions in the data, and the model predicts a larger share of workers shifting away from jobs with high wages in emissions-intensive sectors, thus causing aggregate emissions to fall. My results suggest that using carbon tax revenue for income redistribution may have the unappreciated additional benefit of reducing aggregate emissions—and therefore, has the potential to enhance the effectiveness of the policy in meeting emissions targets.

Literature. I am not the first to recognize—or to model— the distributional impacts of carbon taxation. For example, Rausch, Metcalf, and Reilly (2011) (RMR) use a calibrated version of MIT's US Regional Energy Policy model and a sample of roughly 15,000 households to demonstrate considerable heterogeneity

offset higher energy costs for low-income households, climate adaptation, energy efficiency, carbon sequestration, and research and development programs" (U.S. House of Representatives, 2019).

across demographic groups in the incidence of a carbon tax.[7][8] Hafstead and Williams (2018) use a two-sector general equilibrium model and conclude that the unemployment effects of a carbon tax will be negligible due to growth in clean industries. Using a macroeconomic lifecycle model, Fried, Novan, and Peterman (2021) (FNP) find that, in the welfare-maximizing allocation of carbon tax revenue, two-thirds goes to a reduction in capital-income taxes and one third towards increasing the progressivity of the income tax. My results are complementary to those of FNP. While I abstract from dynamics, my model's geographic and sectoral heterogeneity allows me to capture the relationship between income redistribution and aggregate emissions. In a recent working paper, Castellanos and Heutel (2019) (CH) conclude that alternative assumptions about worker mobility potentially play a significant role in the aggregate employment effects from carbon pricing. RMR and CH make different assumptions about labor mobility; in RMR, workers are mobile across sectors but not locations. CH examines edge cases with perfect mobility and perfect immobility. My work explicitly models (and then estimates) the process by which households make city-sector choices—including moving costs. Even if employment remains constant, my work accommodates the fact that relocating to new jobs and locations is costly.

An extensive empirical literature has demonstrated that environmental regulation has heterogeneous impacts across sectors. Recent work by Yamazaki (2017) finds that more energy-intensive sectors saw larger relative losses in wages

[7]Other examples of research that examines the distributional consequences of carbon pricing includes Goulder, Hafstead, Kim, and Long (2019), Williams III, Gordon, Burtraw, Carbone, and Morgenstern (2015), and Beck, Rivers, Wigle, and Yonezawa (2015). To my knowledge, I am the first to estimate the heterogeneous welfare effects from a carbon tax specifically in a locational discrete-choice setting.

[8]For details on the Regional Energy Policy model, see Yuan, Rausch, Caron, Paltsev, and Reilly (2019).

from the carbon tax in British Columbia, Canada. In British Columbia, overall unemployment decreased due to a shift in demand towards less energy-intensive sectors due to the province's revenue recycling decisions (i.e., transfers that households received from the tax).[9] In complementary work, Yip (2018) finds that the British Columbia carbon tax was implemented mainly at the expense of individuals without a college degree, as these workers are generally in more energy-intensive sectors with fewer outside options.[10] Other research has demonstrated broader distributional effects from environmental regulation. For example, Walker (2013) examines the labor-market impacts of the 1990 amendments to the Clean Air Act (CAA). These amendments involved command-and-control regulations that established thresholds for the maximum allowable ambient concentrations of pollutants. Using a triple-difference approach, Walker finds significant reductions in employment in manufacturing (and other energy-intensive sectors) due to these regulations. Curtis (2014) examines the labor market impacts of the EPA's NO_x budget trading program—which was a type of cap and trade program enacted in 2003—and reaches conclusions similar to Walker (2013). Specifically, sectors that are more energy-intensive experience larger losses in unemployment.

Methodologically, the present paper is closely related to Diamond (2016), Piyapromdee (2021), Colas and Hutchinson (2021), and Colas and Morehouse (2022) (CM). These papers estimate general equilibrium models of location choice in which wages and rents respond endogenously to agents' location choices. The model in CM is designed to measure changes in *residential* carbon emissions from

[9]The tax was supposed to be revenue-neutral. In practice, it did not turn out this way. Yamazaki (2017) notes "tax credits have been exceeding tax revenues since its implementation."

[10]Other recent empirical work has examined the effects of carbon taxes implemented in different countries, such as Martin, De Preux, and Wagner (2014). These authors find minimal adverse unemployment impacts from the UK's carbon tax.

the relaxation of stringent land-use regulations. The present paper departs from the model in CM along three significant dimensions. First, I model energy as a production input for firms. This allows for endogenous wage changes in response to energy price changes that ensue from a carbon tax. Second, firms vary across cities and sectors—with different factor intensities and productivities. The sectoral composition of industries at a given location will impact the welfare of local workers (for reasons described above) and thus will have first-order consequences for household sorting. Lastly, households make a *joint* choice concerning both their city and their sector. To assess the welfare effects of a carbon tax, my model combines empirical insights from a large literature on the distributional consequences of carbon pricing and extends the modeling strategy used in Colas and Morehouse (2022). I build a unified framework that simultaneously analyzes the variation in labor-market responses to carbon taxes across both cities and sectors.

Overall, my objective in this paper is to improve policy-makers' understanding of the distributional consequences of carbon taxes. The empirical structural model I develop allows me to obtain quantitative estimates of the varying incidence levels across cities and sectors. The rest of this paper proceeds as follows. Section 4.2 details the structural model, Section 4.3 provides an overview of the data, Section 4.4 discusses the estimation procedure and how key parameters are identified, Section 4.4.3 discusses the parameter estimates, section 4.5 examines selected counterfactuals, and Section 4.6 concludes.

4.2 Model

This section describes a general equilibrium model where households make a joint, discrete choice over location and sector. Conditional on location and sector, households consume a numeraire good, housing, and energy services. These

different decisions, and the linkages between these related markets, are crucial for determining the net welfare effects of a carbon tax. Labor demand arises from perfectly competitive firms that differ across locations and sectors. Housing demand follows from the location choices of agents, and housing supply is increasing in the price of housing.

More specifically, a household has an exogenously given education level and birthplace. Locations vary in terms of the location-specific consumption goods (amenities) they provide, wages, rents, the marginal utility of energy consumption, and the carbon intensity of the local power plants. Sectors vary in input use intensity and unobservable sectoral amenities (e.g., non-wage employment benefits). Firms with nested CES technologies combine college-educated and non-college-educated labor with electricity and gas to produce outputs within each sector at each location. Each location functions as a small open economy, so wider national and international markets always satisfy any excess demand.

To capture the labor-market effects of a carbon tax, I introduce heterogeneity across locations and sectors of production for the output good. Across locations and sectors, firms vary by all input-use intensities (a set of parameters). Variation in production parameters reflects differences across locations in available labor supply and variation in regional energy prices. Furthermore, across sectors (but not locations), firms vary by their elasticities of substitution across different types of energy inputs and the mix of energy and labor. Some sectors, such as agriculture, can substitute relatively easily between labor and energy. Other sectors, such as construction, cannot.

In my model, a uniform national carbon tax will have location- and sector-specific welfare effects for several reasons. First, locations vary in their marginal utilities of energy consumption (*e.g.* they have different climates), and their power

plants have different carbon intensities. Consequently, the consumption of the same amount of electricity will imply different levels of emissions across locations. Second, a carbon tax will have varying effects on wages. Consumption of energy in production leads to emissions. The specific amount of energy used and carbon emitted depends on the firm's sector and location. After the carbon tax, wages will change via two particular channels. First, as a direct result of the carbon tax, the relative price of energy inputs will increase, so the firm will substitute other inputs and reduce output. Second, there will be adjustments in equilibrium wages to the extent carbon taxes have are differential impacts on labor demand across locations and sectors. Thus, workers in regions/sectors with higher carbon emissions will see a relatively large decrease in wages and will sort towards sectors and locations with lower carbon intensities.

4.2.1 Households. A household consists of one or more individuals. If a household contains more than one working-aged individual, the "agent" refers to the putative household head. Agents are endowed with an education level and native birth state. They receive utility from the consumption of a numeraire good, housing, energy services, and amenities. Each agent makes a one-time decision over locations and employment sectors. Let $j \in J$ index cities, $n \in N$ index sectors, and $e \in \{l, c\}$ index education groups (where l indicates that the household has "less than a college degree" and c indicates that the agent has a 'college degree or greater"). Agent i's utility from living in city j and working in sector n is characterized as:

$$u_i(c, h, x_m | e, j, n) = \alpha_e^c \log c + \alpha_e^H \log h + \sum_m \alpha_{ejn}^m \log x_m + \lambda_{ijn}, \tag{4.1}$$

where c is consumption of the numeraire good, h is consumption of housing, x_m is consumption of fuel type $m \in \{\text{elec}, \text{gas}, \text{oil}\}$ and λ_{ijn} is amenities. I parameterize

amenities as

$$\lambda_{ijn} = f(j, \mathcal{B}_i) + \xi_{ejn} + \sigma_e \epsilon_{ijn}, \tag{4.2}$$

where the function $f(j, \mathcal{B}_i)$ is a function of the agent's birth location, $\mathcal{B}_i$, and location j, ξ_{ejn} is an unobserved component of amenities that all workers share within an education group-city-sector. The term ϵ_{ijn} is an idiosyncratic preference shock drawn from a Type I Extreme Value distribution (EV1) with mean zero and shape parameter σ_e. Variation in ξ_{ejn} captures differences in amenities across locations and sectors. Heterogeneity in ξ_{ejn} across cities (but within the same sector) is driven by heterogeneity in location-specific market or non-market consumption goods such as air quality, crime, schools, or the number of restaurants in the city. Variation in ξ_{ejn} within the same city but across sectors is due to variation in the availability of non-pecuniary benefits across sectors. I parameterize f as:

$$f(j, \mathcal{B}_i) = \gamma_e^{div} \mathbb{I}\left(j \in \mathcal{B}_i^{div}\right) + \gamma_e^{\text{dist}} \phi\left(j, \mathcal{B}_i^{st}\right) + \gamma_e^{\text{dist2}} \phi^2\left(j, \mathcal{B}_i^{st}\right), \tag{4.3}$$

where $\mathbb{I}\left(j \in \mathcal{B}^{div}\right)$ is an indicator for j being in worker i's birth division, $\phi(j, \mathcal{B}_i^{st})$ is the Euclidean distance between location j and the agent's birth state $\mathcal{B}_i^{st}$, and $\phi^2(j, \mathcal{B}_i^{st})$ is the squared Euclidean distance between j and $\mathcal{B}_i^{st}$.[11] As noted by Bayer et al. (2009), individuals tend to have a preference for locations with greater accessibility to their birth state, where I model accessibility simply as distance.

Agents face the following budget constraint

$$w_{ejn} = c + R_j H + \sum_m P_j^m x_m, \tag{4.4}$$

[11] This specification is slightly unusual in that I use an indicator for the individuals' *census division* of birth and not their *state* of birth. The model includes only 70 CBSAs (plus 9 census divisions as outside options). Not every state is represented, but all census divisions are.

where w_{ejn} is the wage level for an agent of education level e in city j and sector n. R_j and P_j^m represent the rental and energy prices (of each type m) in city j, which are constant across sectors and demographic groups.

Maximizing the utility described by equation (4.1) subject to the budget constraint represented by equation (4.4) yields constant shares of income devoted to housing and fuel consumption:

$$\begin{aligned} H_{ejn}^{\star} &= \frac{\alpha_e^H w_{ejn}}{\boldsymbol{\alpha}_{ejn} R_j} \\ x_{ejn}^{m\star} &= \frac{\alpha_{ejn}^m w_{ejn}}{\boldsymbol{\alpha}_{ejn} P_j^m} \qquad \forall m \in \{\text{elec}, \text{gas}, \text{oil}\}, \end{aligned}$$

where to simplify the notation in what follows, I define the parameter $\boldsymbol{\alpha}_{ejn}$ as:

$$\boldsymbol{\alpha}_{ejn} = \alpha_e^c + \alpha_e^H + \sum_m \alpha_{ejn}^m.$$

I then solve for the agent's constrained utility maximization problem to yield the corresponding indirect utility function associated with location j and sector n

$$v_{ijn} = (\boldsymbol{\alpha}_{ejn}) \log(w_{ejn}) - \alpha_e^H \log R_j - \sum_m \alpha_{ejn}^m \log P_j^m + f(j, \mathcal{B}_i) + \hat{\lambda}_{ijn}, \tag{4.5}$$

where I again simplify the notation by defining:

$$\hat{\lambda}_{ijn} = \lambda_{ijn} + \sum_m \alpha_{ejn}^m \log\left(\alpha_{ejn}^m\right).$$

Given the EV1 assumption for the idiosyncratic preference shock (ϵ_{ijn}) the probability that agent i with education level e chooses option jn is given by the familiar conditional logit form:

$$P_{ijn}^e = \frac{\exp(\frac{\bar{v}_{ijn}}{\sigma_e})}{\sum_{j'} \sum_{n'} \exp(\frac{\bar{v}_{ij'n'}}{\sigma_e})}, \tag{4.6}$$

where $\bar{v}_{ijn} = v_{ijn} - \sigma_e \epsilon_{ijn}$.

4.2.2 Firms & Housing Supply. Firms competing in perfectly competitive factor markets combine both labor and energy inputs to produce a sector-specific, tradeable good. I model each location as a small, open economy, so firms treat their output price (denoted by P_n) as exogenous. Firms in city j and sector n produce according to:[12]

$$Y_{jn} = A_{jn} K_{jn}^{\eta} \mathcal{I}_{jn}^{1-\eta},$$

where $\mathcal{I}_{jn}$ is a conventional CES aggregator for energy and labor inputs which is specific to that location and sector, jn:

$$\mathcal{I}_{jn} = \left(\alpha_{jn} \mathcal{E}_{jn}^{\rho_{el}^n} + (1 - \alpha_{jn}) \mathcal{L}_{jn}^{\rho_{el}^n} \right)^{\frac{1}{\rho_{el}^n}}.$$

Furthermore, the CES aggregator is applied to two-component CES sub-aggregators. $\mathcal{E}_{jn}$ aggregates the use of electricity (denoted by E_{jn}) and natural gas (denoted by G_{jn}) by firms. $\mathcal{L}_{jn}$ aggregates the use of workers with a college-degree-or-more education (denoted by C_{jn}) and workers having less than a college degree (denoted by L_{jn}) by firms. Specifically, these sub-component aggregators are parameterized as:

$$\mathcal{E}_{jn} = \left(\zeta_{jn} E_{jn}^{\rho_e^n} + (1 - \zeta_{jn}) G_{jn}^{\rho_e^n} \right)^{\frac{1}{\rho_e^n}}$$

$$\mathcal{L}_{jn} = \left(\theta_{jn} C_{jn}^{\rho_l} + (1 - \theta_{jn}) L_{jn}^{\rho_l} \right)^{\frac{1}{\rho_l}}.$$

Electricity demand by firms generates emissions indirectly (at the regional power plants), so these emissions are sensitive to both use *and* the carbon efficiency of local power plants. Natural gas consumption leads to direct emissions but does not vary by location in the model because I assume the carbon emissions rate of natural gas is the same everywhere.

[12] The recent applied general equilibrium literature has utilized nested CES production functions with various functional forms. See Brockway, Heun, Santos, and Barrett (2017) for a discussion.

Given that I assume factor markets are perfectly competitive, input prices are equal to their marginal products. I assume the supply of capital is perfectly elastic with rental rate $\bar{r}$. The firm chooses its level of capital such that the price of capital is equal to its marginal revenue product.[13] Specifically, the first-order condition for capital utilization yields:

$$K_{jn} = \left(\frac{P_n A_{jn} \eta \mathcal{I}_{jn}^{1-\eta}}{\bar{r}} \right)^{\frac{1}{1-\eta}}. \tag{4.7}$$

Using equation (4.7), I can write the system of first-order conditions to derive the inverse energy and labor demand curves:

$$\begin{aligned}
P_{jn}^{E} &= \mathcal{A}_{jn} \mathcal{I}_{jn}^{1-\rho_{el}^{n}} \mathcal{E}_{jn}^{(\rho_{el}^{n}-\rho_{e}^{n})} \alpha_{jn} \zeta_n E_{jn}^{\rho_{e}^{n}-1} \\
P_{jn}^{G} &= \mathcal{A}_{jn} \mathcal{I}_{jn}^{1-\rho_{el}^{n}} \mathcal{E}_{jn}^{(\rho_{el}^{n}-\rho_{e}^{n})} \alpha_{jn} (1-\zeta_n) G_{jn}^{\rho_{e}^{n}-1} \\
W_{jn}^{C} &= \mathcal{A}_{jn} \mathcal{I}_{jn}^{1-\rho_{el}^{n}} \mathcal{L}_{jn}^{(\rho_{el}^{n}-\rho_{l})} (1-\alpha_{jn}) (\theta_{jn}) C_{jn}^{\rho_{l}-1} \\
W_{jn}^{L} &= \mathcal{A}_{jn} \mathcal{I}_{jn}^{1-\rho_{el}^{n}} \mathcal{L}_{jn}^{(\rho_{el}^{n}-\rho_{l})} (1-\alpha_{jn}) (1-\theta_{jn}) L_{jn}^{\rho_{l}-1},
\end{aligned} \tag{4.8}$$

where

$$\mathcal{A}_{jn} = P_n A_{jn} \left(\frac{A_{jn} \eta}{\bar{r}} \right)^{\frac{\eta}{1-\eta}} (1-\eta).$$

Rents. The housing supply curve is upward sloping with city-specific elasticities and intercepts. Specifically, I parameterize the housing supply curve as:

$$R_j = \bar{K}_j H_j^{\beta_j}. \tag{4.9}$$

Differences in $\bar{K}_j$ across cities reflect differences in local construction costs. The amount of land available for production and the tightness of local land-use restrictions drive variation in β_j (Saiz, 2010). For example, consider a city with less land available for development. All else equal, the marginal cost of developing land that is relatively more scarce will be greater, which will lead to a higher value of β_j.

[13] In Appendix C.0.1.1, I provide more details for the first-order condition derivations.

Taking logs of both sides of equation (4.9) yields:

$$\log(R_j) = \log(\bar{K}_j) + \beta_j \log(H_j). \tag{4.10}$$

The log-log form in equation (4.10) highlights the constant elasticities of the housing supply curves used in the model.

4.2.3 Energy Supply. There are three energy markets in the model: electricity, natural gas, and oil. For each market, demand is endogenous and downward sloping. I assume that natural gas and fuel oil are traded on international markets, and their supply is perfectly elastic. Due to high transmission costs, I assume electricity is traded within NERC regions but not across them. Furthermore, I segment electricity supply into two separate markets: residential and industrial. Electricity supply varies across the type of consumer k due to differences in transmission costs for households compared to firms. Conditional on consumer type, electricity supply varies across local labor markets j due to differences in transmission costs within a NERC region $\mathcal{R}$. I parameterize the inverse supply curve for electricity as:

$$P_{kj}^{\text{elec}} = a_{kj} Q_{\mathcal{R}(j)}^{\mu}, \tag{4.11}$$

where $k \in \{\text{Residential}, \text{Industrial}\}$ indexes the consumer "type", a_{kj} is the consumer-city intercept, $Q_{\mathcal{R}(j)}$ is the electricity supply in NERC region $\mathcal{R}$ (where $R(j)$ maps cities to their corresponding NERC region), and μ is the inverse electricity supply elasticity.

To be clear, the model developed in this paper says nothing about dynamics, and any identified equilibrium is considered a "long-run" equilibrium. Typically in the short run, electricity markets models use "constant-order dispatch curves." In

the long run, however, electricity suppliers can respond to changes in demand by opening new power plants or switching fuel types.[14]

4.2.4 Emissions. In the context of the model, carbon emissions arise from the consumption of energy inputs by agents and firms. Agents consume natural gas, fuel oil, and electricity, whereas firms consume natural gas and electricity. I assume a constant carbon emissions factor of 117 lbs per thousand cubic feet for natural gas and 17 lbs per gallon for fuel oil.[15] Emissions from electricity generation vary across NERC regions, denoted by $\delta_{\mathcal{R}}^{elec}$. I assume that total regional carbon emissions from electricity in a NERC region equals the output-weighted average carbon-emission factors for individual electricity generating units (EGUs). The emissions factor in NERC region $\mathcal{R}$ is therefore given by:

$$\delta_{\mathcal{R}}^{elec} = \sum_{g \in \mathcal{R}} \frac{\text{elec}_g}{\text{elec}_{\mathcal{R}}} \times \frac{\text{CO}_{2,g}}{\text{elec}_g},$$

where elec_g is the amount of electricity produced by a generator g, $\text{elec}_{\mathcal{R}}$ is the total amount of electricity produced in NERC region $\mathcal{R}$ and $\text{CO}_{2,g}$ is the total (yearly) amount of carbon dioxide emitted by generator g. More generally, I write the emissions factor for fuel-type m in city j as:

$$\delta_j^m = \begin{cases} \delta_{\mathcal{R}(j)}^{elec} & \text{if } \; m \in \{\text{elec}\} \\ \delta_m & \text{if } \; m \in \{\text{gas}, \text{oil}\} \end{cases}$$

To obtain aggregate emissions, I multiply the energy consumption in city j and sector n by the respective conversion factors for agents and firms and then sum

[14] Scott (2021b) shows that the EPA's Mercury and Toxics Air standards induced many power plants to convert their coal generators to natural gas.

[15] https://www.eia.gov/tools/faqs/faq.php?id=73t=11

these. Concretely, this is given by:

$$\texttt{Emis} = \sum_j \sum_n \delta_j^m \hat{f}_{jn},$$

where $\hat{f}_{jn} = \sum_m \sum_e N_{ejn} x_{ejn}^m + E_{jn} + G_{jn}$ is the sum of agent fuel consumption ($N_{ejn} x_{ejn}^m$) for each city-sector and firm fuel consumption.

4.2.5 Equilibrium. Equilibrium in the model is achieved when agents and firms make optimal choices and all markets clear. Specifically, an equilibrium requires:

(1) **Utility Maximization.** Each agent must be in a sector and at a location that yields maximal utility, given their constraints. The equilibrium population for each education group-city-sector, $N_{ejn}^\star$, can thus be determined by: $N_{ejn}^\star = N \times s_{ejn}$, where $s_{ejn} = \frac{1}{N} \sum_i P_{ijn}^e$ are the shares computed from the choice probabilities in equation (4.6).

The population distribution determines each city's housing and energy demand. Housing demand in city j is given by the sum of all individual agents' demands. I write total equilibrium housing and energy demand in city j as

$$\begin{aligned} H_j^{\mathcal{D}} &= \sum_{e \in E} \sum_{n \in N} N_{ejn}^\star \times \frac{\alpha_e^H w_{ejn}}{\boldsymbol{\alpha}_{ejn} R_j} \\ x_j^{\mathcal{D}} &= \sum_{e \in E} \sum_{n \in N} N_{ejn}^\star \times \frac{\alpha_{ejn}^m w_{ejn}}{\boldsymbol{\alpha}_{ejn} P_j^m} \qquad \forall m \in \{\text{elec}, \text{gas}, \text{oil}.\} \end{aligned} \tag{4.12}$$

Aggregate labor supply is given by the number of efficiency units of labor supplied by each worker:[16]

$$\mathcal{L}_{ejn}^{\mathcal{S}} = N_{ejn} \times \ell^e. \tag{4.13}$$

[16] Estimation of efficiency units of labor is relatively standard and thus the details can be found in Section C.0.1.2

(2) **Profit Maximization.** Firms maximize profits. This implies that the first-order conditions given by equation (4.8) must be satisfied.

(3) **Market Clearing.** All markets in the model need to clear. Namely, supply and demand must be balanced in the labor market, in the housing market, and in all energy markets.

4.3 Data

I combine data from multiple sources. I obtain individual-level data from the ACS and Census. Data for energy prices and sectoral energy use data come from the Energy Information Association (EIA). I employ data from the Environmental Protection Agency (EPA) to calculate power-plant carbon emissions.

4.3.1 Sources.

Household Data. The time-aggregated 5-year ACS files (Ruggles et al., 2010) have detailed individual-level information for more than 5 million individuals in the US from 2012-2016. Crucially, the ACS has both the location where the individual currently resides (down to the MSA level) and their current occupational sector. Furthermore, the ACS provides information on the individual's monthly rent payments, pre-tax and after-tax wages, and energy expenditures for various fuel types. In estimation, I use repeated cross-sections consisting of official 5% samples of 1990, 2000, and 2010 decennial censuses. I stratify households into one of two education levels: having a college degree (or more) or having less than a college degree. In selecting my sample, I closely follow the strategies used by other researchers and thus relegate the details of the process to Appendix C.0.7.

CBSA Data and Sectoral Data. My measure of a "city" (or location, in the model) is a Core Based Statistical Area (CBSA). CBSAs correspond to

distinct (and relatively closed) labor markets and are the official definition of a metropolitan area for the Office of Management and Budget. For tractability, I restrict the choice set to the 70 largest CBSAs by 1980 population, plus nine additional "outside options"—one for each census division. I categorize workers as participating in one of five sectors: services, construction, agriculture, manufacturing, and an outside-option sector. These five sectors account for roughly 90% of total 2016 employment.[17] My selection of cities and sectors results in a choice set containing 395 elements, where each alternative is a city-sector pair. Recall that I focus on household consumption of three energy types: electricity, natural gas, and fuel oil. The ACS and Census datasets contain information on household expenditures of each of these energy types. I combine household energy expenditure data with state-level energy price data from the EIA to calculate household energy consumption. The methods by which I construct wage, rent, and household energy consumption series are detailed in in Appendices C.0.1.2, C.0.1.3, and C.0.2 respectively.

Energy and Emissions Data. I use sector-level aggregated energy consumption data from the EIA. The EIA's Manufacturing Energy Consumption Survey (MECS) provides information on aggregate manufacturing energy consumption, while EIA's Annual Energy Outlook provides aggregate energy consumption data for construction, agriculture, and other sectors. I also collect data on aggregate electricity and natural gas consumption for the services sector from the EIA. For privacy (and security) reasons, none of my public-use data contain information about sectoral energy demand at the city level. I impute city-sector energy consumption as that city's share of employment (for that

[17] I drop some sectors such as military and other public services from the sample. For more details, see Appendix C.0.7.

sector) multiplied by total sectoral energy consumption. In other words, energy consumption for a city-sector is proportional to the city's level of employment for that sector.[18]

Carbon emissions from electricity vary with the fuel used to generate that electricity. I use plant-level data from the EPA's Emissions Generation Resource Integrated Database (eGRID). Local power plants often trade with each other to meet demand, so I use nine North American Electric Reliability Council (NERC) regions to calculate carbon emissions factors for electricity. While power plants occasionally trade across the borders of these NERC regions, the NERC regions are, for the most part, effectively closed markets (see (Holland & Mansur, 2008)). I calculate the emissions factor as the weighted average of CO_2 emissions per megawatt-hour of electricity across all plants in a given NERC region. A map of these regions and their respective emissions factors is provided in section C.0.5.

4.4 Estimation

In taking my model to the data, I use a combination of calibration and estimation techniques. I focus on the exposition of the labor supply parameters. The labor demand and rent parameters are fairly standard, and are described in Appendix sections C.0.1.2 and C.0.1.3, respectively. Appendix Table C3 provides a full summary of the model's estimation and calibration.

4.4.1 Labor Supply. To estimate the preference parameters in the agent's utility function, I use a two-step estimation procedure that combines maximum likelihood and an instrumental variables approach. The estimation procedure exploits repeated cross-sections of household microdata. Specifically, I use repeated cross-sections of 5% samples of the U.S. Census for

[18]For more details, see Appendix C.0.7.2. This assumption implies that energy-labor ratios will be constant across cities (but not necessarily across industries).

1990, 2000, and 2010 and the 2017 five-year American Community Survey. Let $t \in \{1990, 2000, 2010, 2017\}$ denote the sample year. In what follows, I provide details for each step of the estimation routine.

Step one: Maximum Likelihood. I normalize indirect utility by the scale parameter of the idiosyncratic preference shock. More specifically, I divide equation 4.5 by σ_e to yield:

$$\begin{aligned} v_{ijnt} = \Theta_e^w \log(w_{ejnt}) - \Theta_e^r \log(R_{jt}) - \sum_m \Theta_{ejnt}^m \log P_{jt}^m + \\ \Theta_{et}^{div} \mathbb{I}\left(j \in \mathcal{B}_i^{div}\right) + \Theta_{et}^{\text{dist}} \phi\left(j, \mathcal{B}_i^{st}\right) + \Theta_{et}^{\text{dist2}} \phi^2\left(j, \mathcal{B}_i^{st}\right) + \\ \xi_{ejnt} + \epsilon_{ijnt}, \end{aligned}$$

where the new notation for the preference parameters (Θ) indicates the original parameter multiplied by $\frac{1}{\sigma_e}$. I write the shared (conditional on education group) component of utility associated with each city-sector as:

$$\delta_{ejnt} = \underbrace{\Theta_e^w \log(w_{ejnt}) - \Theta_e^r \log(R_{jt}) - \sum_m \Theta_{ejnt}^m \log P_{jt}^m}_{\text{observed}} + \overbrace{\xi_{ejnt}}^{\text{not observed}} . \tag{4.14}$$

I will refer to the δ_{ejnt} values as the "mean utilities" associated with each alternative. I emphasize the separability between the observable and unobservable components to the mean utilities because this structure is crucial for matching the model's choice shares to the choice shares observed in the data. Given equation (4.14), I can re-write the probability that agent i chooses city-sector jn as:

$$P_i = \frac{\exp\left(\delta_{ejnt} + \Theta_{et}^{div} \mathbb{I}\left(j \in \mathcal{B}_i^{div}\right) + \Theta_{et}^{\text{dist}} \phi\left(j, \mathcal{B}_i^{st}\right) + \Theta_{et}^{\text{dist2}} \phi^2\left(j, \mathcal{B}_i^{st}\right)\right)}{\sum_{j' \in J} \sum_{n' \in N} \exp\left(\delta_{ej'n't} + \Theta_{et}^{div} \mathbb{I}\left(j' \in \mathcal{B}_i^{div}\right) + \Theta_{et}^{\text{dist}} \phi\left(j', \mathcal{B}_i^{st}\right) + \Theta_{et}^{\text{dist2}} \phi^2\left(j', \mathcal{B}_i^{st}\right)\right)}. \tag{4.15}$$

Then, using these choice probabilities, the log-likelihood function is given by:

$$\mathbf{L}(\mathbf{\Theta}_{et}) = \sum_{i=1}^{N^e} \sum_{n \in N} \sum_{j \in J} \mathbb{I}_{ijn} \log(P_i), \tag{4.16}$$

where $\mathbb{I}_{ijn}$ is an indicator equal to one if agent i chooses to live in city j and work in sector n and N^e is the total number of workers of education group e. I jointly estimate the mean utilities δ_{ejnt} and the parameter vector $\mathbf{\Theta}_{et}$ with a nested fixed-point algorithm, proposed in Berry (1994).[19] For details of the implementation of this algorithm, see appendix C.0.6.1.

Step two: Decomposition. In the second step, I decompose the mean utilities to estimate the parameter vector (Θ_e^w, Θ_e^r). To limit the dimensionality of the parameter space, I elect not to estimate the parameters Θ_{ejnt}^m. Instead, I follow Colas and Morehouse (2022), and define $\tilde{\alpha}_{ejnt}^m = \frac{\alpha_{ejnt}^m}{\alpha_{ejt}}$. Given the Cobb-Douglas utility function, the expenditure share of fuel type m is given by $\frac{x_{ejnt}^m \times P_{jt}^m}{w_{ejnt}^m} = \tilde{\alpha}_{ejnt}^m$. I choose values for the $\tilde{\alpha}_{ejnt}^m$ to match the estimated expenditure shares.[20] Next, I rewrite the mean utility as:

$$\delta_{ejnt} = \Theta_e^w \log(\tilde{w}_{ejnt}^{EA}) + \Theta_e^r \log(R_{jt}) + \xi_{ejnt}, \tag{4.17}$$

where $\tilde{w}_{ejnt}^{EA} = \dfrac{\log(w_{ejnt}) - \sum_m \left(\tilde{\alpha}_{ejnt}^m \log(P_{jt})\right)}{1 - \sum_m \tilde{\alpha}_{ejnt}^m}$ is net income adjusted for the energy budget.[21] Taking first differences of equation (4.17) yields my estimating equation:

$$\Delta\delta_{ejn} = \Theta_e^w \Delta \log(\tilde{w}_{ejn}^{EA}) + \Theta_e^r \Delta \log(R_j) + \Delta\xi_{ejn}, \tag{4.18}$$

[19] I implement this algorithm with the Nevo (2000) strategy to speed up the contraction mapping proposed in Berry (1994). With a slight abuse of notation, the Nevo contraction is given by $\exp(\delta_{\tau+1}) = \exp(\delta_\tau) \times \frac{s_{jn}^{\text{data}}}{s_{jn}^{\text{model}}(\delta_\tau)}$ where τ denotes iteration number and $s_{jn}^{\text{model}}(\delta_\tau)$ are the predicted choice shares—which are a function of the mean utilities.

[20] For details on how I impute baseline energy use by city and demographic group, see Appendix C.0.2.

[21] For details of this transformation, see Appendix C.0.3.

where $\Delta\xi_{ejn}$ is the change in the shared unobservable component of amenities. Changes in unobservable amenities at the city or sector level (such as construction of a new park) *mechanically* confound OLS estimates of Θ_e^w and Θ_e^r. For example, consider a city-wide school improvement program.[22] Such a program will induce migration towards this city (due to an increase in the value of choosing this location). Equilibrium wages and rents adjust to the new population level (governed by the equations described in Section 4.2), further affecting utility. Thus, even after taking first differences, Θ_e^w cannot be consistently estimated via pooled OLS due to the endogeneity caused by changes in unobserved amenities.

I use an instrumental variables strategy that exploits exogenous local labor demand shocks. More specifically, I employ an instrument first introduced by Katz and Murphy (1992). The instrument uses historical industry concentration patterns at the city level and interacts them with changes in hours worked across each industry.[23] Specifically, I can write the Katz-Murphy (KM) index for city j between any two sample periods as:

$$\Delta Z_{ejnt} = \sum_{\iota \in n} \omega_{ej\iota}^{1990} \times \left(\Delta \text{Hours}_{e,-j,\iota}\right),$$

where $\omega_{ej\iota}^{1990}$ is the 1990 share of total hours worked by education group e in industry ι in city j as a fraction of the total hours worked by education group e in all industries in city j in 1990. $\Delta\text{Hours}_{e,-j,\iota}$ is the change in national hours worked in all cities other than city j. I construct the differences in national hours (omitting city j), $\Delta\text{Hours}_{e,-j,\iota}$, as decadal differences between each of my sample years.[24]

[22] Diamond (2016) considers the case where residential amenities are endogenous to individual location choices.

[23] For a discussion of shift-share instruments see Goldsmith-Pinkham et al. (2020).

[24] Given that I am estimating the model's labor supply parameters using pooled first differences, I need to estimate mean utilities (and hence individual moving-cost parameters) for each of my cross-sections. Each cross-section has a likelihood function constructed from millions of observations and 395 alternatives. Even with cloud computing, it is not computationally feasible

Finally, to instrument for Θ_e^r, I follow Diamond (2016) and interact the Katz-Murphy index with the elasticity of the housing supply curve. Note that I construct my current instrument such that it generates exogenous labor demand shocks at the *city-sector* level. Rents are assumed to vary only across cities (and not sectors), so I use a slightly different version of the instrument in which I sum over all industries and sectors within a city. Concretely, I write the "city-level" KM index as: $\Delta\tilde{Z}_{ejt} = \sum_n \omega_{ejn}^{1990} \times (\Delta \text{Hours}_{e,-j,n})$. Differences in the responsiveness of housing prices to population changes generate the variation used to identify Θ_e^r. For example, suppose we have two cities that experience identical labor demand shocks. The city with the less-elastic housing supply curve will see rents bid up faster. This variation in rents is assumed to be exogenous to changes in unobservable amenities. After estimating the agent's preference parameters and the mean utilities for each sector-city alternative, I calculate the unobservable component of amenities as the residuals from equation (4.17).

4.4.2 Other Parameters.

Electricity Supply The reduced-form equilibrium expression for residential electricity prices is given by:

$$\log(P_{kj}^{\text{elec}}) = a_{kj} + \mu \times \log(x_j^{elec,\mathcal{D}}),$$

where $x_j^{\text{elec},\mathcal{D}} = \sum_{e \in E} \sum_{n \in N} N_{ejn} \times \frac{\alpha_{ejn}^m w_{ejn}}{\boldsymbol{\alpha}_{ej} P_j^m}$ is electricity demand in city j. I calibrate the electricity supply curve using $\kappa = \frac{1}{2.7}$ following C. Dahl and Duggan (1996).

---

to estimate the individual choice parameters on the entire sample. As a consequence, I make various sample restrictions to limit the number of observations. I describe these in detail in Section C.0.7.

Simplifying this expression yields an expression analogous to equation C.16:[25]

$$\log(P_{kj}^{elec}) = \frac{\mu}{1+\mu}\log\left(\sum_e\sum_n N_{ejn}\frac{\left(\alpha_{ejn}^{\text{elec}} \times w_{ejn})\right)}{\alpha_{ejn}}\right) + a_{kj} \quad \text{for} \quad k = \text{residential}, \tag{4.19}$$

I choose the a_{kj} values to match the data. When k = industrial, I set a_{kj} = $\log(P_{kj}^{\text{elec}}) - \mu \times \log(E_j)$, where $E_j = \sum_n E_{jn}$ is firm energy consumption in city j (aggregated across sectors).

Firms and Rents As does Card (2009), I calibrate the elasticity of substitution between college and non-college workers to 2.5. I set my baseline calibration of energy-labor substitution elasticities to those in Koesler and Schymura (2012).[26] Inter-fuel-substitution elasticities come from Serletis, Timilsina, and Vasetsky (2010).

I solve for input use intensities and total factor productivity (TFP) using relatively standard algebra and thus relegate the details to Appendix C.0.1.2. I calibrate housing supply elasticities to those estimated in Saiz (2010). The reduced-form rental supply can be found in Appendix C.0.1.3.

4.4.3 Parameter Estimates.

Table 8 displays the preference parameters estimates.

[25] I abuse notation here slightly. Technically, the intercept term in equation 4.19 is $\frac{a_j}{1+\kappa}$.

[26] Fried (2018) calibrates a similar parameter, the elasticity of substitution between energy and non-energy inputs, to be close to zero. This requires energy and non-energy inputs to be used in approximately fixed proportions.

	No College			College		
Year	Θ_{lt}^{div}	Θ_{lt}^{dist}	Θ_{lt}^{dist2}	Θ_{ct}^{div}	Θ_{ct}^{dist}	Θ_{ct}^{dist2}
1990	1.696	-3.4318	0.741	1.418	-2.649	0.628
	(0.004)	(0.002)	(0.001)	(0.063)	(0.033)	(0.016)
2000	1.677	-3.438	0.806	1.412	-2.618	0.644
	(0.011)	(0.005)	(0.003)	(0.036)	(0.010)	(0.004)
2010	1.702	-3.226	0.711	1.474	-2.523	0.601
	(0.003)	(0.003)	(0.002)	(0.011)	(0.006)	(0.003)
2017	1.698	-3.218	0.696	1.489	-2.609	0.644
	(0.004)	(0.005)	(0.004)	(0.012)	(0.006)	(0.003)
Income and Rents	No College			College		
Θ_e^w	3.558***			7.0362***		
	(0.591)			(0.815)		
Θ_e^r	-2.160***			-3.731***		
	(0.372)			(0.348)		
Cragg-Donald F-Stat: 14.63						

Table 8. Parameter estimates for household labor supply. Standard errors are in parentheses. Maximum likelihood standard errors are estimated with numerical derivatives. Stars indicate statistical significance: *p<0.05; **p<0.01; ***p<0.001.

All of the parameter estimates have signs that are consistent with the extant literature. For each education group and year, $\Theta_{et}^{div} > 0$ and thus agents receive a utility premium for locating within their birth division. Furthermore, all agents receive disutility for locating farther away from their birth state (Θ_{et}^{dist} ¡ 0). However, the marginal disutility of an additional mile on utility declines with distance (Θ_{et}^{dist2}). The birth-division premium and marginal disutility per mile are also larger in absolute magnitude for college than non-college-educated agents. These estimates suggest that household heads with a college degree are more mobile than those without a college degree.

To my knowledge, none of the existing literature uses the exact functional form I have chosen for individual moving costs. However, Diamond (2016)

estimates a birth-division indicator (Θ_{et}^{div}) for college-educated and non-college-educated workers while Colas and Hutchinson (2021) estimate the distance and distance squared parameter (Θ_{et}^{dist} and Θ_{et}^{dist2}). My estimates for Θ_{et}^{dist2} are very stable across years for both education groups, with the premium being roughly 20% higher for agents without a college degree. My parameter estimates for Θ_{et}^{dist} and Θ_{et}^{dist2} are in a range similar to what is estimated in Colas and Morehouse (2022) (CM). However, CM estimate the parameters for different demographic groups, so a direct comparison is inappropriate.

It is even more difficult to compare my estimates of Θ_e^w and Θ_e^r to the literature, given that I am the first to estimate these parameters for city-sector pairs. My estimates are slightly smaller in magnitude than those of Colas and Hutchinson (2021) and in a similar range to those of Diamond (2016). As do both of these other papers, I find $\Theta_c^w > \Theta_l^w$, which indicates that workers with at least a college degree are more responsive to changes in wages. I also find that $|\Theta_l^r| > |\Theta_c^r|$, which is different from Colas and Hutchinson (2021) but consistent with some specifications of Diamond (2016).

Model Fit. I use the fully estimated model to simulate the baseline equilibrium.[27] In Figure 16, I compare the model's predicted city-sector choice shares by education group to the data:

[27] Details of how I solve for the equilibrium of the model can be found in Section C.0.6.2.

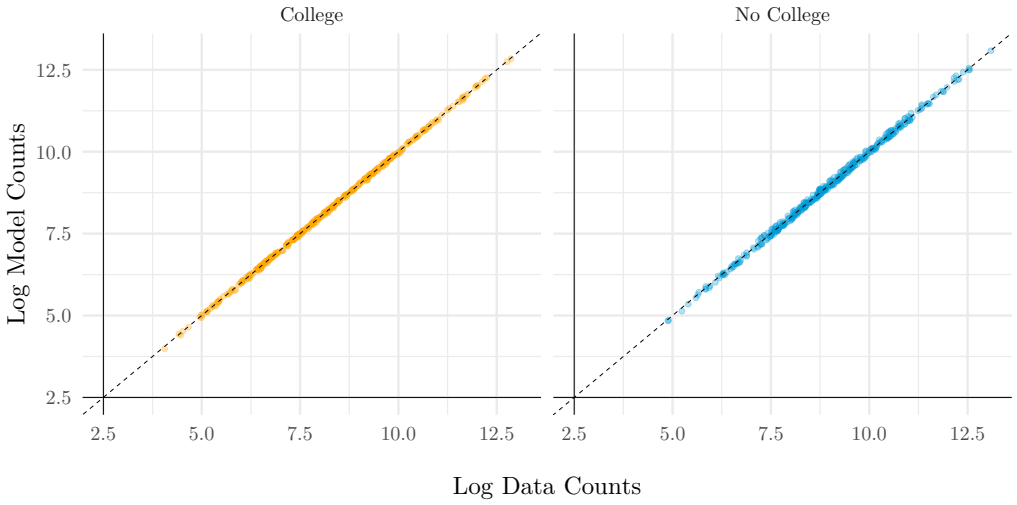


Figure 16. Model Fit Results by education group. For college and non-college-educated workers, these two graphs plot the model's predicted log city-sector shares on the vertical against the corresponding log city-sector shares in the data. The black dashed line is the 45-degree line.

Appendix C.0.2.4 provides analogous scatterplots for household energy consumption. The model appears to fit the data well.

4.5 Counterfactuals

Recently, there has been an increasing legislative effort in the US to pass a carbon tax. In late July of 2019, three carbon tax bills were introduced to Congress: the Climate Action Rebate Act of 2019 (CAR act), the Stemming Warming and Augmenting Pay Act (SWAP Act), and the Raise Wages, Cut Carbon Act of 2019 (RWCC Act). Fundamentally, all three bills are examples of Pigouvian taxes; however, their implementations vary. One of the primary differences between the bills is the use of government revenue. For example, the CAR act proposes distributing 70% of the tax revenue to low-income and middle-income households in the form of lump-sum payments. In contrast, the SWAP and

RWCC acts propose using revenues to reduce payroll taxes.[28] Motivated by these recent examples of proposed legislation, I use my model to assess carbon pricing in conjunction with various transfer schemes.

A carbon tax will affect the price of energy differently for each fuel type. Recall that I assumed the supply curves for fuel oil and natural gas are perfectly elastic, and the emissions rate for each of these fuel types is constant across locations and given by δ^m. With the introduction of a carbon tax, the price of natural gas and fuel oil in city j is given by:

$$\tilde{P}_j^m = P_j^m + \tau\delta^m.$$

For electricity, I construct city-level supply curves as in Equation (4.11). With the carbon tax, the supply curve is given by:

$$P_{kj}^{\text{elec}}(\tau, \delta_{\mathcal{R}(j)}) = a_{kj} Q_{\mathcal{R}(j)}^{\mu} + (\tau \times \delta_{\mathcal{R}(j)}^{\text{elec}}),$$

where I write the electricity supply curve, $P_{kj}^{\text{elec}}(\tau, \delta_{\mathcal{R}(j)})$, as a function of the tax level τ, and the emissions rates $\delta_{\mathcal{R}(j)}$.

4.5.0.1 The Welfare Effects of Carbon Taxes.

In this section, I simulate a carbon tax of \$31 dollars per ton—the Social Cost of Carbon (SCC) as estimated by Nordhaus (2017). I calculate the compensating variation (CV) for agent i as:

$$CV_i = \mathbb{E}[V(\tau > 0)] - \mathbb{E}[V(\tau = 0)] \times \frac{w_{ejn}}{\Theta_e^w}, \tag{4.20}$$

where $V(\tau) = v_{ijn}(\tau, j^\star, n^\star)$ is the indirect utility, given tax level τ, evaluated at equilibrium choices $j^\star$ and $n^\star$ and the expectation $\mathbb{E}$ is taken over the idiosyncratic

[28] For more details of these bills, see https://taxfoundation.org/carbon-tax-bills-introduced-congress/.

preference shocks. Note that because v_{ijn} is measured in log-utility units, $\mathbb{E}[V(\tau > 0)] - \mathbb{E}[V(\tau = 0)]$ is approximately equivalent to percentage change in utility. Multiplying by $\frac{w_{ejn}}{\Theta_e^w}$ converts this to a dollar amount. Agent i's expected utility is given by:

$$\mathbb{E}[V(\tau)] = \bar{\gamma} + \log(\sum_{j' \in J} \sum_{n' \in N} \exp(\Theta_e^w \log(w_{ej'n'}) - \Theta_e^r \log(R_{j'}) - \sum_m \Theta_{ej'n'}^m \log P_{j'}^m + \xi_{ej'n'} + f(j, \mathcal{B}_i; \hat{\theta}_e)))$$

where $\bar{\gamma}$ is Euler's constant and $f(j, \mathcal{B}_i; \hat{\theta}_e)$ is the distance function (equation 4.3) evaluated at the estimated parameter vector $\hat{\theta}_e$. The distribution of CV across cities and sectors is displayed in Figure 17.

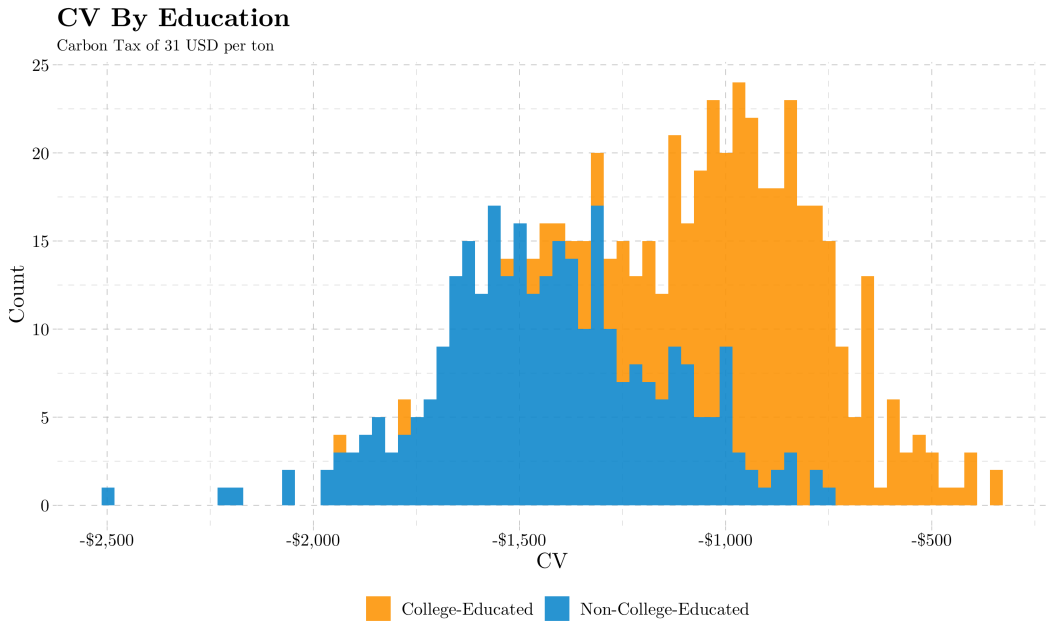


Figure 17. The distribution of compensating variation from a carbon tax of $31.

Figure 17 demonstrates considerable household-level heterogeneity in the burden of a carbon tax. The model predicts that carbon prices are regressive; the mean tax burden for college-educated agents is $926 while the mean tax burden

for non-college-educated agents is $\$1,417$.[29] The tax regressivity is consistent with a large literature that observes that lower-income households spend larger portions of their income on energy, and work in more energy-intensive sectors. Indeed, this is reflected in my estimates of $\tilde{\alpha}^m_{ejn}$ (the budget share of energy type m).[30] Additionally, the model predicts a significant decline in manufacturing employment. Manufacturing, which is relatively carbon-intensive, experiences an 11.1% decrease in employment with a 12.7% reduction in college workers and a 10.4% reduction in non-college workers. The services sector—which is relatively green—experiences an increase in employment. There is a 2.01% increase in aggregate services employment, with a 1.78% increase for college workers and a 2.34% increase for non-college workers.

Figures C4 and C9 in Appendix C.0.5.2 disaggregate Figure 17 by industry and Census Region. Manufacturing has a substantially higher tax burden than other sectors for both college-educated and non-college-educated workers. Cities and sectors in the Midwest and West have marginally higher average compensating variations associated with the tax, comapred to other regions. Table 9 summarizes the consequences of the carbon tax.

[29] I did not impose the tax regressivity on the model's structure in any way.

[30] Summary statistics for these estimates can be found in Table C2.

$\tau = \$31$/ton:	Mean CV ($)	Mean/st.dev CV	%Δ Man. Emp	%Δ Ser. Emp
Total	$-1,221$	-3.14	-11.1	2.01
College	-926	-3.55	-12.7	1.78
Non-College	$-1,417$	-4.16	-10.4	2.34

Table 9. Counterfactual results from a carbon tax of $31 per ton. Mean CV is calculated using equation (4.20). Percent changes in emissions and employment are relative to baseline levels.

Non-Monetized Incidence. While my primary welfare metric is compensating variation, I also examine tax incidence defined as the *percent change* in wages needed to compensate households for the carbon tax.[31] Compensating variation may mask substantial underlying heterogeneity in the tax burden due to differences in wages across cities. For example, San Francisco may experience a small-degree tax burden as a percent of wages. San Francisco has a services-based economy with a mild climate and carbon-efficient power plants. However, San Francisco has high average wages. Thus, the average household in San Francisco may require a large *dollar* transfer in response to the carbon tax. Conversely, Detroit may experience a large loss in utility due to its manufacturing-oriented economy, harsher climate, and relatively dirtier power plants. Detroit also has relatively low wages, which lowers the dollar amount needed to compensate households in Detroit for the carbon tax, all else equal. Despite the potentially larger loss in utility in Detroit, the compensating variation in San Francisco could be higher due to differences in wages across the two cities.

[31] Compensating variation is in levels; the non-monetized incidence is a percent change. I calculate percent change in utility as $CV \times \frac{1}{w_{ejn}}$, or just $(\mathbb{E}[V(\tau > 0)] - \mathbb{E}[V(\tau = 0)]) \times \frac{1}{\Theta_e^w}$.

Indeed, I find compensating variation is larger (in magnitude) in cities that exhibit high wages, such as those on the West Coast and New England. In Figure C5, I replicate figure C4, except I compare non-monetized tax incidence rather than changes in compensating variation. I aggregate these changes to the state level and map them in Figures C6 and C7. These figures reveal that households in the Midwest and South generally experience larger percent decreases in utility from the carbon tax, despite having similar, if not lower, compensating variations.

Migration. Next, I examine how the spatial distribution of households shifts with the carbon tax. More specifically, I calculate the percent change in aggregate population and by education group for each city, aggregated across all sectors. I display the results in Appendix C.0.5.3 at the state level, both in the aggregate and disaggregated by education level.

In general, the model predicts that cities with milder climates and power plants that are less emissions-intensive will experience population increases. For example, the carbon tax induces an inflow of households to Seattle, Washington, increasing the city's population by 1.5%. Washington is part of the WECC NERC region (see Figure C2) and has relatively carbon-efficient power plants, so electricity prices do not increase as much in Seattle as in other cities. Additionally, Seattle's economy is more services-oriented than many other cities. On the other hand, cities such as Cincinnati experience population decreases. Relative to Seattle, Cincinnati's climate requires more household energy, and the electricity generation in Cincinnati is less carbon-efficient than in Seattle. Additionally, Cincinnati has a smaller share of workers in relatively carbon-efficient services-based jobs than Seattle. Overall, the model predicts that the carbon tax will reduce Cincinnati's population by 2.29%.

Despite them being less mobile than college-educated households, the model predicts larger changes in shares of non-college-educated households in most cities. In Figures C12 and C11, I show that the average city experiences a 0.1% decrease in its share of non-college-educated households, and only a .03 % decrease in its share of college-educated households. College-educated workers are more responsive to changes in prices (as estimated in section 4.4), so the relatively larger migration flows of non-college-educated workers demonstrate the greater tax burden borne by these households.

Tax Incidence and Elections. Next, I examine how changes in utility from the carbon tax correlate with political preferences. I focus my analysis on the share of cities voting for Donald Trump during the 2016 presidential election. Donald Trump's voting base was less educated and concentrated in less-populated areas (Doherty, Kiley, & Johnson, 2018)—features that are correlated with higher tax incidence in my analysis.[32]

I combine my tax-incidence estimates with publicly available, county-level presidential election voting data from the MIT Election Data and Science Lab (MIT, 2018). For each CBSA in my sample, I compute the average (across counties within the CBSA) share of the vote going to Donald Trump in the 2016 primary presidential election, weighted by total votes in each county. In Figure C13, I plot the tax incidence (again, defined as the percent change in wages needed to compensate households for the carbon tax) against Trump's vote shares. Additionally, I plot compensating variation against the share of households voting for Trump in Figure C14.

[32]Donald Trump has been consistently skeptical of climate change and policy. On December 6th, 2013, in a tweet that was subsequently deleted alongside his account, Donald Trump stated "Ice storm rolls from Texas to Tennessee - I'm in Los Angeles and it's freezing. Global warming is a total, and very expensive, hoax!"

As expected, counties with a larger share of Trump votes experience larger decreases in utility, on average.[33] The counties with a higher percentage of Trump voters (and thus larger tax incidence) are arguably those least likely to vote in favor of a carbon tax. Further complicating the political feasibility of the tax is that areas with a *lower* share of Trump voters would require a *higher* (in magnitude) compensating variation, on average. This difference between tax incidence and monetized tax incidence (compensating variation) is again driven by differences in wages across cities; cities such as San Francisco have a low share of Trump voters, relatively low tax incidence, but relatively high compensating variation due to the high baseline wages in San Francisco. Overall, these results highlight a particular challenge concerning the political feasibility carbon-pricing legislation.

4.5.1 The Effects of Coal-Fired Electricity. In large part due to significant decreases in natural gas prices with the advent of fracking technologies, the share of electricity generated by coal in the United States has dropped precipitously since the mid-2000s. In the early 2000s, coal-fired generation accounted for over 70% of power nationally. By 2015, coal's share had diminished to under 40% of generation.[34] Coal emits a considerable amount of carbon compared to natural gas, so the declining share of coal will have first-order consequences for the welfare effects of pricing carbon. As coal's share of electricity generation continues to decline, the revenue needed to compensate households for

[33]Note that I did not target political preferences in the estimation, so this result strengthens the validity of the model's parameter estimates.

[34]The decline in the share of electricity generated by coal is for many reasons in addition to the decline of natural gas prices. See Arias, Reinbold, and Restrepo-Echavarria (2017) for a brief discussion.

carbon pricing will also fall. I illustrate the decline in coal nationally and across NERC regions in Figure 18.

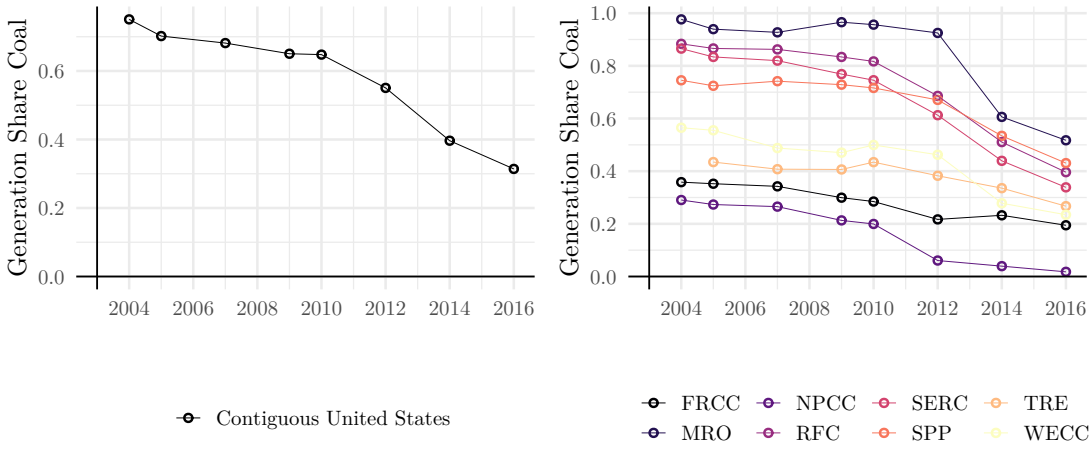


Figure 18. Share of electricity generated by coal nationally and across NERC regions between 2004 and 2016. Shares are computed using annual plant-level generation data from eGRID.

No Coal. Motivated by the remarkable decline in the share of the electricity generated by coal, I use the model to decompose the tax incidence into two components: coal and non-coal. More specifically, I estimate the proportion of the total compensating variation that is attributable to the effect of a carbon tax on coal-fired electricity generation. I recalculate emissions factors for each region in the absence of coal. Without coal, there are considerably lower emissions rates across the US. A map of the changes in emissions factors from dropping coal is in Section C.0.5. With the new, absent-of-coal emissions factors in hand, I resimulate the $31 per ton carbon tax for a "no-coal" economy. I display the distribution of CV, split by education level, in Figure 19.

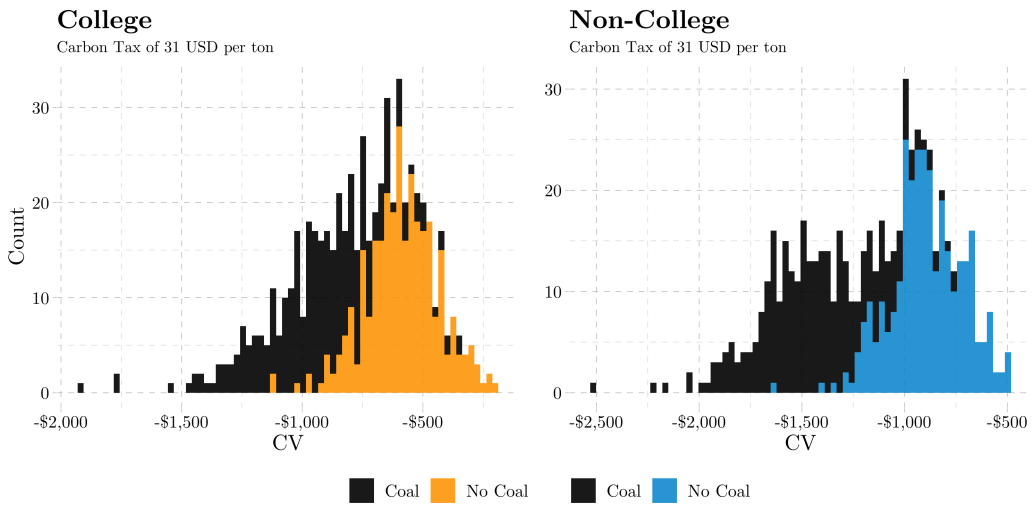


Figure 19. The distribution of compensating variation from a carbon tax of $31 per ton in the absence of coal-fired electricity generation.

Without coal, the mean college and non-college agents experience a $352 and $543 dollar decline in their tax incidence. Percentage-wise, the change is marginally larger for non-college workers as they are generally located in more energy-intensive cities and work in more sectors.

Furthermore, the distribution of incidence without coal-fired electricity has a smaller variance. The distributions of tax incidence have less dispersion due to the geographic distribution of coal; many coal plants are in the Midwest and South. In Appendix C.0.5.2 figure C6 I show that for non-college-educated workers, the change in tax incidence is greatest in the midwest. In contrast, for college-educated workers, the change in incidence is remarkably stable across regions. I display the main results for the "no-coal" scenario in Table 10 below.

$\tau = \$31$/ton: **No Coal**	Mean CV ($)	Mean/st.dev CV	%Δ Man. Emp	%Δ Ser. Emp
Total	−757	−3.11	−8.15	1.51
College	−594	−3.33	−9.56	1.31
Non-College	−874	−4.06	−7.51	1.68

Table 10. Counterfactual results: a $31 per ton carbon tax in the absence of coal. Mean CV is calculated using equation (4.20). Percent changes in emissions and employment are relative to baseline levels.

To be clear, this is purely a decompositional and bounding exercise with respect to the tax incidence. It suggests what could be the distribution of tax incidence without coal. A carbon price will impact the rate at which coal-fired power plants are shut down. In reality, the supply curves of electricity are determined by a "constant dispatch order." The plant uses the generating unit with the fuel type that has the lowest marginal cost, up to a fuel-specific capacity constraint. The plant then switches on the next lowest (fuel-specific) marginal cost generators upon hitting each fuel-specific constraint. Given that coal has a low marginal operating cost, removing coal would be likely to impact significantly the slope of the supply curve, further changing the impact of a carbon tax on electricity prices and tax incidence. I leave it to future work to recover the incidence of carbon pricing with dynamic coal shares and a micro-founded electricity supply curve.

This counterfactual is highly policy-relevant as coal's decline continues. My decomposition demonstrates that a significant proportion of the total compensating variation (roughly 40%) from carbon pricing is attributable to losses in coal-fired electricity. These simulations suggest policymakers may need to compensate

households significantly less to remain indifferent between a carbon tax and no tax in a low-coal future.

4.5.2 Equity and Emissions. Next, I examine the relationship between progressive transfers and aggregate emissions. I provide a motivating example from the data and then extend the baseline model to include government payments.

Wages and Emissions. Variation in household carbon emissions across cities and sectors is well documented. This variation is driven largely by (1) differences in climate across cities, (2) differences in emissions intensities from regional power plants,(3) differences in preferences for energy consumption, and (4) differences in income levels across households. Cities also host different sets of industries which vary in their energy use, and hence their emissions intensities. Thus, any policy that affects the spatial and sectoral population distribution will also impact aggregate emissions.

If lower-wage city-sectors have, on average, higher emissions, then progressive redistribution may *increase* aggregate emissions. This relationship arises because the more carbon-intensive city-sector combinations will receive larger transfers (all else equal), which will induce a larger share of workers to move to these cities and/or into these sectors, raising aggregate emissions. However, the effect of the transfers on aggregate emissions also depends on substitution patterns across cities and sectors. For example, even if the correlation between wages and emissions is positive, progressive transfers may induce a large share of workers into a low-wage yet emissions-intensive city-sector. Thus, the effect of progressive

transfers on aggregate emissions depends on both the correlation between wages and emissions and substitution patterns across city sectors.[35]

In Figure 20, I plot city-sector emissions per capita—defined as the sum, per capita, of firm and household emissions per capita—against city-sector wages.

Figure 20. Carbon emissions at the city-sector level plotted against average wages. On the vertical axis, total emissions per capita are calculated as aggregated firm and household emissions divided by the respective city-sector employment count. On the horizontal axis, wages are estimated by city-sector-education group in Equation (C.3), then averaged by city-sector, weighted by the count of workers in that education group.

As Figure 20 indicates, the average household that lives and works in a lower-wage city-sector also tends to emit less carbon. In Appendix C.0.8, I

[35]There are two exceptions to this: when there are only two alternatives (cities and or sectors), or when wages and emissions are perfectly correlated. In the first case, there is only one alternative for workers to switch into (the alternative with higher transfers), and given the negative correlation between wages and emissions, this alternative will be lower. In the second case, lower-wage alternative have strictly lower emissions. Thus any substitution to a lower wage alternative would reduce aggregate emissions.

disaggregate Figure 20 by education group to reveal that the positive relationship between emissions and wages holds for college-educated and non-college-educated workers considered separately.[36]

The positive cross-city correlation between emissions and wages has important implications for redistributing the revenue from a carbon tax. Often, policymakers and academics propose progressive transfers as a way to alleviate distributional concerns with pricing carbon.[37] However, whether or not these transfers increase or decrease aggregate emissions depends on both the correlation between wages and emissions and substitutions across the agent's choice sets.

Carbon Taxes with Transfer Payments. I extend the model so that the agent receives a transfer equal to $\mathcal{T}(w)$. The agent's post-transfer wage is then $\tilde{w}_{ejn} = w_{ejn} + \mathcal{T}(w)$. One of the primary goals of this paper is to capture the relationship between the progressivity of government transfers and aggregate carbon emissions. Thus, for the transfer function, I use the parsimonious specification employed by Heathcote, Storesletten, and Violante (2017) (henceforth HSV), which is given by:

$$\mathcal{T}(w) = \lambda w_{ejn}^{1-\gamma}$$

where $\lambda > 0$ is the overall level of the reimbursement and $\gamma \geq 1$ indexes the progressivity of the transfers.[38] Note that when $\gamma = 1$, transfers are proportional to wages, and when $\gamma \geq 1$, transfers are decreasing in wages. A higher value of γ implies higher-wage agents will receive smaller transfers. In counterfactuals with transfers, I append the definition of the model's equilibrium to include a balanced-

[36] Additionally, I disaggregate by education group and industry. For some industries, the relationship between emissions and wages is positive.

[37] See Carattini, Carvalho, and Fankhauser (2017) for an example.

[38] If $\lambda < 0$ and $\gamma < 1$ then this is the familiar HSV tax function.

budget condition for the government. The balanced budget assumption implies that λ is endogenously determined in equilibrium; more details are included in Appendix C.0.2.3.

I use the model to estimate numerically the general equilibrium elasticity of aggregate emissions with respect to the relative progressivity of transfers. I define this elasticity as:

$$\epsilon_{\texttt{Emis},\gamma} = \frac{\partial \texttt{Emis}}{\partial \gamma} \frac{\gamma}{\texttt{Emis}}. \tag{4.21}$$

Note that the level of transfers an agent receives may change the agent's city-sector choice, which impacts the agent's emissions and thus aggregate emissions. Equation (4.21) is thus an implicit function, effectively, of all the model's underlying parameters. If labor supply—especially non-college labor supply—is highly mobile and responsive to changes in wages, then an increase in γ will cause a larger change in emissions as more workers migrate or switch sectors. Firm production parameters (and hence labor demand), energy, and rental supply all determine the full extent of the price changes and hence aggregate emissions change. I estimate equation (4.21) numerically by simulating the model under $\gamma \in \{1, 1.2, 1.4, 1.6\}$. I display the result from this exercise graphically in Figure 21 below.

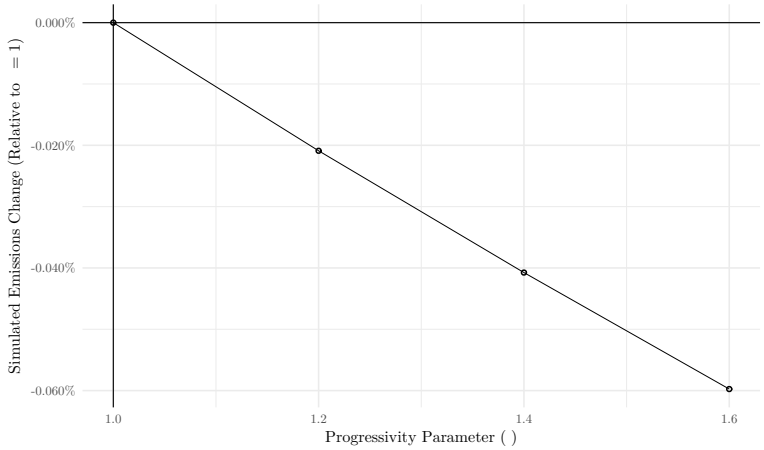


Figure 21. Simulated total emissions relative to lump-sum transfers plotted against the progressivity parameter (γ). Simulated emissions are calculated in equilibrium as the sum of total emissions from agents and firms across all cities and sectors.

Figure 21 illustrates that, in the US, equity-of-transfers and aggregate carbon emissions are inversely related. More specifically, I estimate $\epsilon_{\texttt{Emis},\gamma} = -0.00104$. Put differently, this indicates that a 1% increase in the progressivity of the transfer system would result in approximately a -0.001% decrease in aggregate emissions. In Table 11, I show the change in sectoral composition under a carbon tax with lump-sum transfers ($\gamma = 1$) and a carbon tax with progressive transfers ($\gamma = 1.2$).

$\gamma = \mathbf{1}$	%Δ Man. Emp	%Δ Ser. Emp	%Δ Con. Emp	%Δ Ag. Emp
Total	−11.8	2.42	1.57	−2.78
College	−13.7	1.99	0.07	−3.51
Non-College	−10.9	2.80	1.7	−2.62
$\gamma = \mathbf{1.2}$				
Total	−11.9	2.49	1.36	−1.86
College	−13.8	2.03	0.05	−2.78
Non-College	−11.1	2.91	1.5	−1.65

Table 11. Change in sectoral employment in the aggregate and by education group for a $31 carbon tax with transfers. The change is relative to the same $31 carbon tax, but with no transfers.

As shown in Section 4.3, cities and sectors with lower wages also tend to have lower carbon emissions. Thus, more-progressive transfers induce a larger share of agents to move into greener cities and sectors, lowering aggregate emissions.[39] When $\gamma = 1.2$, I find that manufacturing employment declines by 11.9% (compared to 11.8% with lump-sum transfers) and services employment increases by 2.49% (compared to 2.42% with lump-sum transfers). The model also predicts modest increases in construction employment (1.57%) and decreases in agricultural employment (2.78%). These results suggest that carbon tax policies with progressive transfers have the added benefit of modestly reducing aggregate emissions due to the compositional changes in the workforce induced by these transfers.

4.6 Conclusions

I explore the distributional effects of a uniform national carbon tax. To accomplish this, I estimate a model of worker sorting across cities and sectors, in

[39] If the relationship between wages and emissions were negative, then there would be an equity-emissions *trade-off*.

which imperfectly mobile agents that vary by education level consume three goods: numeraire good, energy, and housing. The model incorporates the incidence of a carbon tax, including endogenous wages, rents, and electricity prices. To take the model to the data, I utilized a combination of estimation and calibration methods.

There are three main takeaways from a set of counterfactual exercises. First, the incidence of a carbon tax exhibits substantial spatial and sectoral heterogeneity—workers in the carbon-intensive manufacturing experience a much greater tax burden than workers in the less carbon-intensive services sector. Overall, I find that a carbon tax has a mean compensating variation of $926 for college-educated workers and $1,417 for non-college-educated workers.

Second, the share of this incidence attributable to coal varies across space. Coal use accounts for a larger share of the total burden in some coal-dependent regions, such as the Midwest, particularly for non-college-educated workers. Other regions, like New England, have a small share of the burden attributable to coal dependence. These findings imply that policymakers will need to adjust expected compensation with via recycling of carbon tax revenues as the grid decarbonizes at potentially different rates across space. Finally, I document a new relationship between equity and emissions: aggregate emissions fall as compensation becomes more equitable. Lower-wage city-sectors are less carbon-intensive, so progressive transfers induce a larger share of workers to move into these cities and sectors. This relationship suggests that progressive transfers may have the added benefit of helping achieve emissions targets.

Despite the detailed heterogeneity included in the model, I have made several simplifying assumptions in my analysis. I make two strong assumptions about capital in the production function. First, I assume the elasticity of substitution between capital and other inputs is one (e.g., the production function

is Cobb-Douglas in capital and other inputs). Second, I assume capital is traded on international markets, and there are no differences in local prices. Complementarity between capital and energy may have important implications for the distribution of carbon tax incidence. Third, I assume perfectly competitive input and output markets. Heterogeneity in market power across cities and sectors could bias my estimates of the tax incidence. In addition, I use natural gas and electricity as primary fuel inputs for firms; while this may be reasonable for many sectors, it certainly underestimates the effects of carbon pricing on agriculture since a large share of agricultural carbon emissions are from livestock. Future work could extend the model by generalizing the production function to incorporate these features.

My results speak to the importance of considering place-based incidence when designing federal policy. Much of the literature has identified negligible employment effects, but I note that even if, on net, the number of jobs is the same or increases, reallocation is costly. My model provides insights concerning the mechanisms via which spatial-sectoral heterogeneity would be created by the tax, and my empirical results substantiate the possibility of costly re-allocations of labor. Understanding heterogeneity in the incidence of a carbon tax is paramount for policymakers looking to reduce the current political headwinds faced by proposals for carbon pricing. As Sallee (2019) notes, "...the failure to create a Pareto improvement is due to a *prediction* problem; lump-sum transfers can only undo the distribution of burdens if they can be targeted precisely." Due to the rich level of heterogeneity in the model, and hence in the simulation results, the analyses this paper may help design policymakers design compensating schemes that can be more precisely targeted.

CHAPTER V

DISSERTATION CONCLUSION

In this dissertation, I have demonstrated the importance of accounting for spatial heterogeneity in environmental policy and related outcomes.

In Chapter 2, we focused on the geography of the US power sector. We first documented that many power plants are located near jurisdictional (county and state) borders. This behavior may indicate strategic siting (intended to export pollutant emissions to neighboring jurisdictions). However, many county and state borders are composed of water (rivers and lakes), a key input for many types of electricity generation. We developed a statistical test to distinguish strategic sites from non-strategic sites. We then used our test and found strong circumstantial evidence that coal-fired power plants have been sited downwind within their own counties and states to export a greater share of their emissions. Motivated by these findings, we employed a state-of-the-science atmospheric dispersion model, HYSPLIT, to document the pervasiveness of the "pollution-transport problem." We provided new statistics on the speed and scale of particulate matter transport from coal-fired power plants.

In Chapter 3, we examined how a locally set policy (land-use regulations) can impact *aggregate* carbon emissions. To facilitate our analysis, we developed and estimated a quantitative spatial equilibrium model. We used the model to reduce existing stringent land-use regulations in California to a reasonable level. Relaxing these land-use regulations mechanically lowers rents, and, in the model, households respond by migrating to California. Since California has a temperate climate and carbon-efficient power plants, households emit less carbon in California than they do elsewhere. Thus, aggregate carbon emissions fall when households move into California (due to lower land-use restrictions).

In Chapter 4, I studied the distributional effects of a national uniform carbon tax in the United States. The incidence of a carbon price on a given household will depend not only on location but also on the employment sector for that household since different sectors use different amounts of carbon-intensive inputs. Thus, a carbon tax may lower wages (or even cause a loss of employment) for workers in carbon-intensive sectors. In addition, if a household lives in an area with an extreme climate and carbon-intensive power plants, the carbon tax will significantly impact their energy bill. To understand how a carbon tax would impact different households, I built and estimated a quantitative spatial equilibrium model. I found that non-college-educated workers in manufacturing bear a significant share of the tax incidence. The model also predicts that states with services-oriented economies, such as California, will experience modest population increases in response to a carbon tax. I simulated a carbon tax with progressive transfers given the tax's regressivity. As the progressivity of the transfers increases, aggregate carbon emissions will fall as households sort into greener cities and sectors.

The results from my dissertation underscore the importance of spatial variation in public policy. As demonstrated in Chapter 2, even local pollutants such as $PM_{2.5}$ can travel long distances very quickly. Our results suggest that the geographic level of regulation is a first-order concern for air-quality management. Furthermore, hypothetical policies such as carbon taxes will have unequal impacts on the population, even if implemented uniformly.

APPENDIX A

CHAPTER 2: APPENDIX

A.0.1 Appendix: Methods.

A.0.1.1 Border-distance calculations. We first project the plant's location and the Census shapefiles into the plant's zone of the Universal Transverse Mercator (UTM) coordinate system. Then we calculate the distance to the plant's nearest county and state border. We use `R`'s `sf` package for these calculations (Pebesma, 2018).

A.0.1.2 Counterfactual grid. If the county and state borders do not impact or correlate with EGUs' locations, then EGU's distances to borders should mirror the overall national distribution of distances to borders.To build this comparison distribution, we cover the contiguous U.S. with a uniform, hexagonal grid of points as illustrated in Appendix Figure A2. The number of grid points is approximately equal to the area covered in square kilometers. We then calculate each point's distances to the nearest county border and the nearest state border.[1] This process produced a nationally representative distribution (for the contiguous U.S.) of distances to state and county borders using a uniform grid of approximately 7.91 million points.[2] This distribution represents the expected distribution of EGUs' distances to borders *if* they were sited in a manner that ignores borders and features that correlate with borders.

The last row of Figure 3 depicts the distribution of distance-to-nearest-border for the uniform grid covering the U.S. This grid's distribution demonstrates that it is *not* the case that all points in the United States are near borders. Only 8% of the U.S. (area-wise) sits with 1 kilometer of a county border (36% within 5 km; 62% within 10 km). For state borders, only 1.1% of the U.S. sits within 1 kilometer (6% within 5 km; 11% within 10 km). These numbers stand in stark contrast to the distributions of EGUs.

A.0.1.3 Borders and water. We calculate the share of each county's and state's borders that coincide with bodies of water in four steps. First, we convert each administrative unit's linear boundaries into a series of points with

[1] Specifically, we work in the counties' UTM zones and subset the grid points to the points *within* the county under consideration—a point's nearest border is always the border of the unit that contains that point. Again, we employ `R`'s `sf` package for these calculations (Pebesma, 2018).

[2] For comparison, the area of the contiguous U.S. is approximately 8.08 million km^2.

50-meter spacing. Second, we calculate the distance to the nearest body of water for each of these boundary points (if the boundary point is within a body of water, then the distance is zero). These bodies of water cover all rivers, lakes, and coastlines including in the U.S. Census's TIGER/Lines shapefiles discussed in *Data*. Third, we designate a boundary point as including water if the nearest body of water is less than 50 meters. This step allows for *near misses* in the Census geography files without including too many false positives. Finally, we smooth this *includes water* indicator variable using a moving-window average of all boundary points within a 2.5 kilometer radius of the given boundary point. This final step allows neighboring boundary points to *vote* on whether the boundary indeed coincides with water—e.g., a single, spurious *includes water* will be overwhelmed by non-water neighbors. The final product is a series of points with 50-meter spacing covering all county and state borders in the contiguous U.S.—with each point measuring whether the boundary substantively coincides with water.

A.0.1.4 EGUs and water. To calculate the distance to the nearest body of water, we include all bodies of water contained in the U.S. Census's areas of water, linear water, and coastline shapefiles, (US Census Bureau, 2016b). After merging these calculated distances with eGRID's EGU characteristics, we build the distribution of distance-to-water for each fuel category.

A.0.1.5 HYSPLIT. The `R` packages `splitr`, `hyspdisp`, and `dispersR` were extremely helpful in developing our computational approach—as was `GNU Parallel` (Tange, 2011).

A.0.2 Appendix: Policy.

A.0.2.1 The Clean Air Act and cross-border pollution. The Clean Air Act (CAA)—often called the "crown-jewel" of environmental regulation in the U.S. (Browning, 2020; Feldman, 2010))—recognizes that cross-border air pollution is a challenge on a scale larger than neighboring counties. The original texts of the 1963 CAA limited federal involvement mainly to (a) resolving trans-boundary pollution issues—when invited by a governor—and (b) funding/guiding research related to air pollution (Edelman, 1966; United States Congress (90^{th}), 1968; United States Senate, Committee on Public Works, Staff Report, 1963). Known as the "good neighbor" provision, section 110 of the CAA explicitly prohibits "any source or other type of emissions activity within the State from emitting any air pollutant in amounts which will (I) contribute significantly to

non-attainment in, or interfere with maintenance by, any other State with respect to any such national primary or secondary ambient air quality standard" (U.S. Environmental Protection Agency, 2013a).[3] Further emphasizing the importance of cross-border pollution transport, in 2011 the U.S. EPA enacted the Cross-State Air Pollution Rule (CSAPR). The CSAPR covers 27 states[4] in the eastern U.S.—especially targeting power-plant emissions of SO_2 and NO_x and their formation of fine-particulate matter ($PM_{2.5}$) and Ozone (O_3) (U.S. Environmental Protection Agency, 2020). The CSAPR links emissions-source states to recipient states—emphasizing non-attainment areas—and creates a budget-and-trading program for emissions within the covered states (U.S. Environmental Protection Agency, 2020). Despite this substantial infrastructure addressing cross-border pollution, disputes regarding trans-border pollution continue—e.g., in 2018 Delaware announced its intent to sue the EPA over emissions from power plants based in Pennsylvania and West Virginia, and in 2019 New York, Connecticut, Delaware, New Jersey, Maryland, Massachusetts, and NYC sued the EPA regarding upwind ozone precursor emissions (Groom, 2019; Sanders, 2018; Volcovici, 2018).

One of the complexities of monitoring and regulating air pollution from coal-fired EGUs is the degree to which emissions can travel long distances from the initial source, polluting distant destinations. In 2018, the average height of a smokestacks attached to a coal-fired EGU in the U.S. was approximately 500 feet, and the maximum was approximately 1,038 feet (calculated from CAMD (2020) data). While tall smokestacks aid in dispersing high concentrations of harmful chemicals, they also substantially increase the transport of emissions to other counties and states (U.S. Government Accountability Office, 2011).

A.0.3 Appendix: Figures.

[3] The CAA also allows states to petition the EPA for reviews of upwind sources (U.S. Environmental Protection Agency, 2013b).

[4] Texas, Oklahoma, Kansas, and Nebraska comprise the western edge of the CSAPR states.

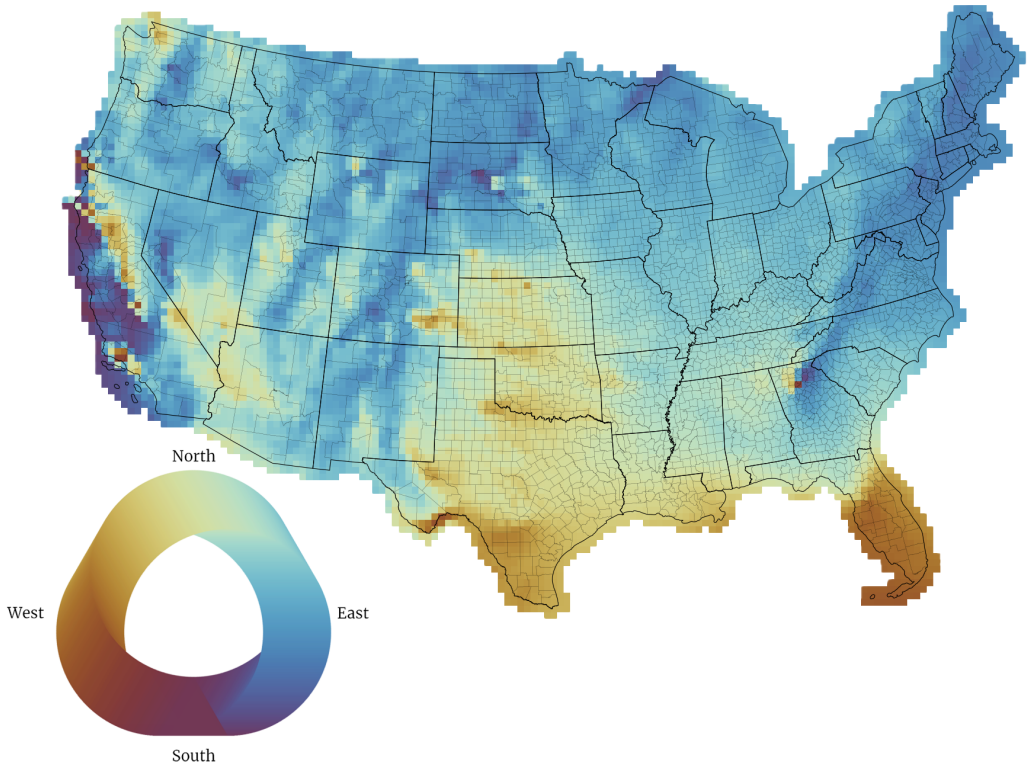


Figure A1. NARR (North American Regional Reanalysis, 2006) prevailing wind directions across the US.

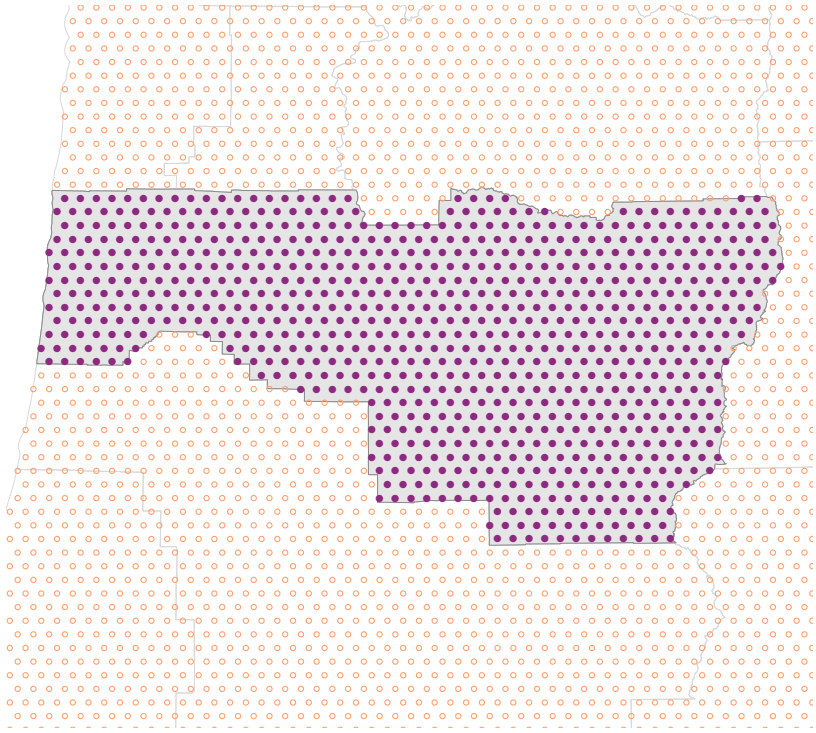


Figure A2. This figure illustrates the uniform grid within our nearest-border calculation. All dots (open and closed) are part of the uniform grid. Closed, dark purple dots are within Lane County, Oregon. We then calculate the shortest distance from each dot to borders of Lane County and of Oregon.

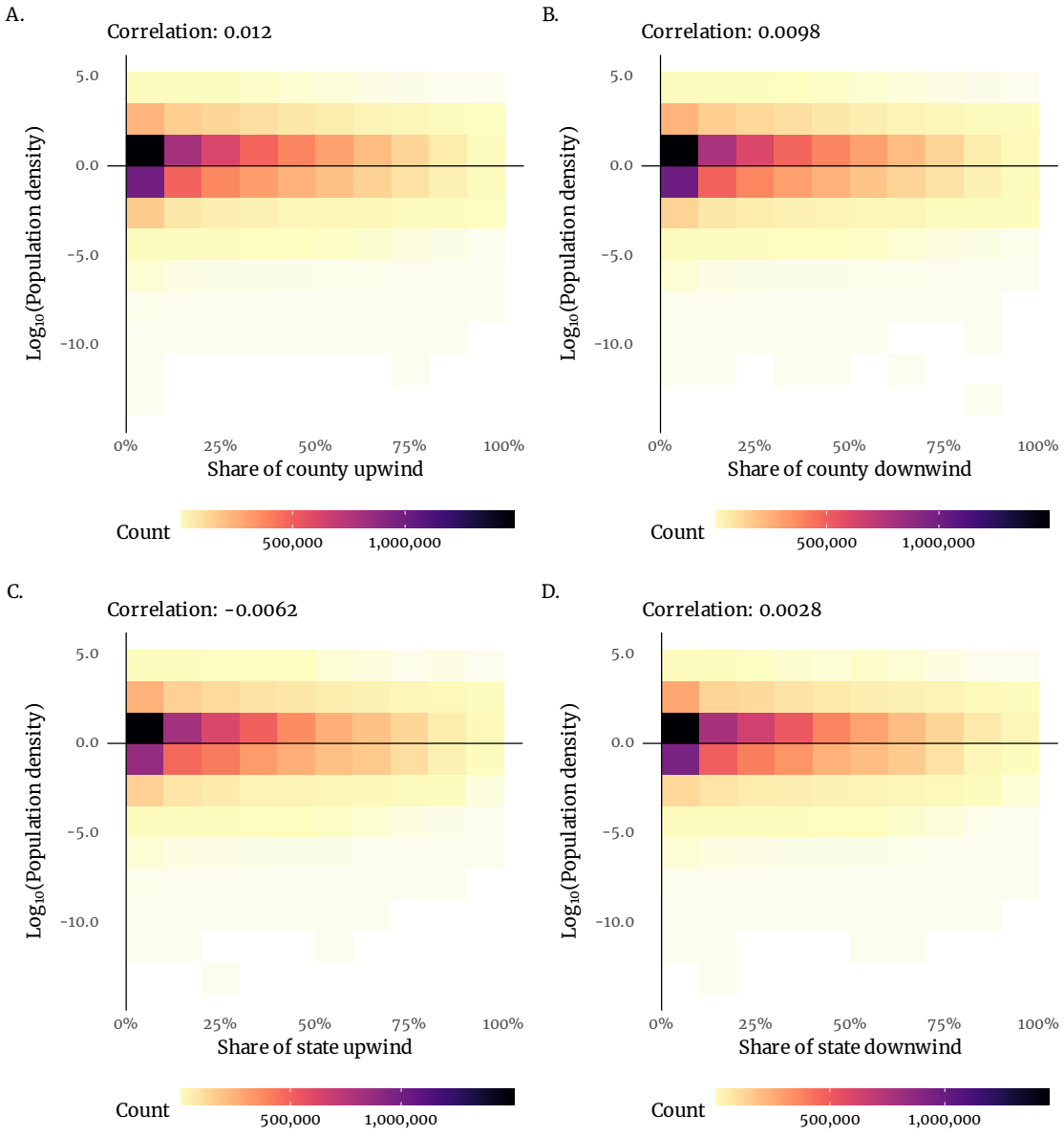


Figure A3. Using a uniform grid that covers the entire contiguous U.S., these figures show the relationship between (1) population density (here, transformed via base-10 log) and (2) the share of the county or state that is upwind or downwind of the grid cell. Due to the large number of grid cells, we use a heat map rather than the typical scatter plot. Note: Because we define *upwind* and *downwind* using a 90-degree angle, very few points in the U.S. are upwind or downwind of more than 75% of their states or counties.

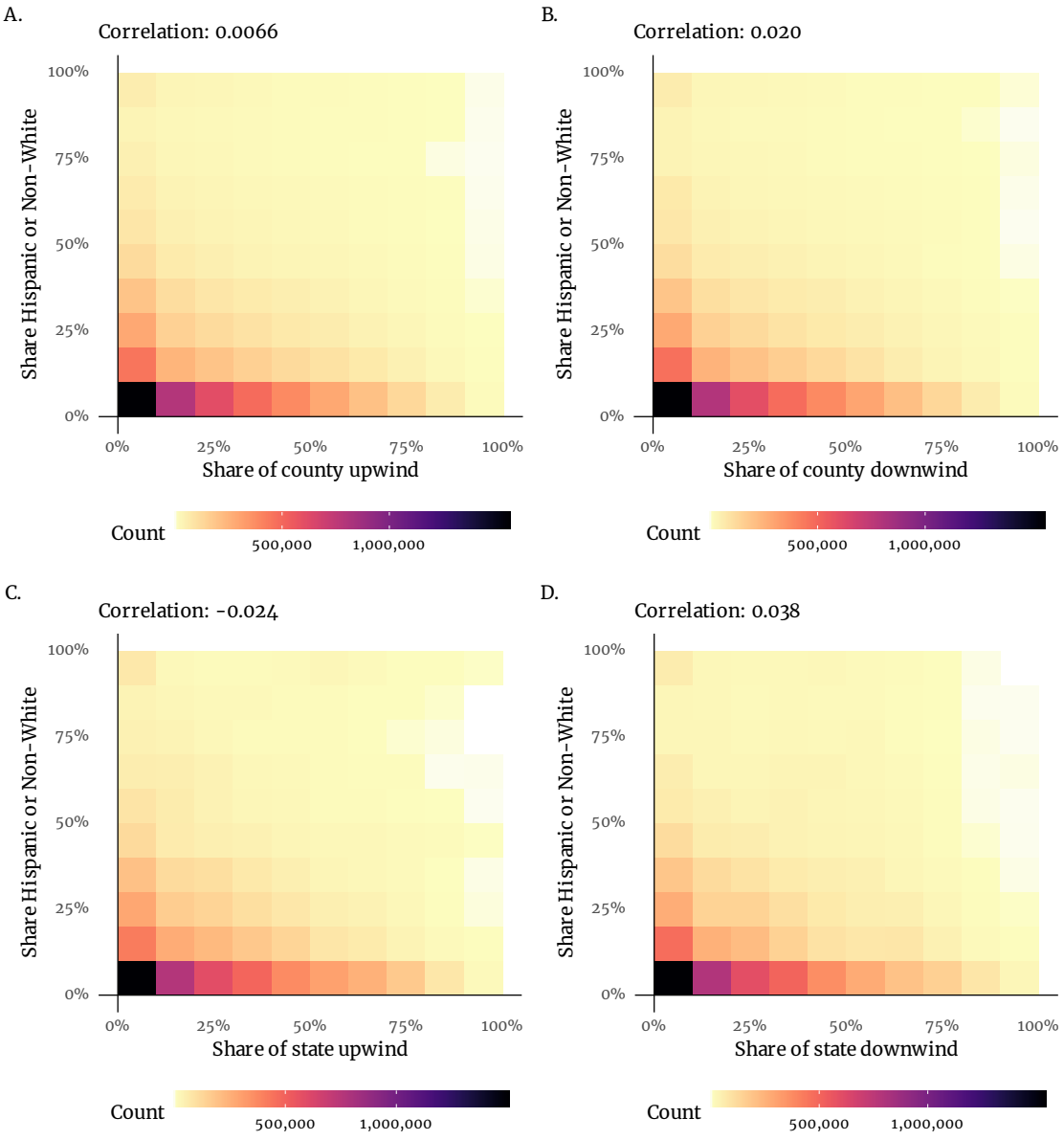


Figure A4. Using a uniform grid that covers the entire contiguous U.S., these figures show the relationship between (1) the share of the cell's population that is Hispanic or non-white and (2) the share of the county or state that is upwind or downwind of the grid cell. Due to the large number of grid cells, we use a heat map rather than the typical scatter plot. Note: Because we define *upwind* and *downwind* using a 90-degree angle, very few points in the U.S. are upwind or downwind of more than 75% of their states or counties.

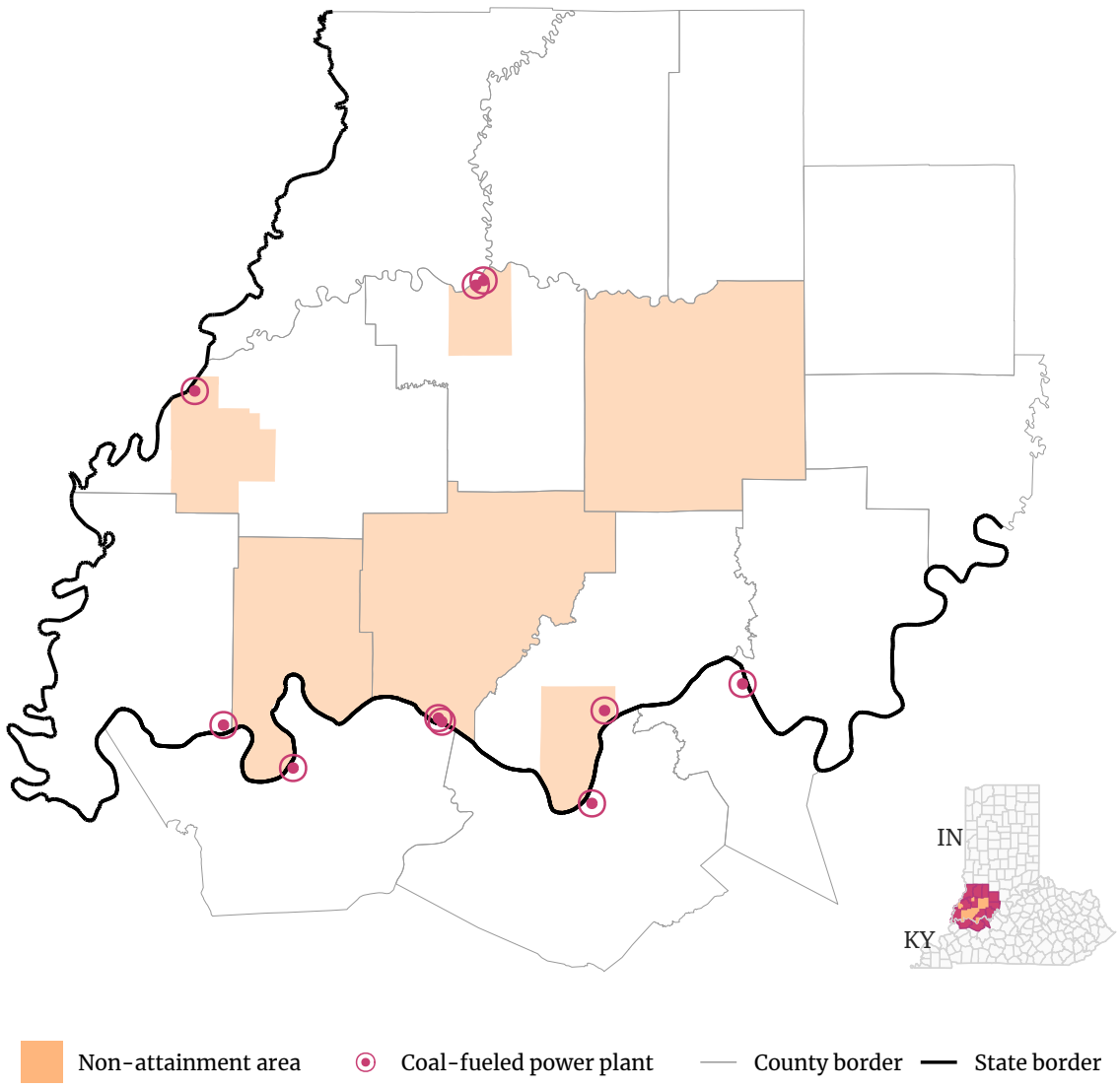


Figure A5. This map illustrates the complexity of the Evansville, Indiana non-attainment area (orange), which covers six counties (3 whole; 3 partial) within Indiana (along its borders with Kentucky to the south and Illinois to the west). Six of the counties form a contiguous area. The remainder of the non-attainment area is formed by islands in three counties that cover nearby coal plants (circled, red dots). As with Figure 9, the non-attainment area is for the 1997 $PM_{2.5}$ standard.

Table A1. Testing EGUs' border distances relative to uniform US grid border distance

	County borders		**State borders**	
Fuel category	K-S test stat.	p-value	K-S test stat.	p-value
Coal	0.248	$< 1 \times 10^{-6}$	0.194	$< 1 \times 10^{-6}$
Gas	0.143	$< 1 \times 10^{-6}$	0.107	$< 1 \times 10^{-6}$
Hydro.	0.477	$< 1 \times 10^{-6}$	0.178	$< 1 \times 10^{-6}$
Solar/Wind	0.037	0.106	0.096	$< 1 \times 10^{-6}$

Testing EGUs' border distances relative to uniform US grid border distance. The columns labeled *K-S test stat.* contain Kolmogorov-Smirnov test statistics testing EGUs' distances to borders against the distribution of distance-to-border built by our uniform national grid. We conduct the tests by EGU fuel category (rows) and administrative level (county and state). The p-values correspond to the adjacent Kolmogorov-Smirnov test statistic.

A.0.4 Appendix: Tables. To test whether the distribution of EGUs' distances to nearest borders is consistent with random sampling from the national grid we employ a simple, non-parametric, Kolmogorov-Smirnov test. The Kolmogorov-Smirnov test is designed to test whether the empirical distribution of a sample statistically differs from a known distribution, which is exactly our goal of this exercise: does the spatial distribution of the EGUs differ from a uniform-national distribution of points?[5] We focus on five major fuel categories: coal, gas, hydropower, and *other renewables* (wind and solar). For each fuel category, we test whether its EGUs' distances to county (or state) borders statistically differ from the distribution of grid points' distances to borders (the grid described above).[6] The results are displayed in A1.

The K-S test resoundingly rejects that null hypothesis that the EGUs' distributions mirror the uniform grid's distribution for each combination of administrative level (county or state) and fuel category (coal, gas, hydro, or solar/wind) with one exception. As one may guess from Figure 3, the one exception is the distance from solar and wind generators to the nearest county border.

[5] Alternatively, the *two-sample* Smirnov test (sometimes called the two-sample Kolmogorov-Smirnov test) tests whether the underlying distributions of two samples statistically differ.

[6] We use `R`'s `base` function `ks.test()`.

This distribution fails to reject the null with a p-value of approximately 0.106 (and a K-S test statistic of 0.037). Except for the solar and wind generators' distances to county borders, we observe overwhelming evidence that EGUs are disproportionately sited near county and state borders—particularly for coal and hydropower units. This observation emphasizes the complexity of monitoring and regulating emissions from EGUs.

Table A2. Robustness to omitting coastal counties: Upwind *vs.* downwind areas for coal and natural gas plants

	(1) **Coal-fueled plants**	(2) **Natural-gas-fueled plants**
Panel a: Siting strategically within **county**		
Count	475	915
Count *strategic*	263	461
Percent *strategic*	55.37%	50.38%
Fisher's exact test of H_o: In-**county** downwind area $\leq$ upwind area		
Under H_o: E[Percent strategic: County] = 50%		
P-value	0.0108	0.4214
Panel b: Siting strategically within **state**		
Count	475	915
Count *strategic*	251	437
Percent *strategic*	52.84%	47.76%
Fisher's exact test of H_o: In-**state** downwind area $\leq$ upwind area		
Under H_o: E[Percent strategic: State] = 50%		
P-value	0.1164	0.9175
Panel c: Siting strategically within **both county *and* state**		
Count	475	915
Count *strategic*	157	230
Percent *strategic*	33.05%	25.14%
Fisher's exact test of H_o: Downwind area $\leq$ upwind area in **county and state**		
Under H_o: E[Percent strategic: County $\wedge$ State] = 25%		
P-value	0.0001	0.4746

Robustness to omitting coastal counties: Upwind *vs.* downwind areas for coal and natural gas plants. By omitting counties on the coast, this table shows the results of Table 1 are not driven by siting in coastal areas. As before, we define a plant's location as "strategic" if the downwind area *within its home county (or state)* is less than its upwind area *within its home county (or state)*. We calculate *downwind* and *upwind* areas based upon 90-degree right triangles with a vertex at the plant pointing up- or down-wind based upon the locally prevailing wind direction. Figure 2 illustrates this calculation. *Sources:* Emissions & Generation Resource Integrated Database (2018) and authors' calculations.

APPENDIX B

CHAPTER 3: APPENDIX

B.0.1 Data and Theory Appendix.

B.0.1.1 Demographic Groups. We drop households living in group quarters and for which the household head is over age 65. A demographic group in our model is defined by the household head's level of education, marital status, age, minority status, and whether there are children in the household. We split education by those that have a college degree. Marital status is defined as either being married or single. Minority status is characterized by whether the individual is white or not. Lastly, very few single individuals in our sample have children. Therefore, we do not differentiate between single households with and without children. In total, this gives us 24 distinct demographic groups.

To better understand which demographic characteristics play the most important roles in determining household-level emissions, we run the following regression of household level emissions on the demographics of a household using data from the 2017 aggregated ACS:

$$Emissions_{ij} = \beta X_i + \gamma_j + \varepsilon_i \tag{B.1}$$

where X_i is the vector of demographic variables, and γ_j is a CBSA level fixed effect.

	Household Emissions (pounds)
White	-183.4***
	(24.23)
College Plus	423.3***
	(18.27)
Old	3,487***
	(24.92)
Married	2,286***
	(23.30)
Has Children	3,378***
	(22.06)
Constant	19,347***
	(33.21)
Observations	2,709,529
R-squared	0.126
CBSA FE	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B1. Regression estimates of (B.1)

Being married, having children, and having an older household head are associated with large values of emissions, while the other demographic variables only play a small role in dictating a household's carbon emissions.

B.0.1.2 Energy Prices. We obtain data on average residential electricity, natural gas, and fuel oil prices by state for 1990, 2000, 2010, and 2017 from the Energy Information Association. For each energy type and year, we assign the average residential price to all CBSAs within a state. Furthermore, for electricity prices, we use the prices given from "full-service providers." Fuel oil prices are reported at a weekly level. We average across weeks to obtain yearly average fuel oil prices. Additionally, as fuel oil is used primarily in the northeast, many states do not report average prices. For states that do not have fuel oil prices in the EIA's dataset, we assign the yearly average across all states that do have prices.

B.0.1.3 NERC Regions. We calculate the emissions factor for each region as a weighted average of the average CO_2 emissions rate in each NERC

region. We weight the average by each plant's total yearly MWh generation as a fraction of the total MWh generation in the region. Figure B1 is a map of the NERC regions for the contiguous United States with the conversion factors.

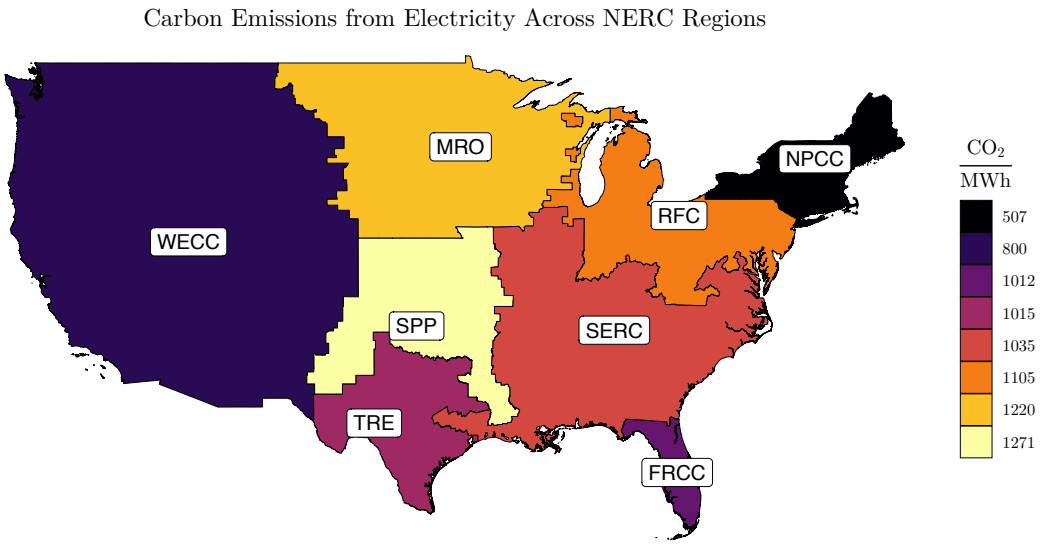


Figure B1. Map of NERC region with regional conversion factors. In the model, there is an additional NERC region for Hawaii (HICC) with an emissions factor of 1522.10.

B.0.1.4 Correction for Rented Homes and Multi-Family Homes.

One concern is that rented homes and multi-family homes are less likely to pay for energy themselves and the proportion of renters and multi-family homes varies across cities. As the ACS and Census contain information only on energy costs, not energy usage, this may lead us to understate under usage in cities with high amounts of renters of residents in multi-family homes. Similar to Glaeser and Kahn (2010), we correct for this using data from the 2015 Residential Energy

Consumption Survey (RECS), which contains data on energy *usage* for a sample of over 5,0000 households.

We use these data to estimate the following regression which compares the energy usage of renters, and those who live in multi-family housing, to owners of single-family housing:

$$\log(E_i^m) = \beta_{MF}^m MultiFamily_i + \beta_{Rent}^m Rent_i + controls + e_i^m \tag{B.2}$$

where controls include controls for household size, number of children, age of household head, whether the household head is white, and division dummies. We then use the coefficients β_{MF}^m and β_{Rent}^m to impute energy usage for households who are renters and who live in multi-family housing. For example, if we estimate that owners of single-family housing in San Francisco use 8 MWh of electricity and estimate $\beta_{MF}^m = .1$, we would impute that owners of multi-family housing use $8 \times 1.1 = 8.8$ MWh of electricity. Finally, we estimate the fraction of renters of single-family housing, and multi-family housing, owners of single-family housing, and owners of multi-family housing, using data from the ACS and Census, and calculate the predicted usage as the weighted average of the estimated predicted usage of owners of single-family housing, and the imputed usage of the other three groups.

B.0.1.5 Fuel Consumption and Population. We assume that the marginal benefit of fuel consumption is exogenous to the population of a given city. As a simple test of the relationship between population and energy consumption, we estimate:

$$\log(\hat{E}_j^m + 1) = \alpha^m + \alpha_1^m \log(\text{Population}_j) + \varepsilon_j \tag{B.3}$$

where $m \in \{Elec, Gas, Fuel\}$ and $\hat{E}_j^m$ is the predicted per-household, selection-corrected energy consumption of type m in city j. Given the selection-correction usages predict zero fuel consumption in certain CBSAs, we use $\log(\hat{E}_j^m + 1)$. The results presented here are not sensitive to this choice. Table B2 provides estimates for the parameters in equation (B.3).

	Dependent variable:		
	Electricity Consumption (MwH)	**Gas Consumption** (1000 ft^3)	**Fuel Consumption** (gal)
log(Population)	−0.012	−0.413	0.049
	(0.136)	(0.310)	(0.036)
Constant	4.276**	8.428**	2.232***
	(1.809)	(4.088)	(0.484)
Observations	70	70	70
Note:			*p<0.1; **p<0.05; ***p<0.01

Table B2. Regression estimates of fuel consumption on population. Heteroskedastic robust standard errors are in parentheses. As the selection-correction usages predict zero fuel consumption in certain CBSAs, we use $\log(\hat{E}_j^m + 1)$. Each observation is a CBSA.

The coefficients on all of the regressions for the energy consumption variables are statistically insignificant. This suggests population increases do not lead to significant changes in the average benefits of energy usage.

B.0.1.6 Equilibrium Definition. In this environment, an equilibrium is characterized by household and firm optimization, and market clearing in the housing and labor markets.[1]

More specifically, as we have shown in Section 3.4.1, given prices, household i's optimal choice maximizes utility.

Household optimization defines housing demand, energy demand, and labor supply. Housing demand in a city j is given by the sum of housing demand across all agents living in that city. We can write this as

$$H_j^D = \sum_d N_{jd} \frac{\alpha_d^H I_{jd}}{R_j \boldsymbol{\alpha}_{jd}}, \tag{B.4}$$

where, as before, N_{jd} is the total number of workers of demographic d who choose to live in city j, and where we allow D and $\mathcal{S}$ superscripts to denote demand and supply quantities, respectively. Similarly, energy demand is the sum of energy

[1] In Section 3.7.2 we consider the case when energy prices are determined in equilibrium. In this case, an equilibrium is also defined by market clearing in the energy markets.

demand of all individuals living in a city:

$$X_j^{mD} = \sum_d N_{jd} \frac{\alpha_{jd}^m I_{jd}}{P_j^m \boldsymbol{\alpha}_{jd}}. \tag{B.5}$$

Labor supply is the sum of efficiency units of labor supplied by all agents of a given skill level in city j. For more-skilled (college-educated) workers:

$$S_j^{\mathcal{S}} = \sum_{d' \in d^S} N_{jd} \ell_{d'}$$

and for less-skilled (non-college-educated) workers:

$$U_j^{\mathcal{S}} = \sum_{d' \in d^U} N_{jd} \ell_{d'}$$

where d^S and d^U are the sets of demographic groups with a college degree and without a college degree, respectively.

Labor demand for skilled and unskilled workers are implicitly defined by (C.3), the first-order conditions of the production firms.

Housing supply is given by (3.14).

Finally, an equilibrium is defined by the two market clearing conditions:

1. Housing Market Clearing: $H_j^{\mathcal{S}} = H_j^D$, for all cities, j.
2. Labor Market Clearing: $S_j^{\mathcal{S}} = S_j^D$ for skilled workers and $U_j^{\mathcal{S}} = U_j^D$ for unskilled workers in all cities.

B.0.1.7 Hedonic Rents. A major concern about producing a measure of housing costs across CBSA's is that they should reflect the user cost of housing. To accommodate this, we use data only on renters as home prices reflect both the current cost and expected future costs. Secondly, it is difficult to compare housing units across CBSA's. Thus, we estimate hedonic regressions of log gross rent on a set of housing characteristics and CBSA fixed effects. Specifically, we control for the number of units in the structure containing the household, the number of bedrooms, the total number of rooms, and household members per room. To generate the rent index, we utilize the predicted values from the hedonic regressions, holding constant the set of housing characteristics and CBSA fixed effects.

B.0.1.8 Estimation: Production Parameters . Let $x \in \{s, u\}$ index worker education/skill levels. Income for workers of demographic d living in

location j is $I_{jd} = W_{jx}\ell_d$, where ℓ_d is the number of efficiency units supplied by workers from demographic group d.

We specify efficiency units as the demographic-specific probability of being employed multiplied by the productivity conditional on being employed. We therefore write

$$\ell_d = E_d\hat{\ell}_d$$

where E_d is the national employment-to-population ratio of workers in demographic group d.

We parameterize $\hat{\ell}_d$ as

$$\log(\hat{\ell}_d) = \beta_x^1 \text{White}\,(d) + \beta_x^2 \text{Over35}\,(d)$$

where $\text{White}\,(d)$ is an indicator variable indicating workers of demographic group d are white and $\text{Over35}\,(d)$ indicates workers of demographic d are over age 35. Therefore $\hat{\ell}_d$ of nonwhite workers below age 35 is normalized to one.

Conditional on working, the log-income of workers of demographic group d and skill level x living in city j is given by

$$\log\left(I_{jd}\right) = \log\left(W_{jx}\right) + \beta_x^1 \text{White}\,(d) + \beta_x^2 \text{Over35}\,(d)\,.$$

We therefore estimate the city-level wage rates and parameters of the efficiency unit of labor using the following individual-level income regression conditional on working:

$$\log I_{ijd} = \gamma_j^x + \hat{\beta}_x^1 \text{White}\,(d) + \hat{\beta}_x^2 \text{Over35}\,(d) + \varepsilon_{ij}$$

where I_{ijd} is the income level of individual i, γ_j^x is a city- by-skill-level fixed effect which is an estimate of $\log\left(W_{jx}\right)$, and ε_{ij} is an individual-level error term.

The remaining unknown parameters of the production function are the elasticity of substitution, ς, the vector of city level total factor productivities, A_j, and the vector of factor intensities, θ_j. We calibrate the elasticity of substitution, $\varsigma = 2$.

Note that the log-wage-ratio in city j is given by

$$\log\left(\frac{W_{js}}{W_{ju}}\right) = -\frac{1}{\varsigma}\log\left(\frac{S_j}{U_j}\right) + \log\left(\frac{\theta_j}{1-\theta_j}\right).$$

As wage levels, labor quantities, and the elasticity of substitution, ς, are already known, the factor intensities θ_j can be solved by using the above equation.

The final set of parameters are the total factor productivities, A_j. These are chosen so that wage levels are equal to those in the data.

B.0.1.9 Calibration: Housing Supply. We know that total demand for housing in city j is given by:

$$H_j = \sum_d N_{jd} \frac{\alpha_d^H I_{jd}}{R_j \boldsymbol{\alpha}_{jd}}, \tag{B.6}$$

where N_{jd} is the total number of workers of demographic d living in city j. Plugging this equation for housing demand into the housing supply curve and rearranging yields the following reduced-form relationship:

$$\log(R_j) = \frac{k_j}{1+k_j} \log(\sum_d N_{jd} \frac{\alpha_d^H I_{jd}}{\boldsymbol{\alpha}_{jd}}) + \zeta_j. \tag{B.7}$$

where $\zeta_j = \frac{\log z_j}{1-k_j}$.

Saiz (2010) estimates the role of physical and regulatory constraints in determining the role of local housing supply elasticities by using labor demand shocks and instruments for housing demand. As in this paper, we set ψ_j^{WRI} equal to the log of the Wharton Regulation Index plus 3, and use Saiz's measure of the unavailable land share (due to geography) for ψ_j^{GEO}. We calibrate ν_1, ν_2 and ν_3 based on the estimates in Saiz (2010).[2] We then choose the values of ζ_j to match the rent levels observed in the data.

B.0.1.10 InMAP and Derivation of the SR matrix. In this section, we provide a broad overview of InMAP and our process for deriving our pollution-transfer matrix that maps electricity generation in a given NERC region to ambient concentration in a given CBSA.

InMAP and ISRM The Intervention Model for Air Pollution (InMAP, Tessum et al. (2017)), is a reduced-complexity air transport model that allows users to estimate how changes in emissions impact concentration nationally. InMAP takes into account atmospheric chemistry, local meteorological conditions (*i.e.* wind), and variables regarding the point of emission (such as stack height and velocity at which the particle was emitted). To estimate particulate-matter concentration, InMAP uses data on emissions of primary $PM_{2.5}$ and secondary pollutants that react with gasses in the air and form $PM_{2.5}$. The secondary

[2] Specifically, we use the estimates from Column (4) of Table III in Saiz (2010), as these are the closest to our specification. Given that the estimate of the coefficient on the interaction between housing supply constraints is quite similar across specifications in Saiz (2010), we do not suspect that our results will be sensitive to the specific estimates we choose.

pollutants used by InMAP are Volatile Organic Compounds (VOC), Nitrogen Oxides (NOx), Ammonia, (NH3), and Sulfur Oxides (SOx). InMAP estimates concentrations for grid cells that vary in size by area population; for urban areas, the grid cells are small, and for rural areas, they are large—expedient for making the model more computationally tractable.

In Goodkind et al. (2019) InMAP is run over 150,000 times to obtain average transfer coefficients for each grid cell—resulting in the InMAP SR matrix (ISRM). Furthermore, ISRM has 3 "height" layers for each of the grid cell; 0 to 57m, 57-379m, and >379m. The Python code provided by Goodkind et al. (2019) uses information about a plant's stack height, and the velocity at which the particle, is emitted to estimate into which of these three height layers the plant's plume of the emissions will fall, at any given time and distance

Derivation of the "transfer matrix" Let $\delta_{R,j}^{PM_{2.5}}$ be the factor that converts electricity produced in region R into concentrations of $PM_{2.5}$ in city j. We calculate $\delta_{R,j}^{PM_{2.5}}$ as an emissions-weighted average of conversion factors for each individual power plant in region R. Let s index an individual source (power plant) and $S(R)$ be the set of all sources within NERC region R. Let $\delta_{s,j}^{PM_{2.5}}$ be the conversion factor between electricity production at source s and $PM_{2.5}$ concentration in city j, given by:

$$\delta_{s,j}^{PM_{2.5}} = \frac{PM_{2.5,s,j}}{x_s^{elec}}$$

where $PM_{2.5,s,j}$ is the ambient air pollution in city j originating form source s (in NERC region R) and x_s^{elec} is the total electricity produced by source s. Then we compute $\delta_{R,j}^{PM_{2.5}}$ as the emissions-weighted average of these source-level conversion factors:

$$\delta_{s,j}^{PM_{2.5}} = \frac{\sum_{s \in S(R)} x_s^{elec} \delta_{s,j}^{PM_{2.5}}}{\sum_{s \in S(R)} x_s^{elec}}$$

Plugging $\delta_{s,j}^{PM_{2.5}}$ into $\delta_{s,j}^{PM_{2.5}}$ yields:

$$\delta_{s,j}^{PM_{2.5}} = \frac{\sum_{s \in S(R)} PM_{2.5,s,j}}{\sum_{s \in S(R)} x_s^{elec}}$$

where $\sum_{s \in S(R)} PM_{2.5,s,j}$ is the average ambient $PM_{2.5}$ concentration in city j, originating from region R and x_s^{elec} is total electricity production in region R.[3]

[3]In practice, we calculate $PM_{2.5,s,j}$ as population-weighted averages within a CBSA.

$\sum_{s \in S(R)} PM_{2.5,s,j}$ is estimated via ISRM by setting pollutant emissions in all regions $R' \neq R$ to zero, and computing the resulting ambient concentration in all cities for emissions from just region R. We note that ISRM has coefficients only for the contiguous United States; thus for Hawaii, we set all transfer coefficients to zero. In the model, this means that the level of particulate matter in Honolulu is fixed and no particulate matter from Honolulu is transferred to the rest of the United States.

B.0.1.11 Derivation of Mean Utility Estimating Equation.

Mean utility is given by

$$\mu_{jdt} = \frac{\left(1 + \alpha_d^H + \sum_m \alpha_{jd}^m\right)}{\sigma_d} \log I_{jdt} - \frac{\alpha_d^H}{\sigma_d} \log R_{jt} - \sum_m \frac{\alpha_{jd}^m}{\sigma_d} \log P_{jt}^m + \hat{\xi}_{jdt}.$$

Recall that we have defined $\tilde{\alpha}_{jd}^m = \frac{\alpha_{jd}^m}{1+\alpha_d^H+\sum_m \alpha_{jd}^m}$. Therefore, it is fairly straightforward to show that

$$\sum_{m'} \alpha_{jd}^{m'} = \frac{\sum_{m'} \tilde{\alpha}_{jd}^{m'} \left(1 + \alpha_d^H\right)}{1 - \sum_{m'} \tilde{\alpha}_{jd}^{m'}}$$

and therefore that

$$\alpha_{jd}^m = \frac{\tilde{\alpha}_{jd}^m \left(1 + \alpha_d^H\right)}{1 - \sum_{m'} \tilde{\alpha}_{jd}^{m'}}.$$

Plugging these identities into the mean utility expression yields

$$\mu_{jdt} = \frac{\left(1 + \alpha_d^H + \frac{\sum_m \tilde{\alpha}_{jd}^m \left(1+\alpha_d^H\right)}{1-\sum_m \tilde{\alpha}_{jd}^m}\right)}{\sigma_d} \log I_{jdt} - \frac{\alpha_d^H}{\sigma_d} \log R_{jt} - \frac{\left(1 + \alpha_d^H\right)}{1 - \sum_{m'} \tilde{\alpha}_{jd}^{m'}} \sum_m \frac{\tilde{\alpha}_{jd}^m}{\sigma_d} \log P_{jt}^m + \hat{\xi}_{jdt}.$$

where can rearrange this to yield

$$\mu_{jdt} = \frac{\left(1 + \alpha_d^H\right)}{\sigma_d} \frac{\log I_{jdt} - \sum_m \tilde{\alpha}_{jd}^m \log P_{jt}^m}{1 - \sum_m \tilde{\alpha}_{jd}^m} - \frac{\left(\alpha_d^H\right)}{\sigma_d} \log R_{jt} + \xi_{jdt}.$$

Defining $\tilde{I}_{jdt} = \frac{\log I_{jdt} - \sum_m \left(\tilde{\alpha}_{jd}^m \log P_{jt}^m\right)}{1-\sum_m \tilde{\alpha}_{jd}^m}$, $\beta_d^w = \frac{\left(1+\alpha_d^H\right)}{\sigma_d}$ and $\beta_d^r = \frac{\left(\alpha_d^H\right)}{\sigma_d}$, we arrive at (3.22):

$$\mu_{jdt} = \beta_d^w \tilde{I}_{jdt} + \beta_d^r \log R_{jt} + \xi_{jdt}.$$

B.0.2 Results Appendix: For Online Publication Only. Table B3 compares various specifications of the selection control function in estimating (3.3), which we use to generate selection-corrected predicted emissions. For each

specification, we estimate the predicted emissions in each CBSA. Then we calculate the population-weighted mean, standard deviation, and correlation with the Wharton Regulation Index across CBSAs.

	Mean	Standard Deviation	Correlation w/ Land Use Restrictions
I. No Selection Correction			
a. Raw Means	24946	5729	-0.18
b. OLS	23711	5526	-0.21
II. Selection Correction			
a. Choice Location and 3 Biggest States			
i. Linear Choice, Linear States, Choice ×State Interactions	25518	5740	-0.28
ii. Linear Choice, Quadrartic States, Choice ×State Interactions	26815	6622	-0.22
iii. Linear Choice, Linear States, No Interactions	23934	5652	-0.28
iv. Linear Choice, Quadrartic States, No Interactions	24107	5488	-0.28
v. Quadrartic Choice, Linear States, Choice ×State Interactions	33185	12114	-0.21
vi. Quadrartic Choice, Quadrartic States, Choice ×State Interactions	32300	11747	-0.24
vii. Quadrartic Choice, Linear States, No Interactions	32581	11328	-0.27
viii. Quadrartic Choice, Quadrartic States, No Interactions	31199	10508	-0.23
b. Choice Location and 5 Biggest States			
i. Linear Choice, Linear States, Choice ×State Interactions	27635	6801	-0.17
ii. Linear Choice, Quadrartic States, Choice ×State Interactions	27339	7073	-0.19
iii. Linear Choice, Linear States, No Interactions	23940	5654	-0.28
iv. Linear Choice, Quadrartic States, No Interactions	24203	5483	-0.26
v. Quadrartic Choice, Linear States, Choice ×State Interactions	32277	11283	-0.17
vi. Quadrartic Choice, Quadrartic States, Choice ×State Interactions	30837	11252	-0.20
vii. Quadrartic Choice, Linear States, No Interactions	32679	11563	-0.28
viii. Quadrartic Choice, Quadrartic States, No Interactions	31150	10662	-0.23
c. Choice Location and Birth States			
i. Linear Choice, Linear Birth State	25327	6147	-0.32
ii. Linear Choice, Quadrartic Birth State	25467	6202	-0.32
iii. Quadratic Choice, Linear Birth State	30878	9608	-0.22
iv. Quadratic Choice, Quadrartic Birth State	30416	9308	-0.23
d. Controls for Climate in Birth State			
i. Linear Choice, Linear States, Choice ×State Interactions	23418	5667	-0.12
ii. Linear Choice, Quadrartic States, Choice ×State Interactions	25281	6524	-0.19
iii. Linear Choice, Linear States, No Interactions	20369	6609	-0.15
iv. Linear Choice, Quadrartic States, No Interactions	22764	6262	-0.25

Table B3. Comparisons of various specifications of selection control function.

B.0.2.1 Comparisons of Specification of Control Function.

Panel I gives the predicted emissions without any selection correction. Row I.a gives simply the mean emissions without including any demographic controls and I.b estimates (3.3) without any selection correction but including demographics controls.

Panel II includes the results with different specification of the control function $M(\cdot)$. Subpanel II.a present estimates in which $M(\cdot)$ is a function of the probability of choosing the state in question, and the probabilities of choosing the three largest states. Row II.a.i consitute our preferred specification, where the selection control function consists of the probability of choosing the state in question entering linearly, the probabilities of choosing the three largest states entering linearly, and the interactions between (a) the probability of choosing the state in question and (b) each of the three largest-state choice probabilities.

The following rows of the table give alternative specifications in which state-choice probabilities enter as a quadratic, in which the probability of choosing the state in question also enters as a quadratic, and the interaction terms are omitted. Subpanel II.b considers an analogous specification except where we include the probabilities of choosing the 5 largest states. Finally, Subpanel II.c considers a control function written as a function of choosing the state in question and choosing the individual's birth state.

Subpanel II.d compare estimates when we also include controls for the average yearly temperature in the state of birth. These specifications are otherwise identical to those in II.a.i through II.a.iv, in which we include controls for the three largest states by population, and the probably of choosing the state in question.

In general, the estimates are relatively similar across specifications. The exception is when the choice probability for the state in question enters as quadratic. In these cases, the standard deviation of the predicted emissions increases. As mentioned before, estimating the intercept of the energy usage equation relies on extrapolating the control function to $P_{is(j)} = 1$. For smaller states, the probability of choosing the state in question is farther from one, so this extrapolation becomes more sensitive to the choice of the control function.

B.0.2.2 Additional Summary Statistics: No Selection Correction.

In this section, we replicate our Table 2 and our main descriptive scatterplots without demographic controls and without the selection correction.

Table B4 gives estimates of energy usage and emissions by CBSA, where estimates of energy use are given simply by the unconditional mean for households living in the CBSA. There are no controls for demographics, and no selection-correction is implemented.

CBSA	Rank	Emissions (1000 lbs)	Gas Emissions (1000 lbs)	Fuel Emissions (1000 lbs)	Electricity Use (MwH)	Electricity Conversion (1000 lbs/MwH)	Electricity Emissions (1000 lbs)
Lowest							
Honolulu, HI	1	12.83	0.47	0.07	8.08	1.52	12.29
Oxnard, CA	2	12.85	5.80	0.17	8.61	0.80	6.89
Riverside, CA	3	13.64	5.59	0.17	9.85	0.80	7.88
Los Angeles, CA	4	14.41	6.06	0.09	10.32	0.80	8.26
San Diego, CA	5	14.87	6.42	0.23	10.27	0.80	8.22
Sacramento, CA	6	15.84	7.28	0.40	10.20	0.80	8.16
Middle							
Atlanta, GA	33	25.24	6.46	0.17	17.97	1.04	18.61
Pittsburgh, PA	34	25.77	11.43	1.35	11.74	1.11	12.98
Akron, OH	35	25.85	12.05	0.58	11.95	1.11	13.21
Birmingham, AL	36	26.10	5.42	0.17	19.81	1.04	20.51
Virginia Beach, VA	37	26.19	6.12	0.71	18.70	1.04	19.36
Houston, TX	38	26.37	4.62	0.08	21.35	1.01	21.67
Highest							
Oklahoma City, OK	65	32.29	8.26	0.20	18.76	1.27	23.84
Detroit, MI	66	32.48	18.72	0.36	12.12	1.11	13.40
Philadelphia, PA	67	33.32	11.39	3.12	17.02	1.11	18.81
Memphis, TN	68	34.45	8.37	0.19	25.01	1.04	25.89
Milwaukee, WI	69	35.22	16.71	0.52	16.28	1.11	17.99
Omaha, NE	70	35.98	15.79	0.28	16.31	1.22	19.91

Table B4. Predicted CBSA level CO_2 emissions by fuel type for the six lowest-emissions cities, the six median cities, and the six highest-emissions cities in 2017. The third column ("Emissions") shows the **unconditional mean** CO_2 emissions from natural gas, fuel oil and electricity for the CBSA. The next two columns show emissions from gas and fuel oil respectively, which are equal to predicted usage multiplied by the appropriate emissions factor. The last three columns show predicted electricity usage, the electricity emissions factor, and predicted electricity emissions, equal to predicted electricity usage multiplied by the emissions factor.

The next figures are replicates of Figures 11 and 12—without selection-corrected energy usage. Figure B2 plots household carbon emissions against the Wharton Index. In the scatterplot on the left, we predict household energy use with a simple OLS regression that controls for demographic groups. In the scatterplot on the right, we predict household energy use with CBSA-level means. Overall, the pattern is qualitatively similar regardless of the specification; California cities have low household carbon emissions and relatively stringent land-use restrictions.

Figure B3 plots household natural gas usage against January temperature and electricity usage against August temperature. In the two scatterplots in the top row, we predict household energy use with a simple OLS regression that controls for demographic groups. In the scatterplots on the bottom row, we predict household energy use with CBSA-level means.

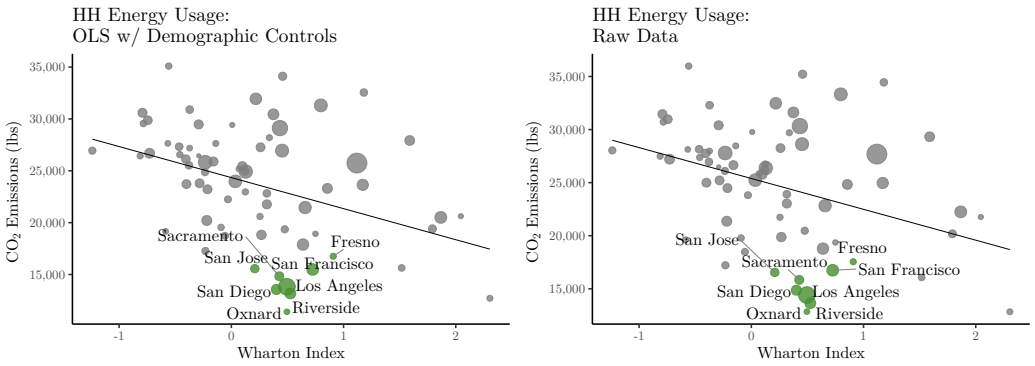


Figure B2. Additional scatterplots: in which CO_2 emissions plotted against the Wharton Index. An observation is a CBSA; a larger circle represents a larger population. California cities are in green.

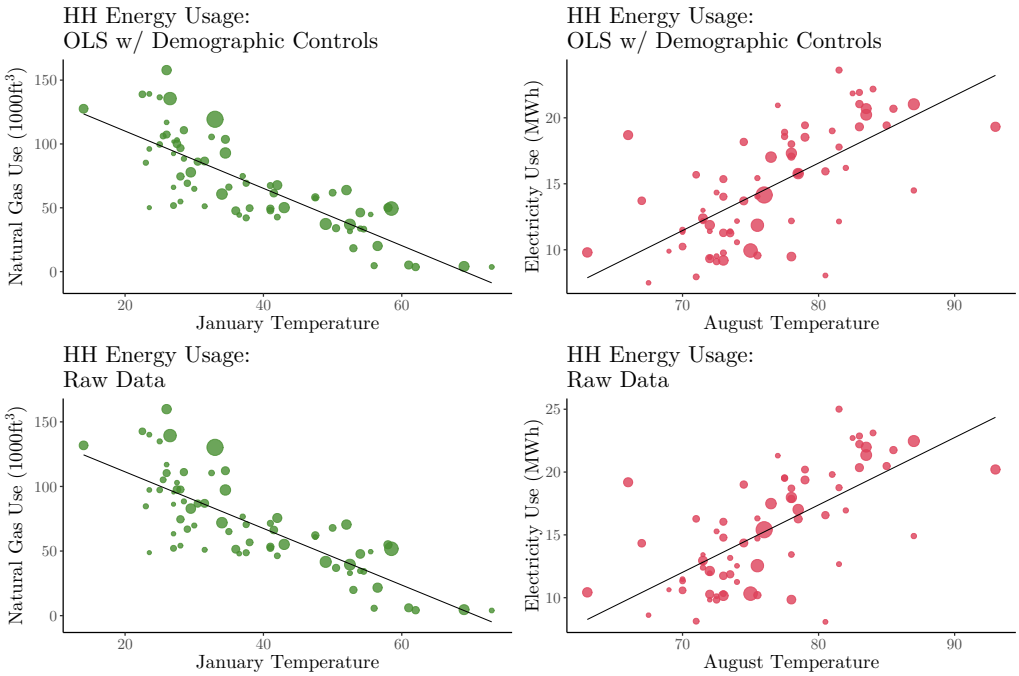


Figure B3. Additional scatterplots in which natural gas and electricity use are plotted against January and August temperatures, respectively. January and August temperature refers to the midpoint between average daily highs and lows for the given month. An observation is a CBSA; a larger circle represents a larger population.

B.0.2.3 $PM_{2.5}$: Additional Results . This section provides additional summary information about $PM_{2.5}$. Figure B4 plots the distribution of **total** $PM_{2.5}$ concentrations across cities and Figure B5 the estimated contribution of household electricity to total $PM_{2.5}$.

From Figure B4, there are a few key takeaways. First, the histogram demonstrates considerable variation across CBSAs in terms of total $PM_{2.5}$. Second, California cities are relatively dispersed throughout the distribution—some are relatively clean, while others have high concentrations of $PM_{2.5}$.

Next, Figure B5 with the city-level ratios of household electricity contribution to total $PM_{2.5}$ illustrates two things. Overall, household electricity contributes fairly little to overall $PM_{2.5}$. Second, the amount by which household electricity use contributes to total $PM_{2.5}$ varies across cities; Portland gets nearly zero percent of its particulate matter emissions from electricity, while Dallas gets roughly 6.5%.

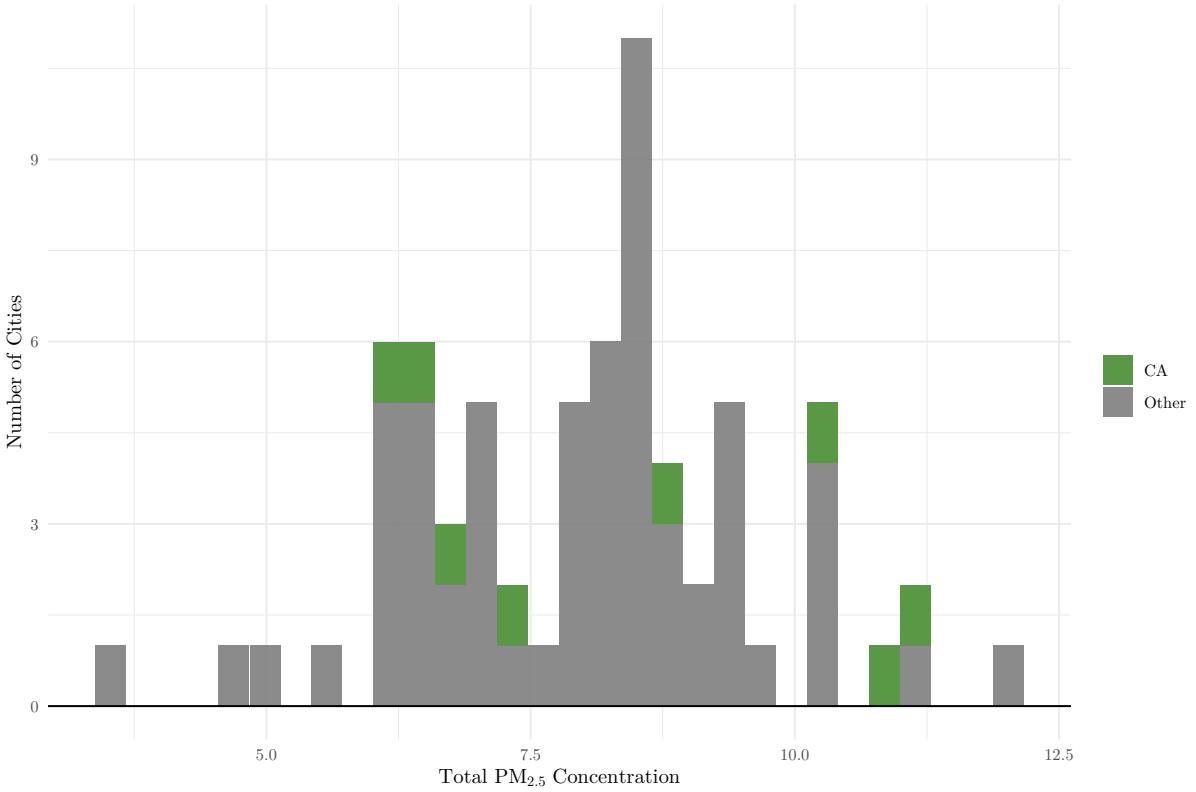


Figure B4. The distribution of 2017 mean $PM_{2.5}$ across CBSAs in our sample.

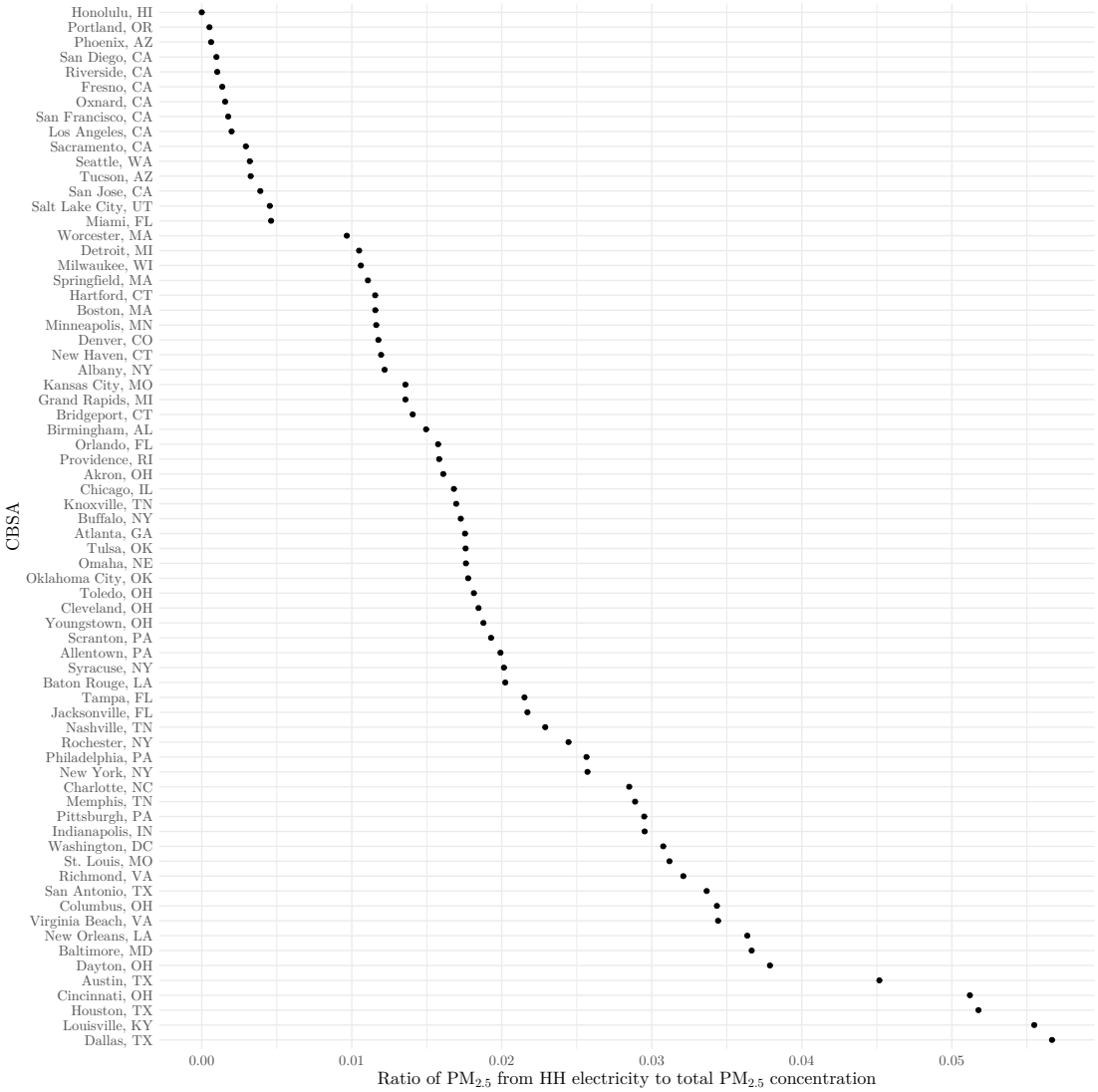


Figure B5. This figure plots the ratio of $PM_{2.5}$ coming from electricity to total $PM_{2.5}$ as measured by the EPA.

Next, Figure B6 plots a histogram of $PM_{2.5}$ changes from the baseline when we set land-use restrictions to the level faced by the median urban household in all cities.

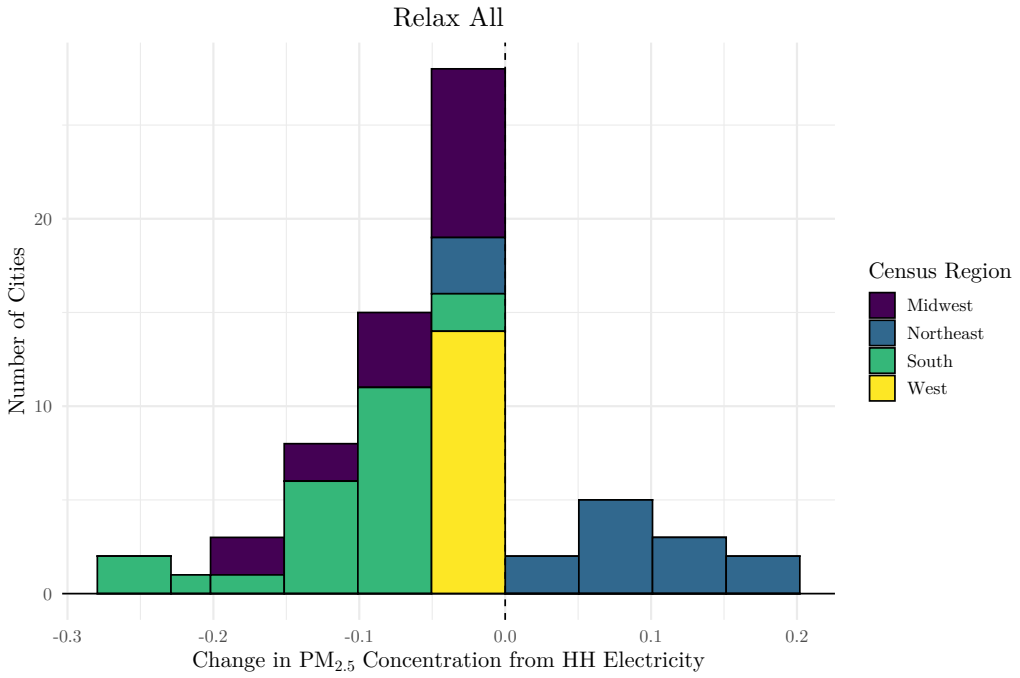


Figure B6. Histogram of CBSA level differences in particulate matter concentration from electricity relative to the case where land-use restrictions in all cities.

B.0.2.4 Robustness of Main Parameter Estimates. Table B5 gives estimates that vary by the age of the household head. The first three columns give estimates for which the head of the households is under 35 years old for (1) single households, (2) married households without children, and (3) married households with children. The next three columns present estimates for which the head of the households is over 35 years old for (4) single households, (5) married households without children, (6) and married households with children. The estimates of β^w and β^r are slightly larger in magnitude for households with older household heads, conditional on marital status and the presence of children.

	Head under 35 years experience			Head over 35 years experience		
		Married			Married	
	Single	No Children	With Children	Single	No Children	With Children
β^w: Adjusted Income	13.37	9.96	7.35	16.81	13.45	7.37
	(3.48)	(2.62)	(2.16)	(4.45)	(3.58)	(2.01)
β^r: Rent	-7.27	-5.02	-5.13	-10.77	-8.76	-4.56
	(2.99)	(2.27)	(1.92)	(3.79)	(3.09)	(1.74)
σ: Idiosyncratic Component	0.16	0.20	0.45	0.17	0.21	0.36
	(0.04)	(0.05)	(0.20)	(0.06)	(0.07)	(0.12)
α^H: Housing Parameter	1.19	1.02	2.30	1.79	1.87	1.62
	(0.56)	(0.51)	(1.37)	(0.83)	(0.90)	(0.79)
Cragg-Donald Wald F Statistic	3.92	4.04	4.27	4.16	4.24	4.29

Table B5. Parameter estimates that vary by age. Standard errors in parentheses.

Table B6 gives estimates using alternative instrumental variables. The first panel presents our baseline estimates. Panel II uses estimates for a model that employs the measure of land use availability from Saiz (2010) in place or the Wharton Land Use Index. Panel III presents estimates when we use both measures as instruments.

I. Baseline Estimates			
		Married	
	Single	No Children	With Children
β^w: Adjusted Income	15.09	11.72	7.33
	(2.80)	(2.19)	(1.47)
β^r: Rent	-9.03	-6.90	-4.82
	(2.40)	(1.89)	(1.29)
σ: Idiosyncratic Component	0.17	0.21	0.40
	(0.03)	(0.04)	(0.11)
α^H: Housing Parameter	1.49	1.44	1.92
	(0.48)	(0.48)	(0.73)
Cragg-Donald Wald F Statistic	8.09	8.28	8.57
II. Available Land Instrument			
		Married	
	Single	No Children	With Children
β^w: Adjusted Income	19.39	18.36	16.14
	(4.66)	(4.52)	(4.07)
β^r: Rent	-9.99	-9.06	-9.22
	(3.00)	(2.88)	(2.62)
σ: Idiosyncratic Component	0.11	0.11	0.14
	(0.03)	(0.03)	(0.04)
α^H: Housing Parameter	1.06	0.98	1.33
	(0.28)	(0.26)	(0.36)
Cragg-Donald Wald F Statistic	5.41	5.14	5.05
III. Both Instruments			
		Married	
	Single	No Children	With Children
β^w: Adjusted Income	15.42	12.40	9.08
	(2.36)	(1.90)	(1.46)
β^r: Rent	-8.43	-6.46	-5.61
	(1.73)	(1.39)	(1.09)
σ: Idiosyncratic Component	0.14	0.17	0.29
	(0.02)	(0.03)	(0.06)
α^H: Housing Parameter	1.20	1.09	1.61
	(0.26)	(0.24)	(0.38)
Cragg-Donald Wald F Statistic	7.91	8.10	8.21

Table B6. Parameter estimates with alternative instrumental variables. Standard errors in parentheses.

B.0.2.5 New Power Plant Development. Table B7 gives the full distribution of emissions and percent of plants that use renewables, split on whether they were constructed before or after 2000.

	Mean Emissions		Percent Renewables	
NERC	Pre-2000's	Post-2000's	Pre-2000's	Post 2000's
ASCC	935.55	842.37	37.38	15.50
FRCC	935.66	857.27	3.65	2.90
HICC	1649.43	461.88	9.22	70.62
MRO	1566.42	188.09	9.49	80.18
NPCC	410.31	747.15	24.42	14.71
RFC	1176.69	850.51	2.18	14.75
SERC	1055.78	941.07	6.16	5.25
SPP	1741.86	521.45	5.93	46.90
TRE	1135.47	620.07	1.18	29.53
WECC	858.24	597.01	40.48	36.47

Table B7. NERC regional mean carbon emissions from plants built before 2000 and after 2000. Emissions rates are measured in lbs/MWh.

B.0.2.6 Counterfactual results with model extensions.

Endogenous Electricity Pricing Table B8 displays the counterfactual results when electricity pricing is endogenous.

	Baseline	Relax Cali	Relax All
I. Percent Total Population			
California Cities	9.1	10.9	7.3
Other West	13.6	13.1	17.1
Midwest	22.2	21.8	10.0
South	37.3	36.6	25.3
Northeast	17.9	17.6	40.3
II. Mean Usage			
Gas (1000 cubic feet)	74.4	74.2	75.1
Electricity (MW h)	17.1	17.0	15.5
Fuel Oil (gallons)	60.4	59.5	133.0
III. Mean Emissions (lbs of CO2)			
Gas	8711	8688	8792
Electricity	16331	16267	14030
Fuel Oil	1622	1599	3572
Total	26664	26553	26394
(%)	100.00	99.6	99.0
IV. Average Log Income			
Skilled	100.0	100.5	112.3
Unskilled	100.0	100.0	100.1
All	100.0	100.2	104.4

Table B8. Counterfactual results with endogenous electricity pricing. Each panel shows the simulated total energy usage, total emissions, average log income, and fraction of total population living in various geographic areas in each specification. See text for details on each simulation.

Effects of Local Pollutants on Utility Table B9 presents counterfactual results in the case where $PM_{2.5}$ enters directly into the utility function. As noted in the text, the results are very similar to the baseline specification, given that changes in household electricity are the only component of the model that changes $PM_{2.5}$—and electricity contributes little to overall $PM_{2.5}$.

	Baseline	Relax Cali	Relax All
I. Percent Total Population			
California Cities	9.1	11.0	7.2
Other West	13.6	13.1	17.8
Midwest	22.2	21.7	9.3
South	37.3	36.6	23.1
Northeast	17.9	17.6	42.6
II. Mean Usage			
Gas (1000 cubic feet)	74.4	74.2	74.9
Electricity (MW h)	17.1	17.0	15.8
Fuel Oil (gallons)	60.4	59.5	138.6
III. Mean Emissions (lbs of CO2)			
Gas	8711	8686	8771
Electricity	16331	16211	13246
Fuel Oil	1622	1598	3722
Total	26664	26495	25738
(%)	100.0	99.4	96.5
IV. Average Log Income			
Skilled	100.0	100.5	113.0
Unskilled	100.0	100.0	100.4
All	100.0	100.2	104.8

Table B9. Counterfactual results with pollution in the utility function. Each panel shows the simulated total energy usage, total emissions, average log income, and fraction of total population living in various geographic areas in each specification. See Section 3.7.3 for details.

B.0.2.7 Birth-State Premium Parameters. Tables B10 through B13 display parameters governing the birth-state premium for each of the years we use in estimation. In all years, households receive a large utility premium for choosing a location in their home state and the amenity value of a location is estimated to be decreasing and convex in distance from the household head's birth state.

B.0.2.8 Demographic Group City Ranks. Table B14 provides selected estimated of ξ_{jdt}, the shared unobservable component of amenities, for the year 2017 for households with heads over the age of 35.

	Unskilled					
	Nonwhite					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	3.14	2.77	2.77	2.94	2.52	2.85
	(0.08)	(0.6)	(0.14)	(0.07)	(0.27)	(0.1)
Distance	-1.78	-1.22	-1.71	-1.78	-1.75	-1.71
	(0.07)	(0.31)	(0.09)	(0.07)	(0.21)	(0.08)
Distance Squared	0.3	0.21	0.29	0.2	0.21	0.19
	(0.01)	(0.03)	(0.01)	(0.01)	(0.03)	(0.01)
	White					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	3.15	3.03	2.94	3.15	3.08	3.13
	(0.02)	(0.06)	(0.02)	(0.01)	(0.02)	(0.01)
Distance	-1.03	-1.41	-2.05	-1.06	-1.03	-1.38
	(0.02)	(0.06)	(0.02)	(0.01)	(0.02)	(0.01)
Distance Squared	0.21	0.31	0.54	0.17	0.09	0.25
	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
	Skilled					
	Nonwhite					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	2.3	2.17	2.43	2.38	2.16	2.37
	(0.47)	(1.99)	(1.1)	(0.4)	(1.21)	(0.46)
Distance	-1.19	-1.04	-1.11	-1.48	-1.09	-1.12
	(0.27)	(0.93)	(0.63)	(0.24)	(0.56)	(0.25)
Distance Squared	0.16	0.12	0.14	0.2	0.11	0.09
	(0.03)	(0.08)	(0.06)	(0.02)	(0.05)	(0.02)
	White					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	2.02	2.04	2.1	2.13	1.92	2.11
	(0.05)	(0.1)	(0.06)	(0.04)	(0.05)	(0.02)
Distance	-2.17	-2.16	-2.35	-1.96	-2.05	-2.11
	(0.04)	(0.08)	(0.06)	(0.03)	(0.04)	(0.02)
Distance Squared	0.62	0.6	0.64	0.52	0.5	0.54
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.00)

Table B10. Parameter Estimates for 1990 Data. Standard errors multiplied by 1000 in parentheses.

	Unskilled					
	Nonwhite					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	3.11	2.62	2.69	2.89	2.66	2.73
	(0.06)	(0.52)	(0.13)	(0.04)	(0.14)	(0.07)
Distance	-1.59	-1.3	-1.55	-1.75	-1.34	-1.71
	(0.05)	(0.29)	(0.09)	(0.03)	(0.11)	(0.05)
Distance Squared	0.27	0.23	0.26	0.27	0.16	0.24
	(0.01)	(0.03)	(0.01)	(0.00)	(0.01)	(0.01)
	White					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	3	3.03	3.16	2.8	2.67	2.9
	(0.02)	(0.08)	(0.03)	(0.01)	(0.01)	(0.01)
Distance	-1.34	-1.37	-1.05	-1.84	-2.07	-1.92
	(0.02)	(0.08)	(0.03)	(0.01)	(0.02)	(0.01)
Distance Squared	0.29	0.32	0.19	0.45	0.49	0.47
	(0.00)	(0.02)	(0.01)	(0.00)	(0.00)	(0.00)
	Skilled					
	Nonwhite					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	2.22	2.02	2.33	2.37	2.15	2.33
	(0.25)	(1.27)	(0.81)	(0.18)	(0.54)	(0.28)
Distance	-1.13	-1.12	-1.19	-1.37	-1.06	-1.02
	(0.14)	(0.65)	(0.45)	(0.11)	(0.27)	(0.15)
Distance Squared	0.17	0.15	0.18	0.2	0.12	0.09
	(0.01)	(0.06)	(0.04)	(0.01)	(0.02)	(0.01)
	White					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	2.08	2.15	2.2	2.21	2.01	2.13
	(0.04)	(0.1)	(0.07)	(0.03)	(0.03)	(0.02)
Distance	-2.04	-2.01	-2.3	-1.73	-1.96	-2
	(0.03)	(0.08)	(0.07)	(0.02)	(0.03)	(0.02)
Distance Squared	0.59	0.55	0.63	0.46	0.51	0.53
	(0.01)	(0.02)	(0.01)	(0.00)	(0.01)	(0.00)

Table B11. Parameter Estimates for 2000 Data. Standard errors multiplied by 1000 in parentheses.

	Unskilled					
	Nonwhite					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	3.09	2.57	2.57	2.87	2.76	2.85
	(0.07)	(0.75)	(0.2)	(0.03)	(0.11)	(0.09)
Distance	-1.61	-1.19	-1.49	-1.65	-1.3	-1.27
	(0.06)	(0.41)	(0.13)	(0.03)	(0.09)	(0.06)
Distance Squared	0.28	0.22	0.25	0.25	0.16	0.15
	(0.01)	(0.04)	(0.01)	(0.00)	(0.01)	(0.01)
	White					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	2.83	2.85	2.72	2.77	2.63	2.74
	(0.02)	(0.1)	(0.04)	(0.01)	(0.01)	(0.01)
Distance	-1.79	-1.36	-2.12	-1.8	-2.17	-2.02
	(0.02)	(0.09)	(0.04)	(0.01)	(0.01)	(0.01)
Distance Squared	0.41	0.31	0.5	0.45	0.57	0.49
	(0.01)	(0.02)	(0.01)	(0.00)	(0.00)	(0.00)
	Skilled					
	Nonwhite					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	2.31	1.99	2.16	2.44	2.18	2.21
	(0.19)	(0.95)	(0.58)	(0.12)	(0.31)	(0.2)
Distance	-1.03	-1.06	-1.51	-1.25	-1.13	-1.1
	(0.11)	(0.48)	(0.35)	(0.07)	(0.17)	(0.1)
Distance Squared	0.15	0.15	0.24	0.18	0.13	0.11
	(0.01)	(0.05)	(0.03)	(0.01)	(0.02)	(0.01)
	White					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	2.15	2.19	2.37	2.2	2.02	2.09
	(0.04)	(0.08)	(0.06)	(0.02)	(0.02)	(0.01)
Distance	-2.04	-2.03	-2.02	-1.8	-1.89	-2.16
	(0.03)	(0.07)	(0.06)	(0.01)	(0.02)	(0.01)
Distance Squared	0.57	0.55	0.51	0.48	0.49	0.6
	(0.01)	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)

Table B12. Parameter Estimates for 2010 Data. Standard errors multiplied by 1000 in parentheses.

	Less than College					
	Nonwhite					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	3.14	2.34	2.68	2.98	2.75	2.94
	(0.07)	(0.68)	(0.25)	(0.03)	(0.11)	(0.09)
Distance	-1.58	-1.22	-1.31	-1.62	-1.54	-1.17
	(0.06)	(0.33)	(0.16)	(0.03)	(0.09)	(0.06)
Distance Squared	0.27	0.22	0.21	0.24	0.22	0.12
	(0.01)	(0.03)	(0.02)	(0.00)	(0.01)	(0.01)
	White					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	2.89	2.84	2.84	2.7	2.66	2.82
	(0.02)	(0.11)	(0.04)	(0.01)	(0.01)	(0.01)
Distance	-1.8	-1.42	-1.6	-2.04	-2.16	-1.87
	(0.03)	(0.11)	(0.04)	(0.01)	(0.01)	(0.01)
Distance Squared	0.43	0.34	0.34	0.52	0.56	0.43
	(0.01)	(0.04)	(0.01)	(0.00)	(0.00)	(0.00)
	College or More					
	Nonwhite					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	2.35	2.01	2.38	2.53	2.04	2.26
	(0.15)	(0.74)	(0.54)	(0.09)	(0.24)	(0.16)
Distance	-0.91	-0.91	-1.09	-1.25	-1.38	-1.05
	(0.08)	(0.36)	(0.32)	(0.06)	(0.13)	(0.08)
Distance Squared	0.14	0.13	0.15	0.18	0.19	0.11
	(0.01)	(0.04)	(0.03)	(0.01)	(0.01)	(0.01)
	White					
	Young			Old		
	Single	Married w/o Children	Married w/ Children	Single	Married w/o Children	Married w/ Children
Birthstate Premium	2.21	2.23	2.37	2.3	2.02	2.13
	(0.03)	(0.07)	(0.05)	(0.02)	(0.02)	(0.01)
Distance	-1.99	-1.92	-2.3	-1.8	-1.94	-2.2
	(0.02)	(0.06)	(0.06)	(0.01)	(0.02)	(0.01)
Distance Squared	0.58	0.52	0.62	0.5	0.51	0.61
	(0.01)	(0.02)	(0.02)	(0.00)	(0.00)	(0.00)

Table B13. Parameter Estimates for 2017 data. Standard errors multiplied by 1000 in parentheses.

Panel (a): White	College or more		Less than College	
Rank	Single (no kids)	Married (with kids)	Single (no kids)	Married (with kids)
1	Miami, FL	Portland, OR	San Diego, CA	Seattle, WA
2	Portland, OR	Miami, FL	Miami, FL	Portland, OR
3	Los Angeles, CA	Seattle, WA	Portland, OR	Los Angeles, CA
4	San Diego, CA	Los Angeles, CA	Seattle, WA	Honolulu, HI
5	Orlando, FL	San Diego, CA	Oxnard, CA	San Diego, CA
66	Youngstown, OH	Memphis, TN	Springfield, MA	Memphis, TN
67	Bridgeport, CT	Worcester, MA	Worcester, MA	Springfield, MA
68	Memphis, TN	Springfield, MA	Albany, NY	Worcester, MA
69	Worcester, MA	Syracuse, NY	Rochester, NY	Albany, NY
70	Syracuse, NY	Youngstown, OH	Syracuse, NY	Syracuse, NY
Panel (b): Non-white	College or more		Less than College	
Rank	Single (no kids)	Married (with kids)	Single (no kids)	Married (with kids)
1	Los Angeles, CA	Los Angeles, CA	Los Angeles, CA	Los Angeles, CA
2	San Francisco, CA	Honolulu, HI	Miami, FL	Honolulu, HI
3	Miami, FL	Miami, FL	San Francisco, CA	Seattle, WA
4	Honolulu, HI	San Francisco, CA	San Diego, CA	San Francisco, CA
5	San Diego, CA	San Diego, CA	Seattle, WA	Portland, OR
66	Albany, NY	Knoxville, TN	Springfield, MA	Springfield, MA
67	Memphis, TN	Syracuse, NY	Syracuse, NY	Albany, NY
68	Syracuse, NY	Springfield, MA	Albany, NY	Syracuse, NY
69	Rochester, NY	Scranton, PA	Milwaukee, WI	Rochester, NY
70	Milwaukee, WI	Youngstown, OH	Rochester, NY	Milwaukee, WI

Table B14. Demographic group city ranks according to the shared component of amenities. Ranks are by unobservable component of amenities for households with older household heads.

B.0.2.9 Methane Emissions. As an alternative to carbon-dioxide emissions, we also explore the relationship between the stringency of land use regulation and methane emissions. Methane is a global issue; while it is odorless and thus not considered a local pollutant, it is classified as a greenhouse gas. According to the Bernstein et al. (2008), pound for pound, methane has 25 times the global warming potential over a 100 year period compared to carbon dioxide.

The relationship between the Wharton Index and methane emissions is quite similar to that of carbon dioxide emissions. Cities with more stringent land use restrictions tend to have lower methane emissions.

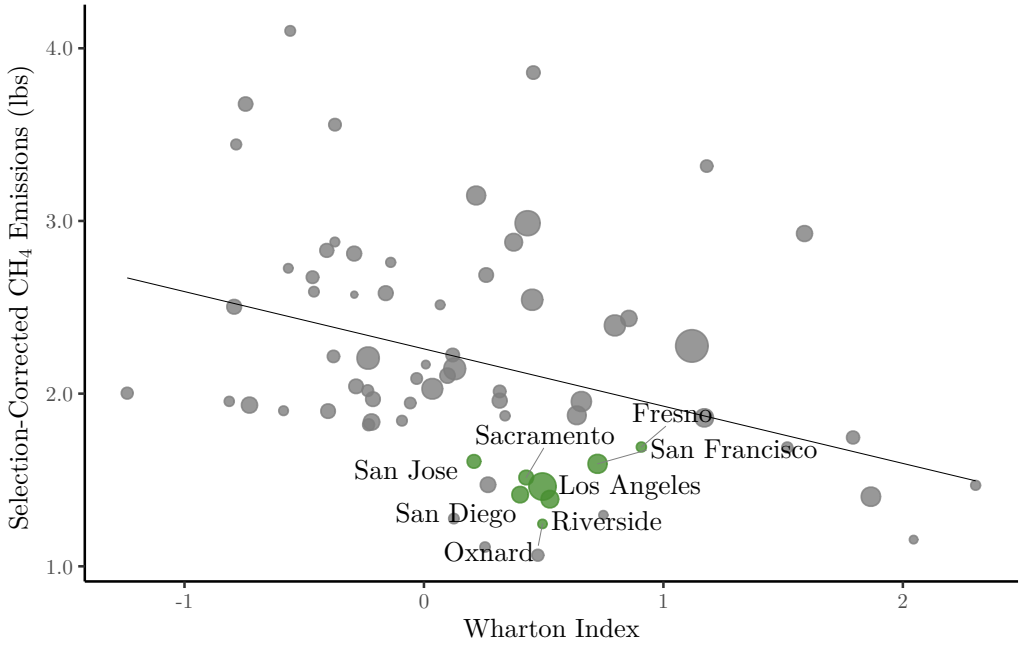


Figure B7. Methane emissions regressed on Wharton Index. Each observation is a CBSA. Size of each observation reflects population of CBSA. CBSAs in green are CA cities.

Methane emissions come from two sources: natural gas and electricity generation. Combustion of natural gas does not produce methane; however, natural gas is composed of 70% methane. Furthermore, natural gas leakages are estimated to be 1.4% according to the EPA. To impute the amount of methane emitted from natural gas ise, we use a conversion factor of 0.7*0.014 = 0.0098. As with carbon dioxide, methane emissions from electricity vary by NERC region. We compute the weighted emissions rate for methane in the same manner as we did for carbon dioxide emissions. Table B15 provides an array of city-level energy consumption, ranked on methane emissions.

CBSA	Rank	Emissions (1000 lbs)	Gas Emissions (1000 lbs)	Electricity Use (MwH)	Electricity Conversion (1000 lbs per MwH)	Electricity Emissions (1000 lbs)
Lowest						
Hartford, CT	1	1.06	0.29	13.48	0.06	0.77
New Haven, CT	2	1.11	0.29	14.32	0.06	0.82
Worcester, MA	3	1.15	0.39	13.33	0.06	0.76
Oxnard, CA	4	1.25	0.56	10.26	0.07	0.69
Bridgeport, CT	5	1.28	0.42	15.06	0.06	0.86
Springfield, MA	6	1.30	0.56	12.83	0.06	0.73
Middle						
New Orleans, LA	33	2.00	0.41	21.38	0.07	1.59
Jacksonville, FL	34	2.01	0.06	25.92	0.08	1.96
Birmingham, AL	35	2.02	0.52	20.15	0.07	1.50
Atlanta, GA	36	2.03	0.36	22.45	0.07	1.67
Austin, TX	37	2.04	0.34	22.00	0.08	1.70
Salt Lake City, UT	38	2.09	1.26	12.36	0.07	0.83
Highest						
Memphis, TN	65	3.32	0.95	31.89	0.07	2.37
Tulsa, OK	66	3.44	1.12	21.60	0.11	2.32
Oklahoma City, OK	67	3.56	1.06	23.21	0.11	2.50
Indianapolis, IN	68	3.68	2.08	18.26	0.09	1.60
Milwaukee, WI	69	3.86	1.96	21.72	0.09	1.90
Omaha, NE	70	4.10	1.55	22.84	0.11	2.55

Table B15. Predicted CBSA-level methane emissions by fuel type for the six lowest-emissions cities, the six median cities, and the six highest-emissions cities. The third column ("Emissions") shows the sum of selection-corrected **methane** emissions from natural gas, fuel oil, and electricity for the CBSA. The next two columns show emissions from gas and fuel oil respectively, which are equal to predicted usage multiplied by the appropriate emissions factor. The last three columns show predicted electricity usage, the electricity emissions factor, and predicted electricity emissions, equal to predicted electricity usage multiplied by the emissions factor. Use is measured in 1000 pounds per megawatt hour.

Our main counterfactual was to relax land use restrictions in California cities to the national median. To do this, we simulated how demand for energy services changed as a result of the changes in rental prices from the relaxation of land-use restrictions. To estimate average CBSA-level emissions, we multiplied the respective usages by the local emissions factors for each source of carbon dioxide. We can use the same simulation to examine the changes in methane emissions by using the relevant conversion factors for methane emissions. Table B16 demonstrates how methane emissions change as a result of our simulation.

	Baseline	Relax Cali	Relax All
II. Emissions (lbs of Methane)			
Gas	0.78	0.78	0.79
Electricity	1.33	1.32	1.16
Fuel Oil	0.00	0.00	0.00
Total	2.11	2.10	1.95

Table B16. Counterfactual results for methane emissions. Each column shows the amount of methane emitted from each energy source under various counterfactual scenarios.

Similar to the summary statistics for carbon emissions, we include a table without the selection correction—and additional scatterplots using OLS with demographic controls and the raw data.

CBSA	Rank	Emissions (1000 lbs)	Gas Emissions (1000 lbs)	Fuel Emissions (1000 lbs)	Electricity Use (MwH)	Electricity Conversion (1000 lbs/MwH)	Electricity Emissions (1000 lbs)
Lowest							
Honolulu, HI	1	12.83	0.47	0.07	8.08	1.52	12.29
Oxnard, CA	2	12.85	5.80	0.17	8.61	0.80	6.89
Riverside, CA	3	13.64	5.59	0.17	9.85	0.80	7.88
Los Angeles, CA	4	14.41	6.06	0.09	10.32	0.80	8.26
San Diego, CA	5	14.87	6.42	0.23	10.27	0.80	8.22
Sacramento, CA	6	15.84	7.28	0.40	10.20	0.80	8.16
Middle							
Atlanta, GA	33	25.24	6.46	0.17	17.97	1.04	18.61
Pittsburgh, PA	34	25.77	11.43	1.35	11.74	1.11	12.98
Akron, OH	35	25.85	12.05	0.58	11.95	1.11	13.21
Birmingham, AL	36	26.10	5.42	0.17	19.81	1.04	20.51
Virginia Beach, VA	37	26.19	6.12	0.71	18.70	1.04	19.36
Houston, TX	38	26.37	4.62	0.08	21.35	1.01	21.67
Highest							
Oklahoma City, OK	65	32.29	8.26	0.20	18.76	1.27	23.84
Detroit, MI	66	32.48	18.72	0.36	12.12	1.11	13.40
Philadelphia, PA	67	33.32	11.39	3.12	17.02	1.11	18.81
Memphis, TN	68	34.45	8.37	0.19	25.01	1.04	25.89
Milwaukee, WI	69	35.22	16.71	0.52	16.28	1.11	17.99
Omaha, NE	70	35.98	15.79	0.28	16.31	1.22	19.91

Table B17. Predicted CBSA-level methane (CH_4) emissions by fuel type for the six lowest-emissions cities, the six median cities, and the six highest-emissions cities in 2017. The third column ("Emissions") shows the **unconditional mean** methane emissions from natural gas and electricity for the CBSA. The next two columns show emissions from gas and fuel oil respectively, which are equal to predicted usage multiplied by the appropriate emissions factor. The last three columns show predicted electricity usage, the electricity emissions factor, and predicted electricity emissions, equal to predicted electricity usage multiplied by the emissions factor.

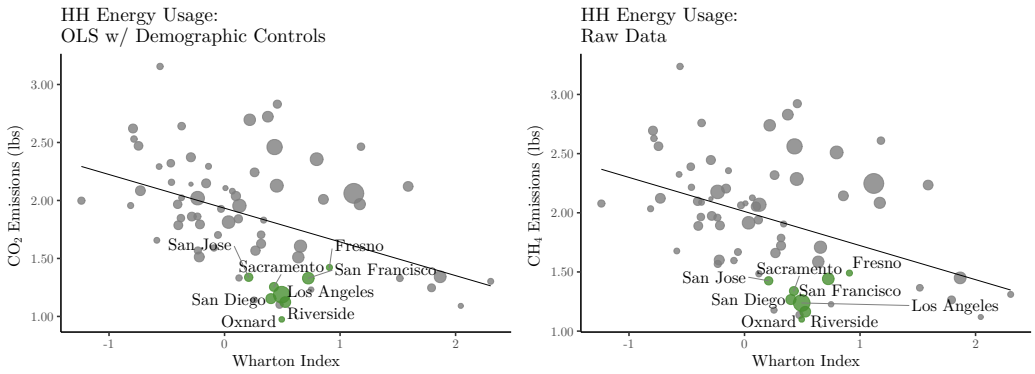


Figure B8. Methane emissions regressed on the Wharton Index. Each observation is a CBSA. The size of each observation reflects the population of CBSA. CBSAs in green are in California.

APPENDIX C

CHAPTER 4: APPENDIX

C.0.1 Model Appendix.

C.0.1.1 Firm FOC Derivation. This section derives the first-order conditions for the firms. Recall that firms in city j and sector n produce according to

$$Y_{jn} = A_{jn} K_{jn}^{\eta} \mathcal{I}_{jn}^{1-\eta},$$

where $\mathcal{I}_{jn}$ is the CES aggregator between energy and labor inputs. Specifically,

$$\mathcal{I}_{jn} = \left(\alpha_{jn} \mathcal{E}_{jn}^{\rho_{el}^n} + (1-\alpha_{jn}) \mathcal{L}_{jn}^{\rho_{el}^n} \right)^{\frac{1}{\rho_{el}^n}},$$

where

$$\mathcal{E}_{jn} = \left(\zeta_{jn} E_{jn}^{\rho_e^n} + (1-\zeta_{jn}) G_{jn}^{\rho_e^n} \right)^{\frac{1}{\rho_e^n}}$$

$$\mathcal{L}_{jn} = \left(\theta_{jn} C_{jn}^{\rho_l} + (1-\theta_{jn}) L_{jn}^{\rho_l} \right)^{\frac{1}{\rho_l}}.$$

Note that I assume factor markets are perfectly competitive so input prices are equal to their marginal revenue products. The firm's profit function is given by:

$$\pi_{jn} = P_n A_{jn} K_{jn}^{\eta} \mathcal{I}_{jn}^{1-\eta} - W_{jn}^C C_{jn} - W_{jn}^L L_{jn} - P_{jn}^G G_{jn} - P_{jn}^E E_{jn}. \tag{C.1}$$

Differentiating equation C.1 with respect to each of the firm's inputs yields the following expressions for energy prices (denoted by P_{jn}^E and P_{jn}^G) and pre-tax wages (W_{jn}^C and W_{jn}^L):

$$P_{jn}^E = \left(\frac{Y_{jn}}{\mathcal{I}_{jn}} \right) \mathcal{I}_{jn}^{1-\rho_{el}^n} \alpha_{jn} (1-\eta) \zeta_{jn} E_{jn}^{\rho_e^n - 1}$$

$$P_{jn}^f = \left(\frac{Y_{jn}}{\mathcal{I}_{jn}} \right) \mathcal{I}_{jn}^{1-\rho_{el}^n} \alpha_{jn} (1-\eta)(1-\zeta_{jn}) G_{jn}^{\rho_e^n - 1}$$

$$W_{jn}^C = \left(\frac{Y_{jn}}{\mathcal{I}_{jn}} \right) \mathcal{I}_{jn}^{1-\rho_{el}^n} (1-\alpha_{jn})(1-\eta)(\theta_{jn}) C_{jn}^{\rho_l - 1}$$

$$W_{jn}^L = \left(\frac{Y_{jn}}{\mathcal{I}_{jn}} \right) \mathcal{I}_{jn}^{1-\rho_{el}^n} (1-\alpha_{jn})(1-\eta)(1-\theta_{jn}) L_{jn}^{\rho_l - 1}.$$

I assume that capital supply is perfectly elastic and has rental rate $\bar{r}$. The firm chooses its level of capital such that the price is equal to the marginal product. Specifically, this is given by

$$K_{jn} = \left(\frac{P_n A_{jn} \eta \mathcal{I}_{jn}^{1-\eta}}{\bar{r}} \right)^{\frac{1}{1-\eta}}. \tag{C.2}$$

Plugging equation C.2 into the FOCs and rearranging yields the desired FOCs as in equation 4.8.

C.0.1.2 Firm Parameters. In this section, I outline my estimation and calibration of the production function parameters. I break the following discussion into two sections: labor and energy.

Labor Parameters. I need city-sector wages, labor elasticity of substitution (σ_l), and factor intensities (θ_{jn}'s). Let $e \in \{C, L\}$ index worker education levels. A worker of demographic d's income in location j and sector n is given by

$$I_{ejn} = W_{ejn}\ell^e \tag{C.3}$$

where ℓ^e is the number of efficiency units of labor supplied by a worker of education level e. I parameterize ℓ^e as the probability that a worker of education level e in city j and sector n is unemployed (denoted by π_{ejn}) multiplied by the efficiency units. That is

$$\ell^e = \pi_{ejn}\hat{\ell^e}, \tag{C.4}$$

where I parameterize $\hat{\ell^e}$ as

$$\hat{\ell^e} = \text{white}_i^{\beta_1^e}\text{over35}_i^{\beta_2^e}. \tag{C.5}$$

Conditional on working, the workers' pre-tax income is given by substituting equation C.4 into equation C.3 and taking logs:

$$\log(I_{ejn}) = \log(W_{ejn}) + \beta_1^e \log(\text{white}_i) + \beta_2^e \log(\text{over35}_i). \tag{C.6}$$

I thus estimate city-sector-education wages using equation C.6. Specifically, I estimate:

$$\log(I_{ejn}) = \nu_{ejn} + \beta_1^e \log(\text{white}_i) + \beta_2^e \log(\text{over35}_i) + \varepsilon_{ijn}, \tag{C.7}$$

where ν_{ejn} is a city-sector fixed effect (which estimates $\log(W_{ejn})$). I estimate equation C.7 using individual data from the ACS. To account for variation in unemployment across city-sectors, I then weight the estimated wages by the employment rate in the city-sector. This is calculated directly from the ACS data.

The remaining labor parameters to be calibrated are the labor elasticity of substitution σ_l and the labor input use intensities, θ_{jn}. Note that the log wage ratio is given by

$$\underbrace{\log\left(\frac{I^C_{jn}}{I^L_{jn}}\right)}_{\text{Estimated}} = \underbrace{-\frac{1}{\sigma_l}}_{\text{Calibrated}} \underbrace{\log\left(\frac{C_{jn}}{L_{jn}}\right)}_{\text{Data}} + \underbrace{\log\left(\frac{\theta_{jn}}{1-\theta_{jn}}\right)}_{\text{Unknown}}. \tag{C.8}$$

Note that the only unknowns in equation C.8 are the θ_{jn} values. By rearranging, I can solve for these as:

$$\theta_{jn} = \frac{B_{jn}}{1+B_{jn}} \tag{C.9}$$

where

$$B_{jn} = \left(\frac{I^C_{jn}}{I^L_{jnl}}\right)\left(\frac{C_{jn}}{L_{jn}}\right)^{\sigma_l}.$$

Energy Parameters. Similar to the above, the log ratio of energy prices is given by

$$\underbrace{\log\left(\frac{P^E_{jn}}{P^G_{jn}}\right)}_{\text{data}} = \underbrace{-\frac{1}{\sigma_e}}_{\text{calibrated}} \underbrace{\log\left(\frac{E_{jn}}{G_{jn}}\right)}_{\text{data}} + \underbrace{\log\left(\frac{\zeta_{jn}}{1-\zeta_{jn}}\right)}_{\text{unknown}}. \tag{C.10}$$

As with the labor parameters, I solve for the factor intensities using equation C.10. Specifically,

$$\zeta_{jn} = \frac{Z_{jn}}{1+Z_{jn}}, \tag{C.11}$$

where

$$Z_{jn} = \left(\frac{P^E_{jn}}{P^f_{jn}}\right)\left(\frac{E_{jn}}{G_{jn}}\right)^{\sigma_e}.$$

After recovering the ζ_n and θ_{jn}, I can recover the final set of input intensities: the α_{jn}. The ratio of the price of electricity to college educated labor is given by:

$$\underbrace{\log\left(\frac{P^E_{jn}}{W^C_{jn}}\right)}_{\text{data}} = \underbrace{\log\left(\frac{E^{\rho^n_e-1}_{jn}}{C^{\rho_l-1}_{jn}}\right)}_{\text{data}} + \underbrace{\log\left(\frac{\mathcal{E}^{\rho_e l^n-\rho^n_e}_{jn}}{\mathcal{L}^{\rho^n_{el}-\rho_l}_{jn}}\right)}_{\text{data}} + \underbrace{\log\left(\frac{\zeta_{jn}}{\theta_{jn}}\right)}_{\text{solved above}} + \underbrace{\log\left(\frac{\alpha_{jn}}{1-\alpha_{jn}}\right)}_{\text{nknown}}. \tag{C.12}$$

Thus I solve for these as:

$$\alpha_{jn} = \frac{Q_{jn}}{1 + Q_{jn}}, \tag{C.13}$$

where

$$Q_{jn} = \left(\frac{P_{jn}^{E}}{W_{jn}^{C}}\right)\left(\frac{C_{jn}^{\rho_l - 1}}{E_{jn}^{\rho_e^n - 1}}\right)\left(\frac{\theta_{jn}}{\zeta_{jn}}\right)\left(\frac{\mathcal{L}_{jn}^{\rho_{el}^n - \rho_l}}{\mathcal{E}_{jn}^{\rho_{el}^n - \rho_e^n}}\right). \tag{C.14}$$

The final set of parameters are the A_{jn} values. I pick these to match the data (i.e. I invert the first-order conditions, now that I have all variables except for the total factor productivities).

C.0.1.3 Rent Parameters. **Hedonic Rents.** I construct a rental index for each city in my sample to make comparisons in prices across CBSA'S more sensible. I regress individual gross log rent on a set of CBSA fixed effects and housing characteristics. I include the number of bedrooms, the number of rooms, household members per room, and the total number of units in the structure containing the household. Specifically, the equation I estimate to explain rental rates for household i is given by:

$$\log(R_i) = \beta_{CBSA(i)} + \beta_2 \text{Rooms}_i + \beta_3 \text{Units}_i + \beta_4 \text{Bedrooms}_i + \beta_5 \left(\frac{\text{members}_i}{\text{rooms}_i}\right) + \varepsilon_i. \tag{C.15}$$

I then take the averages of these characteristics across all CBSA's and hold them constant, and use the predicted value for each CBSA (generated from the fixed effects) as the hedonic housing rental index.

Rent Parameters. After obtaining the rental index, I can calculate the remaining parameters of the housing supply curve. The reduced-form relationship for the housing supply curve is given by

$$\log(R_j) = \frac{\beta_j}{1 + \beta_j} \log\left(\sum_e \sum_n N_{ejn} \frac{\left(\alpha_e^H \times w_{ejn}\right)}{\alpha_{ejn}}\right) + \eta_j \tag{C.16}$$

I first calculate R_j using equation C.15. I then calibrate β_j using the values in Saiz (2010).[1] For the two CBSAs in my model in which are not included among Saiz's elasticities (Honolulu & Sacramento), I compute $\beta_j = \beta_0 + \beta_1 \text{WRI}_j$ where WRI_j is

[1] Specifically, I use the estimates reported in Table VI.

the Wharton Regulation Index for city $j \in \{\text{Sac}, \text{Hono}\}$. I take parameter estimates for β_0 and β_1 from column VI, Table III of Saiz (2010). I then choose the η_j value to match the data (following the standard approach in the literature).

C.0.2 Household Energy.

C.0.2.1 Baseline Consumption. I follow Glaeser and Kahn (2010) very closely in constructing household emissions across CBSAs. I estimate household level regressions using CBSA fixed effects to impute the predicted energy use by CBSA. Specifically, I estimate:

$$x_i^m = \gamma_{\text{CBSA(i)}} + \beta_1 \log(\text{Income}_i) + \beta_2 \text{HHsize}_i + \beta_2 \text{Agehead}_i + \varepsilon_i. \tag{C.17}$$

where x_i^m is household i's consumption of fuel type $m \in \{\text{gas, elec, oil}\}$, $\gamma_{\text{CBSA(i)}}$ is a fixed effect for the household's CBSA, Income_i is the household's income obtained from equation C.8 and the other variables are the same controls used in Glaeser and Kahn (2010). I adjust the estimated coefficients by the composition of a city's single unit/multi-unit and home ownership/rental composition to address the concern that rented homes and multi-family homes are less likely to pay separately for energy. The ACS has flags for whether the household owns or rents their household, and if they live in single or multi-family housing. I reweight the estimated coefficients by the fraction of each of the four groups in every CBSA.

C.0.2.2 Energy Expenditure Shares. This section contains summary statistics for the model's energy expenditure shares: $\tilde{\alpha}^m = \frac{P^m \times x^m}{W}$. The estimates can be found in Table C2.

Expenditure Share on:	**College**	**Non-College**
Electricity		
Mean (SD)	0.025 (0.013)	0.046 (0.018)
Range	[0.005, 0.084]	[0.014, 0.133]
Natural Gas		
Mean (SD)	0.03 (0.03)	0.04 (0.05)
Range	[0.00, 0.39]	[0.00, 0.36]
Fuel-Oil		
Mean (SD)	0.001 (0.003)	0.003 (0.005)
Range	[0.000, 0.021]	[0.000, 0.025]

Table C2. Estimated energy-expenditure shares for college-educated and non-college-educated workers, respectively. Summary statistics are taken over the choice set (N=395). Estimates of energy use of type m come from equation C.17. Estimates of wages come from equation C.7. State-level electricity prices come from EIA data.

C.0.2.3 HSV Transfers. Let post-transfer income be determined by:

$$\tilde{w}_{ij} = w_{ij} + \lambda w_{ij}^{1-\gamma},$$

where $\lambda > 0$ is the overall level of the reimbursement and $\gamma > 1$ indexes the progressivity of the transfers. Note that the government's total revenue from a carbon tax of size τ is given by $\mathcal{T} = \tau \sum_n \sum_j \delta_j \hat{f}_{jn}$ where again $\hat{f}_{jn}$ is total energy use in city-sector jn. Denote the household's transfer as $g_i = w_{ij}^{1-\gamma}$. In equilibrium, the sum of government transfers, $\mathcal{G} = \sum_i g_i$ must be equal to total carbon tax revenue. I can solve for λ that satisfies the government budget constraint, $\lambda^\star$, by equating total payments to total revenues:

$$\lambda^\star \sum_e \sum_j N_{ejn}^\star w_{ejn}^{1-\gamma} = \tau \sum_n \sum_j \sum_m \delta_j^m \hat{f}_{jn}^m$$

$$\lambda^\star = \frac{\tau \sum_n \sum_j \sum_m \delta_j^m \hat{f}_{jn}^m}{\sum_e \sum_j N_{ejn}^\star w_{ejn}^{1-\gamma}}.$$

Note that the allocation of workers across cities, N_j, and energy consumption, $\hat{f}_{jn}$, are functions of λ in equilibrium. Computationally, I guess a value of λ_g and the shares, calculate the *implied* $\lambda^\star$ and check whether $\lambda^\star = \lambda_g$.

If the condition does not hold, I update my guess as a convex combination of my original guess and $\lambda^\star$.

C.0.2.4 Model Fit: Energy Consumption. Figure C1 demonstrates model's baseline fit for the three different fuel types for households: electricity, natural gas, and fuel oil.

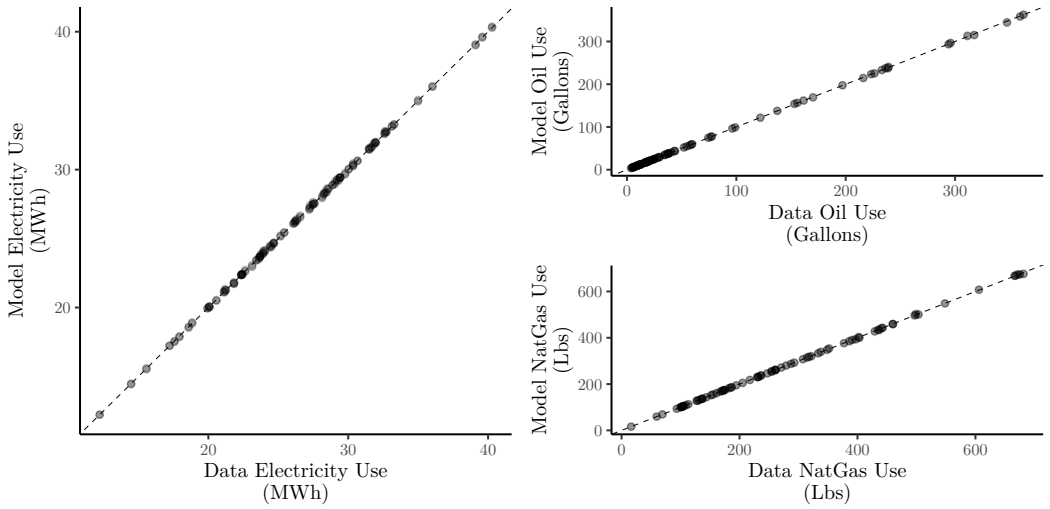


Figure C1. Baseline model fit for electricity, natural gas, and fuel-oil. The x-axis in each graph gives city-sector fuel consumption (predicted from equation C.17), and the y-axis has the baseline equilibrium energy consumption in the model.

C.0.3 Energy-Adjusted Income. In this section, I derive the "energy-adjusted budget income," $w^{EA}_{ejnt} = \dfrac{\log(w_{ejnt}) - \sum_m \left(\tilde{\alpha}^m_{ejnt} \log(P_{jt})\right)}{1 - \sum_m \tilde{\alpha}^m_{ejnt}}$, used in estimating equation 4.18, following Colas and Morehouse (2022). Recall that the mean utility is given by:

$$\delta_{ejnt} = \left(\frac{1 + \alpha^H_e + \sum_m \alpha^m_{ej}}{\sigma_e}\right) \log(w_{ejnt}) - \frac{\alpha^h_e}{\sigma_e} \log(R_{jt}) - \sum_m \frac{\alpha^m_{ej}}{\sigma_e} \log P^m_{jt} + \xi_{ejnt},$$

Note that $\tilde{\alpha}^m_{ejnt} = \frac{\alpha^m_{ejnt}}{\alpha_{ejnt}}$ implies that $\sum_{m'} \tilde{\alpha}^{m'}_{ejt} = \frac{\sum_{m'} \tilde{\alpha}^{m'}_{ejt}(1+\alpha^h_e)}{1-\sum_{m'} \alpha^{m'}_{ejt}}$ and thus $\alpha^m_{ejt} = \frac{\tilde{\alpha}^m_{ejt}(1+\alpha^h_e)}{1-\sum_{m'} \tilde{\alpha_{ejt}}^{m'}}$. I can plug these into the equation for mean utility to get:

$$\delta_{ejnt} = \left(\frac{1 + \alpha_e^h + \frac{\tilde{\alpha}_{ejt}^m (1+\alpha_e^h)}{1-\sum_{m'} \tilde{\alpha_{ejt}}^{m'}}}{\sigma_e} \right) \log(w_{ejnt}) - \frac{\alpha_e^h}{\sigma_e} \log(R_{jt}) -$$
$$\frac{(1+\alpha_e^h)}{1 - 1 - \sum_{m'} \tilde{\alpha_{ejt}}^{m'}} \sum_m \frac{\tilde{\alpha}_{ej}^m}{\sigma_e} \log P_{jt}^m + \xi_{ejnt}.$$

Rearranging yields:

$$\delta_{ejnt} = \Theta_{et}^w \log(w_{ejnt}^{EA}) + \Theta_{et}^r \log(R_j) + \epsilon_{ejn},$$

where $w_{ejnt}^{EA} = \frac{\log(w_{ejnt}) - \sum_m \left(\tilde{\alpha}_{ejnt}^m \log(P_{jt}) \right)}{1 - \sum_m \tilde{\alpha}_{ejnt}^m}$,$\Theta_e^w = \frac{1+\alpha_e^h}{\sigma_e}$, $\Theta_e^r = \frac{\alpha_e^h}{\sigma_e}$. I use the notation w^{EA} to denote "energy-adjusted" wages. Given estimates for Θ_e^w and Θ_e^r, I can then solve for α_e^h and σ_e.

Parameter Name	Notation	Source
Labor Supply		
Moving Costs	$\Theta_e^{div}, \Theta_e^{dist}, \Theta_e^{dist2}$	MLE (equation 4.16)
Marginal Utility of Income/Rents	Θ_e^w, Θ_e^r	IV (equation 4.18)
Variance of pref. shock	σ_e	Algebra
Housing Parameter	α_e^H	Algebra
Utility of Energy param.	α_{ejn}^m	Algebra
Wage Index	W_{ejn}	OLS (equation C.7)
Firm Parameters		
Energy-Lab EoS	σ_{el}^n	(Koesler & Schymura, 2012)
Gas- Elec EoS	σ_e^n	(Serletis et al., 2010)
College-No College EoS	σ_l	(Card, 2009)
Factor Intensities	$\alpha_{jn}, \theta_{jn}, \zeta_{jn}$	Relative demand curves
TFP	A_{jn}	Firm FOCs
Energy and Rent Parameters		
Intercepts	β_j, a_{mj}	Algebra
Energy Supply Elasticity	κ	(C. Dahl & Duggan, 1996)
Rent Supply Elasticity	γ_j	(Saiz, 2010)
Carbon Emissions Factor	δ^m	Data
Rent Index	R_j	OLS (equation C.15)

Table C3. Overview of calibration and estimation strategy for the model's parameters. For shorthand, I abbreviate elasticity of substitution as EoS and total factor productivity as TFP.

C.0.4 Estimation Summary.

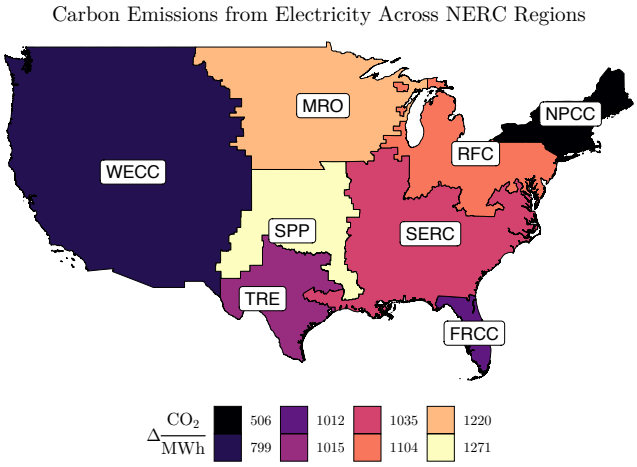


Figure C2. NERC regional electricity emissions factors. Emissions factors are calculated as output-weighted averages of individual plant's emissions rates.

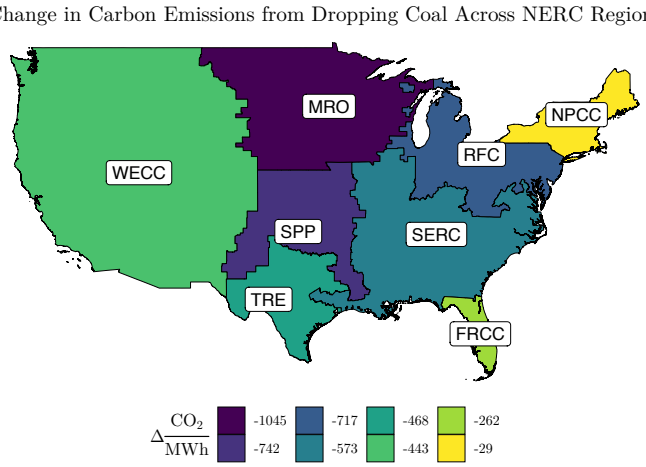


Figure C3. For each NERC region, emissions factors from electricity in the absence of coal-fired power plants. Emissions factors are calculated as output-weighted averages of individual plants' emissions rates, for all plants in a region excluding coal.

C.0.5 NERC Region Emissions Factors.

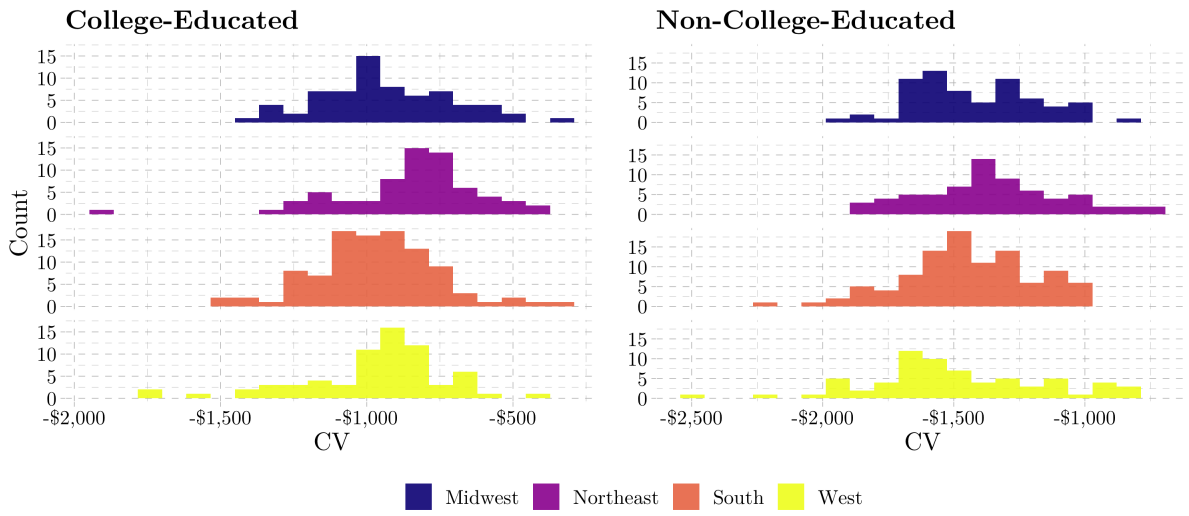


Figure C4. Monetized regional tax incidence. For each Census Region, the distribution of mean compensating variation across cities and sectors from a $31 per ton carbon tax. CV is measured in dollars.

C.0.5.1 Heterogeneity in Incidence Across Space.

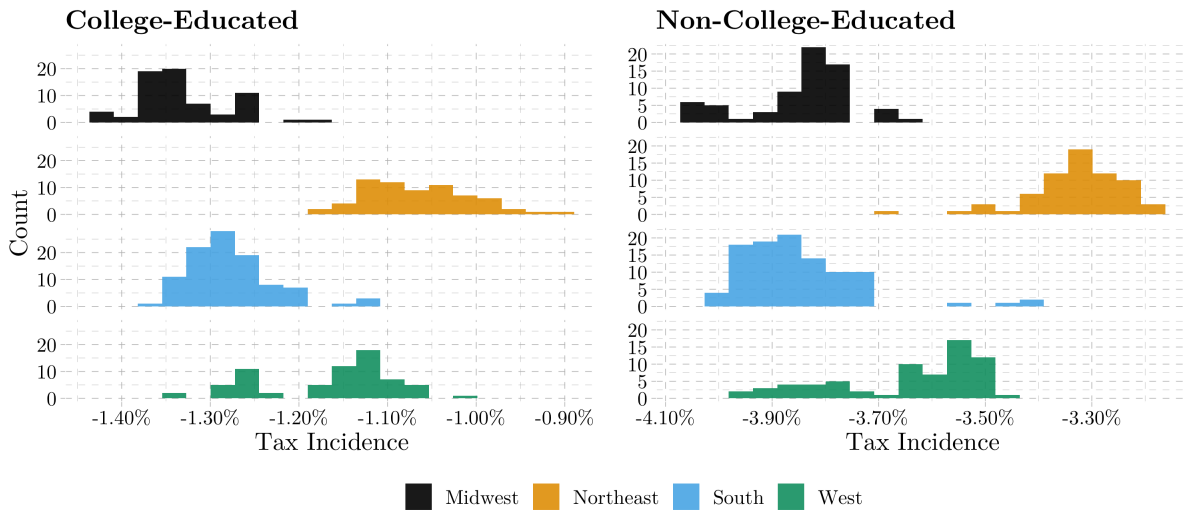


Figure C5. Non-Monetized tax incidence. For each Census Region, distribution of tax incidence (non-monetized) in utility across cities and sectors from a $31 per ton carbon tax.

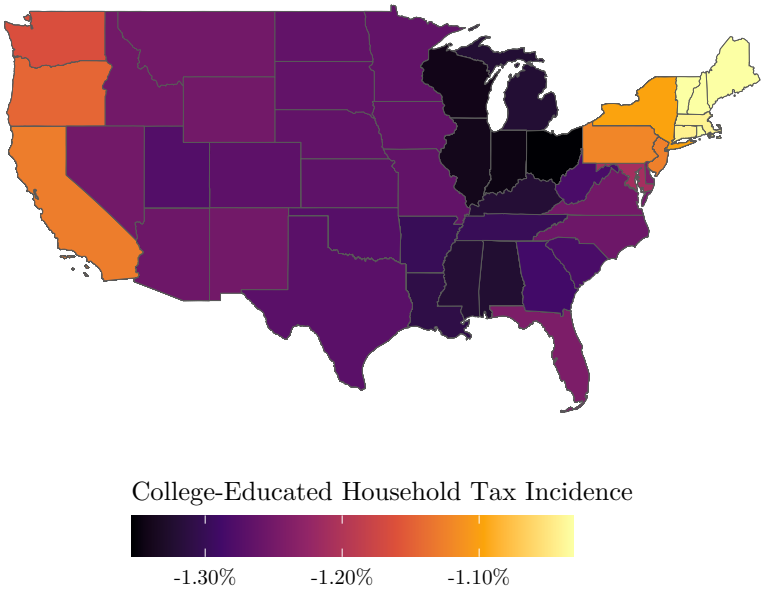


Figure C6. Non-Monetized college-educated regional tax incidence as a percent of baseline wages. Tax incidence (non-monetized) across states from a $31 per ton carbon tax for college educated households.

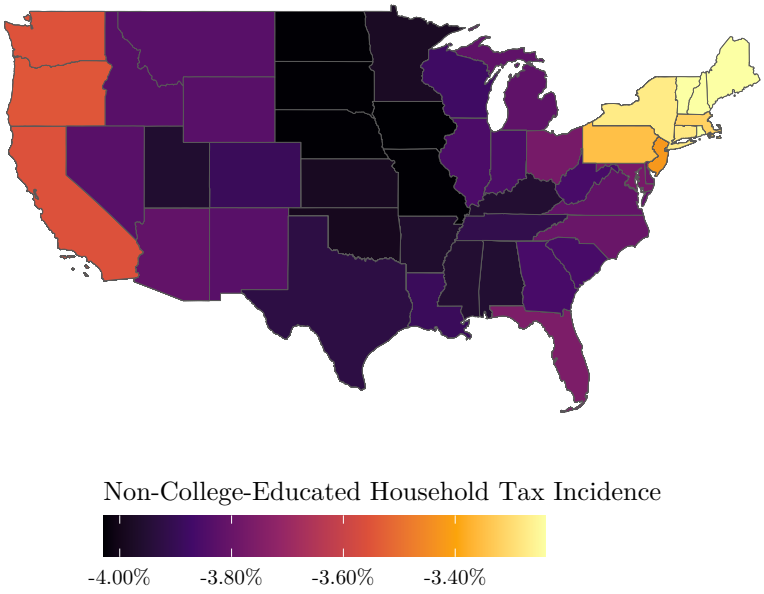


Figure C7. Non-Monetized non-college-educated regional tax incidence as a percent of baseline wages. Tax incidence (non-monetized) across states from a $31 per ton carbon tax for non-college educated households.

$\tau = \$31$/ton: **No Transfers**	Midwest	Northeast	South	West
Overall	$-1,290$	$1,195$	$1,255$	$1,291$
	(309)	(315)	(321)	(355)
College	971	967	962	178
	(181)	(218)	(178)	(257)
Non-College	$-1,471$	$1,404$	$1,455$	$1,508$
	(203)	(235)	(230)	(257)

Table C4. Monetized incidence across regions. By Census Region, compensating variation across cities and sectors from a $31 per ton carbon tax. Standard deviations are in parentheses.

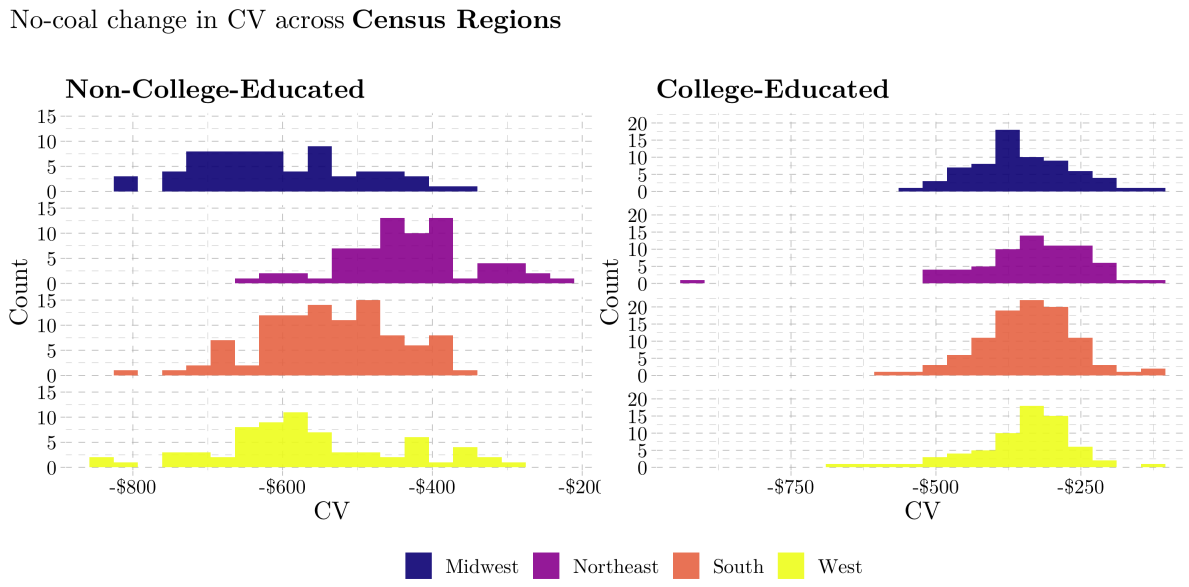


Table C5. No coal monetized tax incidence. By Census Region, the change in compensating variation from "dropping" coal from a $31 per ton carbon tax.

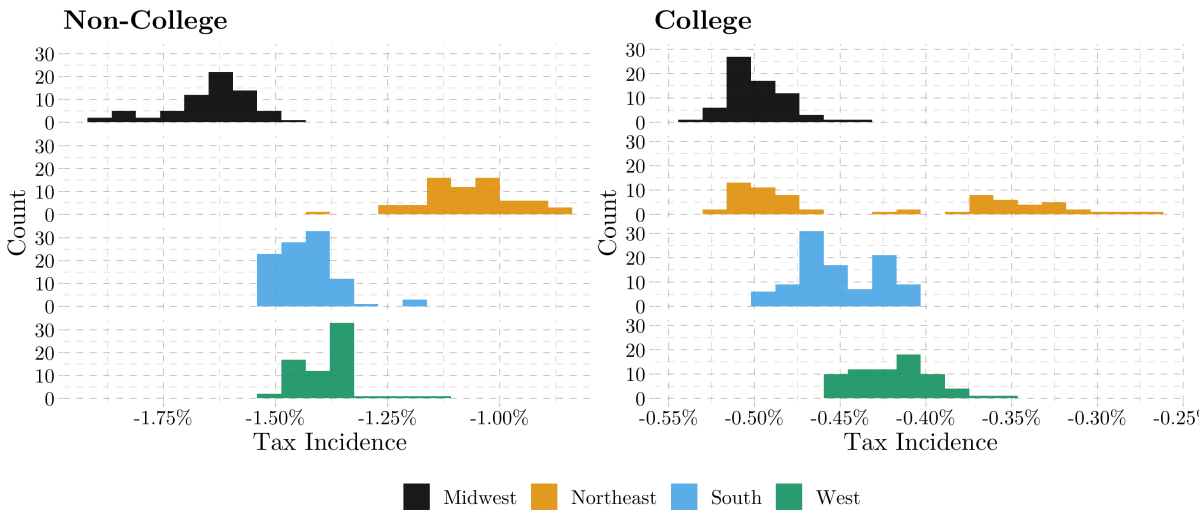


Table C6. No-coal non-monetized tax incidence. By Census Region, the percent change in tax incidence (non-monetized) from "dropping" coal from a $31 per ton carbon tax.

C.0.5.2 Heterogeneity in Incidence Across Industries. In this section, I decompose the monetized (compensating variation) and non-monetized incidence from a $31 per ton carbon tax.

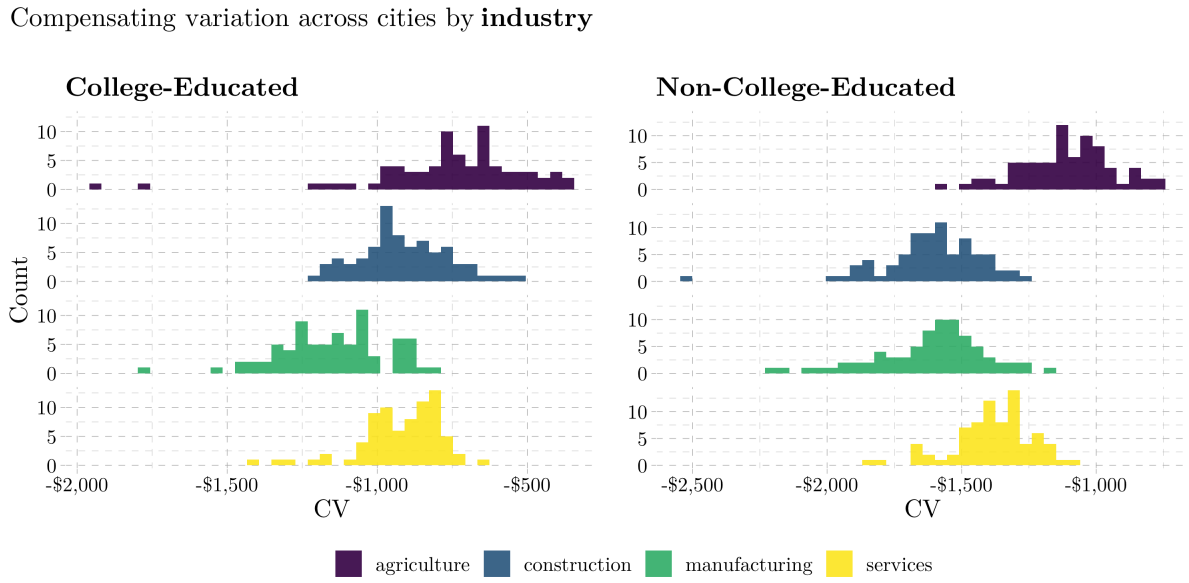


Figure C8. Monetized tax incidence from a carbon tax. By industry, the distribution of mean compensating variation across cities by industries from a $31 per ton carbon tax. CV is measured in dollars.

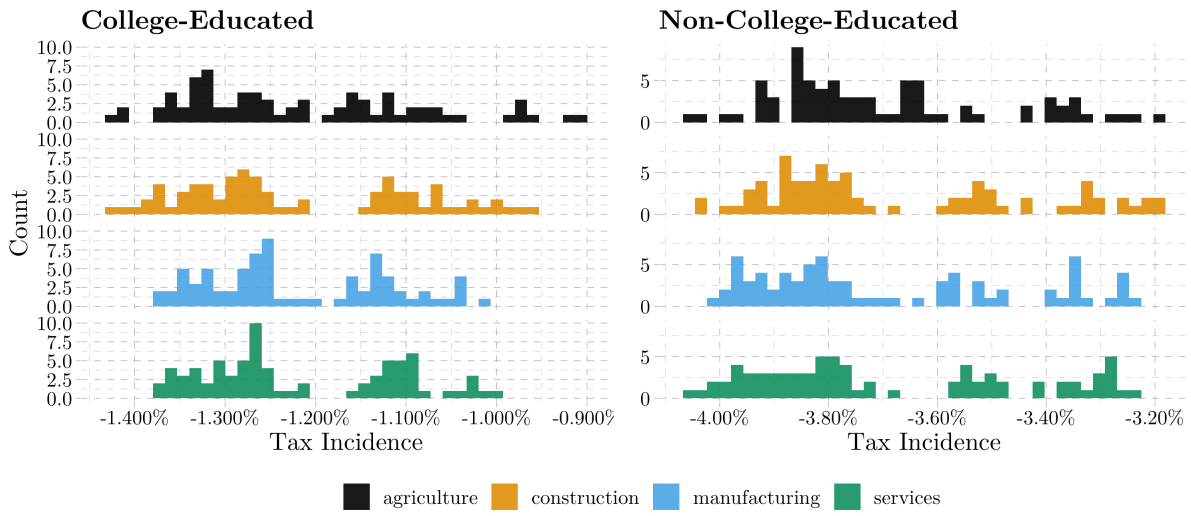


Figure C9. Non-Monetized industry tax incidence. By industry, the distribution of tax incidence (non-monetized) across cities from a $31 per ton carbon tax.

$\tau = \$31$/ton: **No Transfers**	Manufacturing	Services	Construction	Agriculture
Overall	$-1,446$	$1,156$	$1,491$	$1,012$
	(263)	(271)	(304)	(217)
College	$1,170$	949	916	708
	(224)	(187)	(151)	(171)
Non-College	$-1,565$	$1,360$	$1,598$	$1,094$
	(174)	(168)	(180)	(142)

Table C7. Industry compensating variation. By industry, mean compensating variation across cities from a $31 per ton carbon tax. Standard deviations are in parentheses.

C.0.5.3 Migration. In this section, I map out changes in population across cities from a $31 per ton carbon tax. I aggregate the changes in population level across CBSAs to the state level. My sample only includes 34 states, I use changes in population at the Census-Division level for states with missing data, excluding the population changes from the included CBSAs within the respective census division.

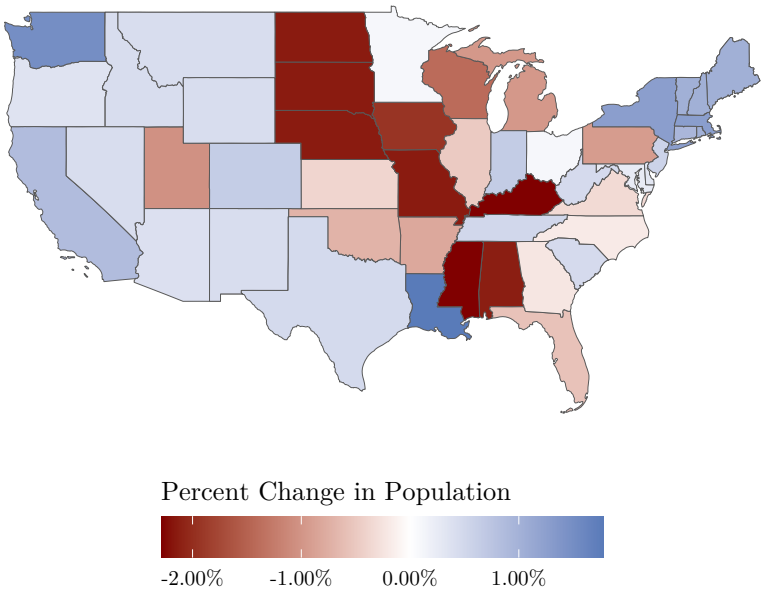


Figure C10. Migration from a carbon tax. Population changes are computed using the equilibrium arising from the model using a $31 per ton carbon tax.

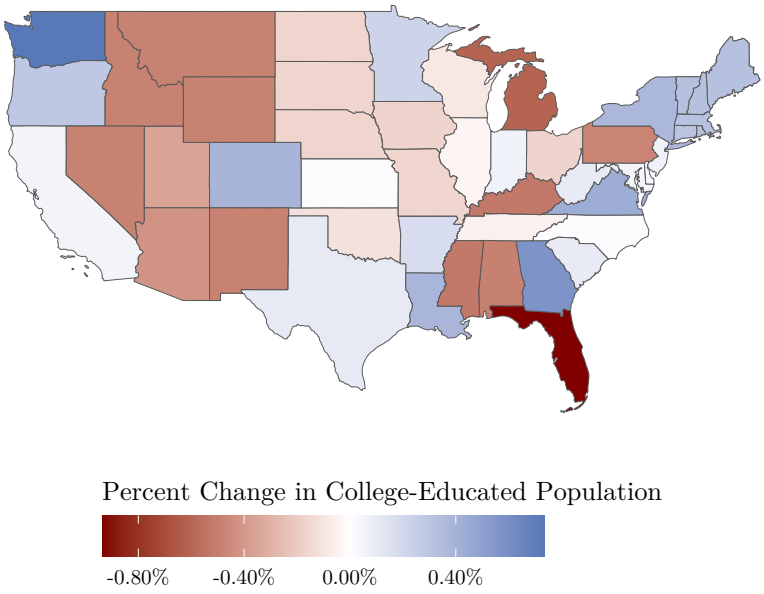


Figure C11. College-educated worker migration resulting from a carbon tax. Population changes are computed using the equilibrium arising from the model using a $31 per ton carbon tax.

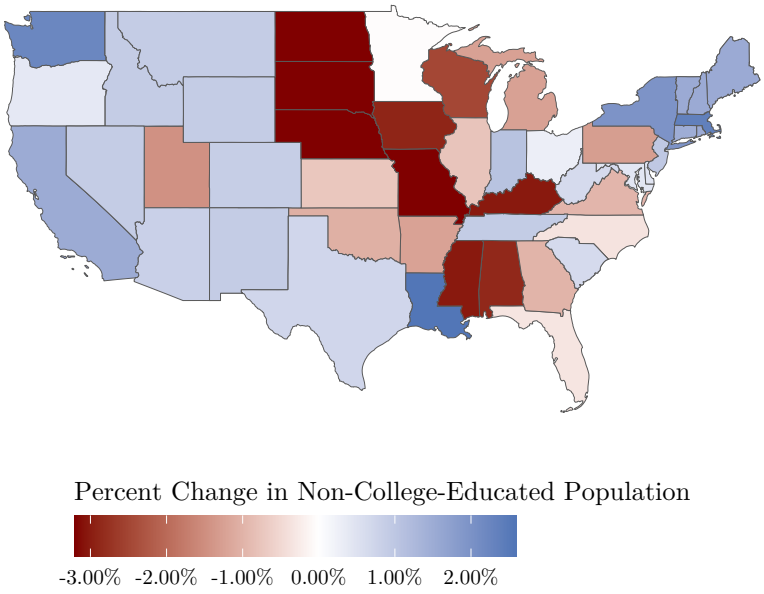


Figure C12. Non-College migration from a carbon tax. Population changes are computed using the equilibrium arising from the model using a $31 per ton carbon tax.

C.0.5.4 Voting and Tax Incidence. This section provides plots that explore the relationship between predicted carbon tax incidence and the share of each given CBSA that voted for Trump in the 2016 Presidential election.

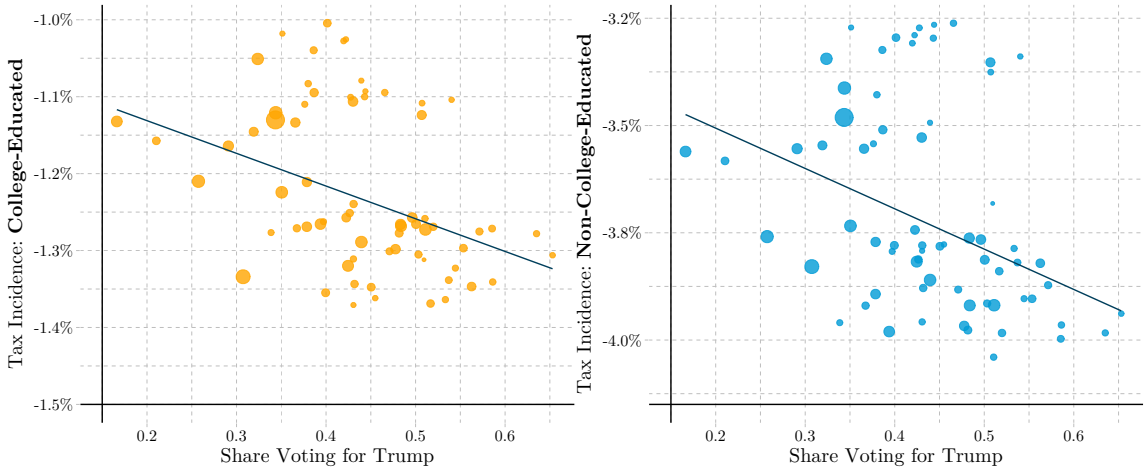


Figure C13. Tax Incidence and voting. An observation is a CBSA; the size of each observation reflects the total number of *voters* in the 2016 presidential election in the CBSA. Tax incidence is measured as the relative change in utility (measured in percent of income) from a $31 per ton carbon tax, calculated using equation 4.20.

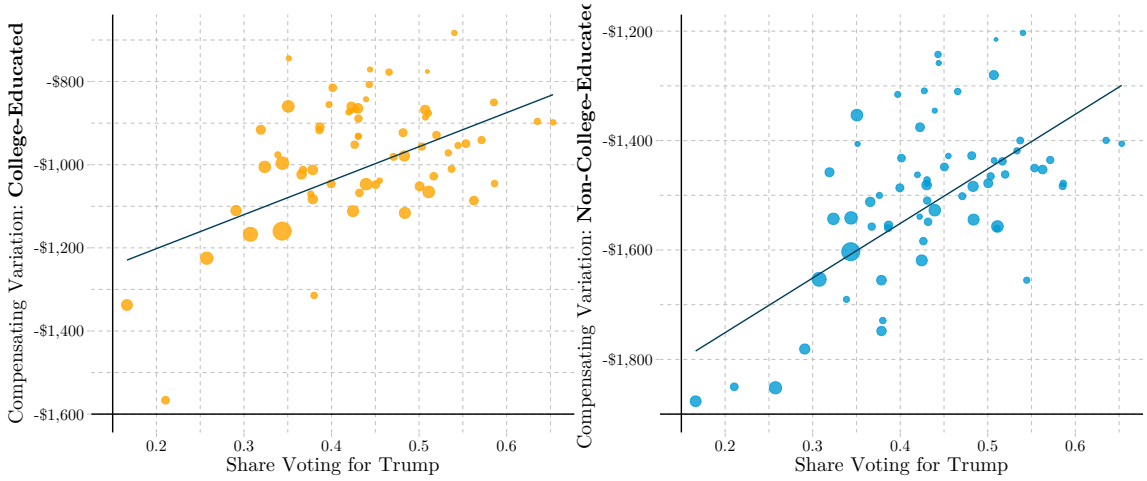


Figure C14. Compensating variation and voting. An observation is a CBSA; the size of each observation reflects the total number of *voters* in the 2016 presidential election in the CBSA. Compensating variation is measured as the average dollar amount required to make a household indifferent to a $31 per ton carbon tax, calculated using equation 4.20.

C.0.6 Computational Appendix.

C.0.6.1 Nested Fixed Point Algorithm. In this section, I outline the nested-fixed-point algorithm I use to obtain the preference parameters:

$$\boldsymbol{\Theta}^* = \sum_{i=1}^{N^d} \sum_{n \in N} \sum_{j \in J} \mathbb{I}_i(j,n) \log(P_i),$$

where P_i is defined in equation 4.15. In what follows, let t denote iteration number (**not time**). I proceed in the following manner:

Outer-Loop

(1) Start with an arbitrary guess for the parameter vector, $\boldsymbol{\Theta}^{*,0}$

Inner-Loop

(2) Guess an arbitrary level of mean utilities for each city-sector pair and education level, δ^0_{ejn}

(3) Given the guess of the parameter vector and mean utility, compute the model's predicted share of agents in each sector-city pair. This is given by

$$S^0_{\text{model}_{ejn}} = \sum_{i}^{N^d} P_i,$$

(4) Use the Nevo (Nevo, 2000) improvement of the contraction mapping from Berry (1994):

$$\exp(\delta^1_{ejn}) = \exp(\delta^0_{ejn}) \times \left(\frac{S_{\text{data}_{ejn}}}{S^0_{\text{model}_{ejn}}} \right), \tag{C.18}$$

where $S_{\text{data}_{ejn}}$ is the share of agents of education level e that choose city j and sector n in the data.

(5) Check for convergence of equation C.18. Specifically, I check whether

$$\sup \left| \delta^1_{ejn} - \delta^0_{ejn} \right| < \epsilon.$$

If the equation hasn't converged, update the new guess of δ to δ^1 and go back to step (3). Repeat 3-5 until δ has converged. This ends the inner-loop.

(6) After obtaining the unique δ_{ejn} for the guess of $\mathbf{\Theta}^{*,0}$, I can compute the value of the likelihood function. Check to see if the log-likelihood function is maximized. If not, go back to step (1). To update the guess of $\mathbf{\Theta}^{*}$, I use the Nelder-Mead algorithm (Nelder & Mead, 1965).

C.0.6.2 Equilibrium Simulation. In this section, I outline how I solve for the counterfactual equilibrium.

(1) Guess a vector of wages, rents, residential electricity prices, industrial electricity demand, and industrial natural gas demand.

(2) Given these guesses, calculated labor supply in each city using the implied choice shares multiplied by the total population of college-educated and non-college-educated workers.

(3) Calculate the value of the labor aggregator, the energy aggregator, and finally, the input aggregator using the implied populations from step 2 and the guesses of industrial energy consumption.

(4) Calculate total housing demand using equation. C.16.

(5) Check whether the new vectors of wages, rents, and residential energy prices and industrial energy consumption from steps 3 and 4 are within ϵ of the guess made in step 1. If not, return to step 1, using updated guesses that are convex combinations of the old guess and new prices that come from the firm's FOC and rent equations. To update the guess of industrial gas demand, I check whether the price from the firm's FOC matches the prices in the data. Otherwise, I update the new guess of industrial gas to be larger than the old. If so, I update the new guess to be smaller. For electricity, I use the same process but compare the firm's FOC prices to city-level supply prices (according to the supply curve) and update accordingly.

C.0.7 Data Appendix.

C.0.7.1 Labor Supply & Demand. **Cleaning.** My sample consists of all non-military, non-institutionalized, employed, (16-64) individuals. I drop all observations with missing or negative incomes. Additionally, I drop all workers not in the 5 sectors described in the model section. A household is defined by the ACS identifying variable *SERIAL*, which assigns all individuals in a household the same number. The decision-maker in the model is the "household head" (given by the IPUMS indicator variable *RELATE*).

Geography. I construct the labor quantities supplied and demanded from the 5-year aggregated ACS 2012-2016 data. My geographic unit of observation is a "Core-Based Statistical Area" (CBSA). I follow closely the strategy used in other literature, in constructing the sample, to make my analysis more comparable to other papers. Specifically, I choose the 70 largest CBSAs, as defined by the population in 1980. I then map other individuals, who do not live in one of these 70 CBSAs into their corresponding broader census division, creating an additional nine alternatives. Thus the model consists of 79 unique, geographic alternatives. All wage estimates and alternative counts are weighted by the exogenous sampling weights provided by the ACS.

Industry. To construct the model's industries, I use the "INDNAICS" variable from the ACS. While the ACS has NAICS codes at the 6-digit (least aggregated), I constrain the model to contain 5 sectors: agriculture, manufacturing, services, construction, and an outside option sector.

C.0.7.2 City-Sector Energy Use. In this section, I detail how I assign energy consumption by city-sector to firms in the baseline case. I observe energy consumption by sector at the national level from the EIA data. Additionally, the EIA has utility-level electricity and natural gas use by aggregated industrial, commercial, and residential sectors. I construct a sector n's consumption of electricity (natural gas) as proportional to the city-sectors' share of aggregate employment. Note that for sectoral employment, L_n, I use BLS data. For the numerator, L_{jn}, I use counts from my sample. Given that E_n is aggregated across all US cities, my strategy using the fraction of workers in my sample across the whole US adjusts firm energy consumption to match my subset of households.

Specifically, electricity and gas consumption bu firms is given by:

$$E_{jn} = \frac{L_{jn}}{L_n} * E_n$$
$$G_{jn} = \frac{L_{jn}}{L_n} * G_n.$$

C.0.8 Additional Scatterplots. In this section, I disaggregate Figure 20 by education group in Figure C15 and then by both education group and sector in Figure C16.

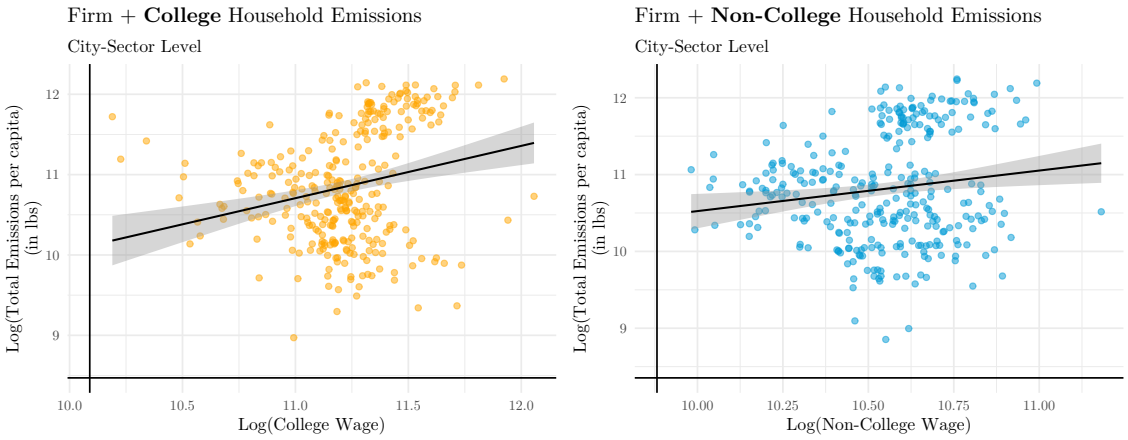


Figure C15. Total emissions per capita for each city-sector, plotted against wages. Wages are constructed as city level averages estimated from equation C.7. Emissions are the sum of firm emissions per capita and household emissions for a given education group.

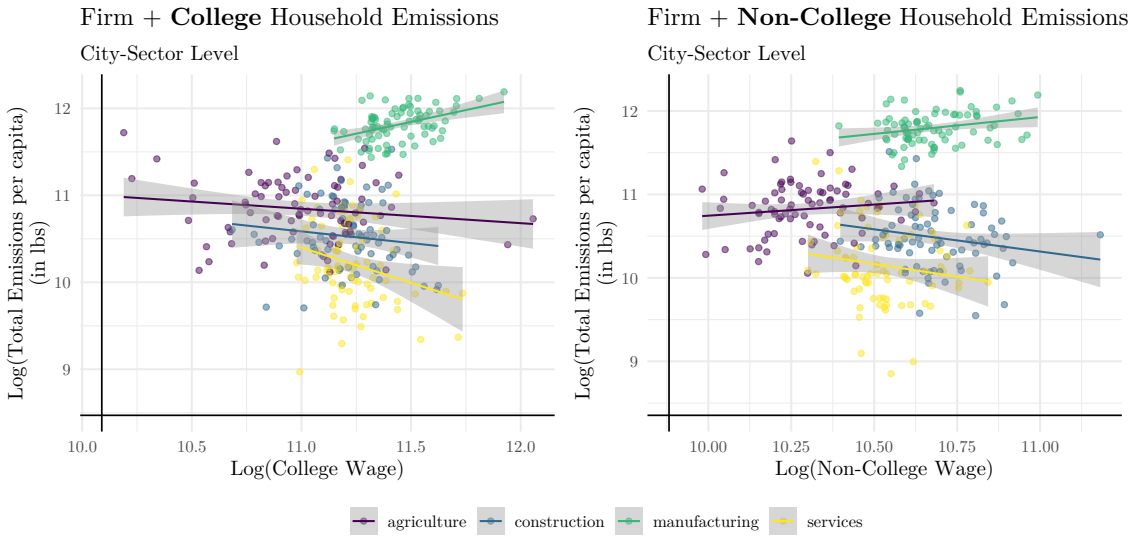


Figure C16. Total emissions per capita by city across industries plotted against wages. Wages are constructed as city-level averages estimated from equation C.7. Emissions are the sum of firm emissions per capita and household emissions for a given education group.

C.0.9 Sector Energy Demand. The EIA classifies the Industrial and Commercial sector in the following manner:

– **Industrial sector**: An energy-consuming sector that consists of all facilities and equipment used for producing, processing, or assembling goods. The industrial sector encompasses the following types of activity manufacturing (NAICS codes 31-33); agriculture, forestry, fishing and hunting (NAICS code 11); mining, including oil and gas extraction (NAICS code 21); and construction (NAICS code 23). Overall energy use in this sector is largely for process heat and cooling and powering machinery, with lesser amounts used for facility heating, air conditioning, and lighting. Fossil fuels are also used as raw material inputs to manufactured products. Note: This sector includes generators that produce electricity

and/or useful thermal output primarily to support the above-mentioned industrial activities. Various EIA programs differ in sectoral coverage.

– **Commercial Sector**: An energy-consuming sector that consists of service-providing facilities and equipment of businesses; Federal, State, and local governments; and other private and public organizations, such as religious, social, or fraternal groups. The commercial sector includes institutional living quarters. It also includes sewage treatment facilities. Common uses of energy associated with this sector include space heating, water heating, air conditioning, lighting, refrigeration, cooking, and running a wide variety of other equipment. Note: This sector includes generators that produce electricity and/or useful thermal output primarily to support the activities of the above-mentioned commercial establishments.

REFERENCES CITED

Ailshire, J. A., & Crimmins, E. M. (2014). Fine particulate matter air pollution and cognitive function among older U.S. adults. *American Journal of Epidemiology*, *180*(4), 359–366.

Albouy, D. (2012). *Are big cities bad places to live? Estimating quality of life across metropolitan areas* (Tech. Rep.). National Bureau of Economic Research.

Albouy, D., & Ehrlich, G. (2018). Housing productivity and the social cost of land-use restrictions. *Journal of Urban Economics*, *107*, 101–120.

Albouy, D., & Stuart, B. (2014). *Urban population and amenities: The neoclassical model of location* (Tech. Rep.). National Bureau of Economic Research.

Aldy, J. E., Kotchen, M., Evans, M. F., Fowlie, M., Levinson, A., & Palmer, K. (2020). *Co-benefits and regulatory impact analysis: Theory and evidence from federal air quality regulations* (Tech. Rep.). National Bureau of Economic Research.

Anderson, M. L. (2019, October). As the wind blows: The effects of long-term exposure to air pollution on mortality. *Journal of the European Economic Association*, *18*(4), 1886–1927. Retrieved from https://doi.org/10.1093/jeea/jvz051 doi: 10.1093/jeea/jvz051

Arias, M. A., Reinbold, B., & Restrepo-Echavarria, P. (2017). The decline of coal. *Economic Research-Federal Reserve Bank of St. Louis*.

Barwick, P. J., Li, S., Rao, D., & Zahur, N. B. (2018). The morbidity cost of air pollution: Evidence from consumer spending in China. *NBER Working Paper No. 24688*. Retrieved from https://doi.org/10.3386/w24688 doi: 10.3386/w24688

Baum-Snow, N., & Pavan, R. (2013). Inequality and city size. *Review of Economics and Statistics*, *95*(5), 1535-1548.

Bayer, P., Keohane, N., & Timmins, C. (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management*, *58*(1), 1–14.

Beck, M., Rivers, N., Wigle, R., & Yonezawa, H. (2015). Carbon tax and revenue recycling: Impacts on households in British Columbia. *Resource and Energy Economics*, *41*, 40–69.

Becker, R., & Henderson, V. (2000). Effects of air quality regulations on polluting industries. *Journal of Political Economy*, *108*(2), 379–421. Retrieved from https://doi.org/10.1086/262123 doi: 10.1086/262123

Bernstein, L., Bosch, P., Canziani, O., Chen, Z., Christ, R., Davidson, O., Hare, W., Huq, S., Karoly, D., Kattsov, V., et al. (2008). *Climate change 2007: Synthesis report: An assessment of the intergovernmental panel on climate change* (Tech. Rep.). IPCC.

Berry, S. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242–262.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, *63*(4), 841-890.

Berry, S., Levinsohn, J., & Pakes, A. (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of Political Economy*, *112*(1), 68–105.

Blomquist, G. C., Berger, M. C., & Hoehn, J. P. (1988). New estimates of quality of life in urban areas. *The American Economic Review*, 89–107.

Borck, R. (2016). Will skyscrapers save the planet? Building height limits and urban greenhouse gas emissions. *Regional Science and Urban Economics*, *58*, 13–25.

Bound, J., & Holzer, H. J. (2000). Demand shifts, population adjustments, and labor market outcomes during the 1980s. *Journal of Labor Economics*, *18*(1), 20-54.

Brockway, P. E., Heun, M. K., Santos, J., & Barrett, J. R. (2017). Energy-extended CES aggregate production: Current aspects of their specification and econometric estimation. *Energies*, *10*(2), 202.

Browning, D. (2020, 04). Don't celebrate Earth Day. Fight for it. *The New York Times*. Retrieved 2020-07-14, from https://www.nytimes.com/2020/04/21/opinion/trump-epa-earth-day.html

CAMD. (2020). *US EPA, Clean Air Markets Division.* Retrieved from https://www.epa.gov/airmarkets

Carattini, S., Carvalho, M., & Fankhauser, S. (2017). How to make carbon taxes more acceptable. *London: Grantham Research Institute on Climate Change and the Environment, and Centre for Climate Change Economics and Policy, London School of Economics and Political Science.*

Card, D. (2009). Immigration and inequality. *American Economic Review*, *99*(2), 1–21.

Castellanos, K. A., & Heutel, G. (2019). *Unemployment, labor mobility, and climate policy* (Tech. Rep.). National Bureau of Economic Research.

Chamberlain, G. (1986). Asymptotic efficiency in semi-parametric models with censoring. *Journal of Econometrics*, *32*(2), 189–218.

Chay, K. Y., & Greenstone, M. (2003). The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. *The Quarterly Journal of Economics*, *118*(3), 1121–1167.

Chay, K. Y., & Greenstone, M. (2005). Does air quality matter? evidence from the housing market. *Journal of Political Economy*, *113*(2), 376–424.

Cherniwchan, J., Copeland, B. R., & Taylor, M. S. (2017, August). Trade and the environment: New methods, measurements, and results. *Annual Review of Economics*, *9*(1), 59–85. Retrieved from https://doi.org/10.1146/annurev-economics-063016-103756 doi: 10.1146/annurev-economics-063016-103756

Climate Leadership Council. (2021). *Economists' statement.* (https://clcouncil.org/economists-statement/)

Colas, M. (2019). Dynamic responses to immigration. *Opportunity and Inclusive Growth Institute*.

Colas, M., & Hutchinson, K. (2021, May). Heterogeneous workers and federal income taxes in a spatial equilibrium. *American Economic Journal: Economic Policy*, *13*(2), 100-134. Retrieved from https://www.aeaweb.org/articles?id=10.1257/pol.20180529 doi: 10.1257/pol.20180529

Colas, M., & Morehouse, J. M. (2022). The environmental cost of land-use restrictions. *Quantitative Economics*, *13*(1), 179–223.

Cole, M. A. (2004). Trade, the pollution haven hypothesis and the environmental Kuznets curve: examining the linkages. *Ecological Economics*, *48*(1), 71–81.

Conover, W. J. (1971). *Practical nonparametric statistics*. New York: John Wiley & Sons.

Currie, J., Davis, L., Greenstone, M., & Walker, R. (2015). Environmental health risks and housing values: Evidence from 1,600 toxic plant openings and closings. *American Economic Review*, *105*(2), 678–709.

Curtis, E. M. (2014). *Who loses under power plant cap-and-trade programs?* (Tech. Rep.). National Bureau of Economic Research.

Dahl, C., & Duggan, T. E. (1996). U.S. energy product supply elasticities: A survey and application to the U.S. oil market. *Resource and Energy Economics*, *18*(3), 243–263.

Dahl, G. B. (2002). Mobility and the return to education: Testing a Roy model with multiple markets. *Econometrica*, *70*(6), 2367–2420.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2019, December). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, *109*(12), 4178–4219. Retrieved from https://doi.org/10.1257/aer.20180279 doi: 10.1257/aer.20180279

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, *106*(3), 479–524.

Doherty, C., Kiley, J., & Johnson, B. (2018). An examination of the 2016 electorate, based on validated voters. *Pew Research Center*.

Draxler, R., Stunder, B., Rolph, G., Stein, A., & Taylor, A. (2020). HYSPLIT user's guide (5th ed.) [Computer software manual].

Draxler, R. R., & Hess, G. D. (1998). An overview of the HYSPLIT(4) modelling system for trajectories. *Australian Meteorological Magazine*, *47*(4), 295–308.

Edelman, S. (1966). The law of federal air pollution control. *Journal of the Air Pollution Control Association*, *16*(10), 523–525. doi: 10.1080/00022470.1966.10468510

EIA. (2021). *Electricity Data Browser.* Retrieved from https://www.eia.gov/electricity/data/browser/

Emissions & Generation Resource Integrated Database. (2018). *Emissions & generation resource integrated database.* Retrieved 2020-07-03, from https://www.epa.gov/energy/emissions-generation-resource-integrated-database-egrid

EPA. (2019). Our nation's air: Status and trends through 2018.

EPA. (2020, Oct). *Co-benefits risk assessment (COBRA) health impacts screening and mapping tool.* Environmental Protection Agency. Retrieved from https://www.epa.gov/statelocalenergy/co-benefits-risk-assessment-cobra-health-impacts-screening-and-mappingtool

Fan, Q., Fisher-Vanden, K., & Klaiber, H. A. (2018). Climate change, migration, and regional economic impacts in the United States. *Journal of the Association of Environmental and Resource Economists*, *5*(3), 643–671.

Feldman, S. (2010, 10). Report: Business groups say Clean Air Act has been a 'very good investment'. *Reuters*. Retrieved 2020-07-14, from https://www.reuters.com/article/idUS212115315520101008

Fisher, R. A. (1934). *Statistical methods for research workers* (5th ed.). Edinburgh: Oliver & Boyd.

Fisher, R. A. (1935). *The design of experiments* (8th ed.). New York: Hafner Publishing Company, Inc.

Fowlie, M., Rubin, E. A., & Walker, R. (2019). *Bringing satellite-based air quality estimates down to earth* (Tech. Rep.). National Bureau of Economic Research.

Fragkias, M., Lobo, J., Strumsky, D., & Seto, K. C. (2013). Does size matter? scaling of CO2 emissions and U.S. urban areas. *PLoS One*, *8*(6), e64727.

Freeman, R., Liang, W., Song, R., & Timmins, C. (2019). Willingness to pay for clean air in China. *Journal of Environmental Economics and Management*, *94*, 188–216. Retrieved from https://doi.org/10.1016/j.jeem.2019.01.005 doi: 10.1016/j.jeem.2019.01.005

Fried, S. (2018). Climate policy and innovation: A quantitative macroeconomic analysis. *American Economic Journal: Macroeconomics*, *10*(1), 90–118.

Fried, S., Novan, K. M., & Peterman, W. (2021). Recycling carbon tax revenue to maximize welfare. *FEDS Working Paper*.

Gaigné, C., Riou, S., & Thisse, J.-F. (2012). Are compact cities environmentally friendly? *Journal of Urban Economics*, *72*(2-3), 123–136.

Gates, E. (1980, 06). No one wants backyard nuclear dump. *The Daily Press*. Retrieved 2021-10-27, from https://www.newspapers.com/image/?clipping_id=24587650

Glaeser, E. L., Gyourko, J., & Saks, R. (2005). Why is Manhattan so expensive? regulation and the rise in housing prices. *The Journal of Law and Economics*, *48*(2), 331–369.

Glaeser, E. L., & Kahn, M. E. (2010). The greenness of cities: Carbon dioxide emissions and urban development. *Journal of Urban Economics*, *67*(3), 404–418.

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, *110*(8), 2586–2624.

Goodkind, A. L., Tessum, C. W., Coggins, J. S., Hill, J. D., & Marshall, J. D. (2019). Fine-scale damage estimates of particulate matter air pollution reveal opportunities for location-specific mitigation of emissions. *Proceedings of the National Academy of Sciences*, *116*(18), 8775–8780.

Goulder, L. H., Hafstead, M. A., Kim, G., & Long, X. (2019). Impacts of a carbon tax across U.S. household income groups: What are the equity-efficiency trade-offs? *Journal of Public Economics*, *175*, 44–64.

Grainger, C., & Ruangmas, T. (2017). Who wins from emissions trading? Evidence from California. *Environmental and Resource Economics*, *71*(3), 703–727. doi: 10.1007/s10640-017-0180-1

Grainger, C., Schreiber, A., & Chang, W. (2018). Do regulators strategically avoid pollution hotspots when siting monitors? Evidence from remote sensing of air pollution. *University of Madison–Wisconsin Working Paper*.

Gray, W. B. (1997). Manufacturing plant location: Does state pollution regulation matter? *NBER Working Paper No. 5880*. Retrieved from https://www.nber.org/papers/w5880.pdf

Groom, N. (2019, 01). Six U.S. states sue Trump's EPA over interstate smog pollution rule. *Reuters*. Retrieved 2020-07-14, from https://www.reuters.com/article/us-usa-epa-lawsuit-smog/six-us-states-sue-trumps-epa-over-interstate-smog-pollution-rule-idUSKCN1PP34D

Gyourko, J., Saiz, A., & Summers, A. A. (2008). A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies*, *45*(3), 693-729.

Hafstead, M. A., & Williams, R. C. (2018). Unemployment and environmental regulation in general equilibrium. *Journal of Public Economics*, *160*, 50–65.

Heathcote, J., Storesletten, K., & Violante, G. L. (2017). Optimal tax progressivity: An analytical framework. *The Quarterly Journal of Economics*, *132*(4), 1693–1754.

Heckman, J. (1990). Varieties of selection bias. *The American Economic Review*, *80*(2), 313–318.

Henneman, L. R., Choirat, C., Ivey, C., Cummiskey, K., & Zigler, C. M. (2019). Characterizing population exposure to coal emissions sources in the United States using the HyADS model. *Atmospheric Environment*, *203*, 271–280. doi: 10.1016/j.atmosenv.2019.01.043

Henneman, L. R., Choirat, C., & Zigler, C. M. (2019). Accountability assessment of health improvements in the United States associated with reduced coal emissions between 2005 and 2012. *Epidemiology*, *30*(4), 477–485. doi: 10.1097/ede.0000000000001024

Henneman, L. R., Mickley, L., & Zigler, C. (2019). Air pollution accountability of energy transitions: The relative importance of point source emissions and wind fields in exposure changes. *Environmental Research Letters*. doi: 10.17605/OSF.IO/B8PA6

Herkenhoff, K. F., Ohanian, L. E., & Prescott, E. C. (2018). Tarnishing the Golden and Empire States: Land-use restrictions and the U.S. economic slowdown. *Journal of Monetary Economics*, *93*, 89–109.

Hernandez-Cortes, D., & Meng, K. C. (2020). Do environmental markets cause environmental injustice? Evidence from California's carbon market. *NBER Working Paper No. 27205*. Retrieved from https://www.nber.org/papers/w27205

Holland, S. P., & Mansur, E. T. (2008). Is real-time pricing green? the environmental impacts of electricity demand variance. *The Review of Economics and Statistics*, *90*(3), 550–561.

Holland, S. P., Mansur, E. T., Muller, N. Z., & Yates, A. J. (2019). Distributional effects of air pollution from electric vehicle adoption. *Journal of the Association of Environmental and Resource Economists*, *6*(S1), S65–S94. Retrieved from https://doi.org/10.1086/701188 doi: 10.1086/701188

Hsieh, C.-T., & Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, *11*(2), 1–39.

Imbens, G. W., & Rubin, D. B. (2015). *Causal inference for statistics, social, and biomedical sciences: An introduction*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139025751.006

Jeppesen, T., & Folmer, H. (2001). The confusing relationship between environmental policy and location behaviour of firms: A methodological review of selected case studies. *The Annals of Regional Science*, *35*(4), 523–546. Retrieved from https://doi.org/10.1007/s001680100055 doi: 10.1007/s001680100055

Jeppesen, T., List, J. A., & Folmer, H. (2002). Environmental regulations and new plant location decisions: Evidence from a meta-analysis. *Journal of Regional Science*, *42*(1), 19–49. Retrieved from https://doi.org/10.1111/1467-9787.00248 doi: 10.1111/1467-9787.00248

Jones, C., & Kammen, D. M. (2014). Spatial distribution of U.S. household carbon footprints reveals suburbanization undermines greenhouse gas benefits of urban population density. *Environmental science & technology*, *48*(2), 895–902.

Joskow, P. L., & Tirole, J. (2000). Transmission rights and market power on electric power networks. *The Rand Journal of Economics*, 450–487.

Kahn, M. E. (1995). A revealed preference approach to ranking city quality of life. *Journal of Urban Economics*, *38*(2), 221–235.

Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *The Quarterly Journal of Economics*, *107*(1), 35–78.

Kline, P., & Moretti, E. (2014). People, places, and public policy: Some simple welfare economics of local economic development programs. *Annual Review of Economics*, *6*(1), 629–662.

Koesler, S., & Schymura, M. (2012). Substitution elasticities in a CES production framework—an empirical analysis on the basis of non-linear least squares estimations. *ZEW-Centre for European Economic Research Discussion Paper*(12-007).

Lambiri, D., Biagi, B., & Royuela, V. (2007). Quality of life in the economic and urban economic literature. *Social Indicators Research*, *84*(1), 1.

Larson, W., & Yezer, A. (2015). The energy implications of city size and density. *Journal of Urban Economics*, *90*, 35–49.

Lawler, J. J., Lewis, D. J., Nelson, E., Plantinga, A. J., Polasky, S., Withey, J. C., Helmers, D. P., Martinuzzi, S., Pennington, D., & Radeloff, V. C. (2014). Projected land-use change impacts on ecosystem services in the United States. *Proceedings of the National Academy of Sciences*, *111*(20), 7492–7497.

Levinson, A. (1996). Environmental regulations and manufacturers' location choices: Evidence from the Census of Manufactures. *Journal of Public Economics*, *62*(1-2), 5–29. Retrieved from https://doi.org/10.1016/0047-2727(96)01572-1 doi: 10.1016/0047-2727(96)01572-1

Levinson, A. (2008). Pollution haven hypothesis. In *The new palgrave dictionary of economics* (pp. 1–5). Palgrave Macmillan UK. Retrieved from https://doi.org/10.1057/978-1-349-95121-5_2693-1 doi: 10.1057/978-1-349-95121-5_2693-1

List, J. A., Millimet, D. L., Fredriksson, P. G., & McHone, W. W. (2003). Effects of environmental regulations on manufacturing plant births: Evidence from a propensity score matching estimator. *Review of Economics and Statistics*, *85*(4), 944–952. Retrieved from https://doi.org/10.1162/003465303772815844 doi: 10.1162/003465303772815844

Livezey, E. T. (1980, 11). Hazardous waste. *The Christian Science Monitor*. Retrieved 2021-10-27, from https://www.csmonitor.com/1980/1106/110653.html

Lyubich, E. (2022). The role of people vs. places in individual carbon emissions. *HAAS Working Paper 324*.

Ma, X., Sun, D., & Cheung, K. (1999). Energy and reserve dispatch in a multi-zone electricity market. *IEEE Transactions on Power Systems*, *14*(3), 913–919.

Mangum, K. (2016). The role of housing in urban carbon emissions. *Andrew Young School of Policy Studies Research Paper Series*(16-15).

Mani, M., Pargal, S., & Huq, M. (1997). Does environmental regulation matter? Determinants of the location of new manufacturing plants in India in 1994. *World Bank Working Paper 1718*. Retrieved from https://doi.org/10.1596/1813-9450-1718 doi: 10.1596/1813-9450-1718

Martin, R., De Preux, L. B., & Wagner, U. J. (2014). The impact of a carbon tax on manufacturing: Evidence from microdata. *Journal of Public Economics*, *117*, 1–14.

McConnell, V. D., & Schwab, R. M. (1990). The impact of environmental regulation on industry location decisions: The motor vehicle industry. *Land Economics*, *66*(1), 67–81. doi: 10.2307/3146684

Mendelevitch, R., Hauenstein, C., & Holz, F. (2019). The death spiral of coal in the U.S.: Will changes in U.S. policy turn the tide? *Climate Policy*, *19*(10), 1310–1324.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., & Shi, W. (2006, 03). North american regional reanalysis. *Bulletin of the American Meteorological Society*, *87*(3), 343–360. Retrieved from https://doi.org/10.1175/bams-87-3-343 doi: 10.1175/bams-87-3-343

Millimet, D. L., & List, J. A. (2003). A natural experiment on the 'Race to the Bottom' hypothesis: Testing for stochastic dominance in temporal pollution trends. *Oxford Bulletin of Economics and Statistics*, *65*(4), 395–420. Retrieved from https://doi.org/10.1111/1468-0084.t01-1-00054 doi: 10.1111/1468-0084.t01-1-00054

Millimet, D. L., & Roy, J. (2015, February). Empirical tests of the Pollution Haven Hypothesis when environmental regulation is endogenous. *Journal of Applied Econometrics*, *31*(4), 652–677. Retrieved from https://doi.org/10.1002/jae.2451 doi: 10.1002/jae.2451

MIT. (2018). *County Presidential Election Returns 2000-2020.* Harvard Dataverse. Retrieved from https://doi.org/10.7910/DVN/VOQCHQ doi: 10.7910/DVN/VOQCHQ

Mitchell, R. C., & Carson, R. T. (1986). Property rights, protest, and the siting of hazardous waste facilities. *The American Economic Review*, *76*(2), 285–290.

Monogan, J. E., Konisky, D. M., & Woods, N. D. (2017). Gone with the wind: Federalism and the strategic location of air polluters. *American Journal of Political Science*, *61*(2), 257–270.

Morehouse, J., & Rubin, E. (2021). Downwind and out: The strategic dispersion of power plants and their pollution. *Available at SSRN 3915247*.

Moretti, E. (2011). Local labor markets. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 4, pp. 1237–1313). Amsterdam: Elsevier.

Mu, Y., Rubin, E., & Zou, E. (2021). What's missing in environmental (self-)monitoring: Evidence from strategic shutdowns of pollution monitors. *NBER Working Paper No. 28735*.

Muehlenbachs, L., Spiller, E., & Timmins, C. (2015). The housing market impacts of shale gas development. *American Economic Review*, *105*(12), 3633–59.

Muller, N. Z., & Mendelsohn, R. (2007). Measuring the damages of air pollution in the United States. *Journal of Environmental Economics and Management*, *54*(1), 1–14.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The computer journal*, *7*(4), 308–313.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, 395–421.

Nordhaus, W. D. (2017). Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, *114*(7), 1518–1523.

North American Regional Reanalysis. (2006). *North american regional reanalysis.* Retrieved 2020-07-03, from https://psl.noaa.gov/data/gridded/data.narr.pressure.html

Notowidigdo, M. J. (2013). *The incidence of local labor demand shocks.* (Unpublished, University of Chicago Booth School of Business)

Oates, W. E. (1972). *Fiscal federalism.* Harcourt.

Oates, W. E. (1999). An essay on fiscal federalism. *Journal of Economic Literature*, *37*(3), 1120–1149.

Oates, W. E. (2002). *A reconsideration of environmental federalism* (J. A. List & A. de Zeeuw, Eds.). UK: Edward Elgar.

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439–446. Retrieved from https://doi.org/10.32614/RJ-2018-009 doi: 10.32614/RJ-2018-009

Piyapromdee, S. (2019). *The impact of immigration on wages, internal migration and welfare* (Tech. Rep.). Working paper.

Piyapromdee, S. (2021). The impact of immigration on wages, internal migration, and welfare. *The Review of Economic Studies*, *88*(1), 406–453.

Preonas, L. (2019). Market power in coal shipping and implications for u.s. climate policy. *Working paper*.

Quigley, J. M., & Raphael, S. (2005). Regulation and the high cost of housing in California. *American Economic Review*, *95*(2), 323–328.

Rangel, M. A., & Vogl, T. S. (2019, October). Agricultural fires and health at birth. *The Review of Economics and Statistics*, *101*(4), 616–630. Retrieved from https://doi.org/10.1162/rest_a_00806 doi: 10.1162/rest_a_00806

Rausch, S., Metcalf, G. E., & Reilly, J. M. (2011). Distributional impacts of carbon pricing: A general equilibrium approach with micro-data for households. *Energy Economics*, *33*, S20–S33.

Ren, X., Hall, D. L., Vinciguerra, T., Benish, S. E., Stratton, P. R., Ahn, D., Hansford, J. R., Cohen, M. D., Sahu, S., He, H., et al. (2019). Methane emissions from the marcellus shale in southwestern pennsylvania and northern west virginia based on airborne measurements. *Journal of Geophysical Research: Atmospheres*, *124*(3), 1862–1878.

Revesz, R. L. (1996). Federalism and interstate environmental externalities. *University of Pennsylvania Law Review*, *144*(6), 2341–2416.

Ruggles, S., Alexander, J. T., Genadek, K., Goeken, R., Schroeder, M. B., & Sobek, M. (2010). *Integrated public use microdata series: Version 5.0 [machine-readable database].* (Minneapolis: University of Minnesota)

Ryaboshapko, A., Bullock, O. R., Christensen, J., Cohen, M., Dastoor, A., Ilyin, I., Petersen, G., Syrakov, D., Artz, R. S., Davignon, D., Draxler, R. R., & Munthe, J. (2007). Intercomparison study of atmospheric mercury models: 1. comparison of models with short-term measurements. *Science of The Total Environment*, *376*(1-3), 228–240. doi: 10.1016/j.scitotenv.2007.01.072

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, *125*(3), 1253–1296.

Sallee, J. M. (2019). *Pigou creates losers: On the implausibility of achieving Pareto improvements from efficiency-enhancing policies* (Tech. Rep.). National Bureau of Economic Research.

Sanders, L. (2018, 01). EPA to be sued over power plant pollution spreading into neighboring state. *Newsweek*. Retrieved 2020-07-14, from https://www.newsweek.com/epa-sued-over-power-plant-pollution-spreading-neighboring-state-768445

Schlenker, W., & Walker, W. R. (2016, October). Airports, air pollution, and contemporaneous health. *The Review of Economic Studies*, *83*(2), 768–809. Retrieved from https://doi.org/10.1093/restud/rdv043 doi: 10.1093/restud/rdv043

Scott, J. B. (2021a). Positive spillovers from infrastructure investment: How pipeline expansions encourage fuel switching. *The Review of Economics and Statistics*, 1–43.

Scott, J. B. (2021b). *Positive spillovers from infrastructure investment: How pipeline expansions encourage fuel switching* (Tech. Rep.). Working Paper.

Sergi, B., Azevedo, I., Davis, S. J., & Muller, N. Z. (2020, October). Regional and county flows of particulate matter damage in the U.S. *Environmental Research Letters*, *15*(10), 104073. Retrieved from https://doi.org/10.1088/1748-9326/abb429 doi: 10.1088/1748-9326/abb429

Serletis, A., Timilsina, G. R., & Vasetsky, O. (2010). Interfuel substitution in the United States. *Energy Economics*, *32*(3), 737–745.

Shadbegian, R., & Wolverton, A. (2010). Location decisions of u.s. polluting plants: Theory, empirical evidence, and consequences. *U.S. EPA Working Paper #10-05*.

Stein, A. F., Draxler, R. R., Rolph, G. D., Stunder, B. J. B., Cohen, M. D., & Ngan, F. (2015). NOAA's HYSPLIT atmospheric transport and dispersion modeling system. *Bulletin of the American Meteorological Society*, *96*(12), 2059–2077. doi: 10.1175/bams-d-14-00110.1

Stein, A. F., Rolph, G. D., Draxler, R. R., Stunder, B., & Ruminski, M. (2009). Verification of the NOAA smoke forecasting system: Model sensitivity to the injection height. *Weather and Forecasting*, *24*(2), 379–394.

Stunder, B. J. B., Heffter, J. L., & Draxler, R. R. (2007). Airborne volcanic ash forecast area reliability. *Weather and Forecasting*, *22*(5), 1132–1139.

Suárez Serrato, J. C., & Zidar, O. (2016). Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms. *American Economic Review*, *106*(9), 2582–2624.

Sullivan, D. M. (2016). The true cost of air pollution: Evidence from house prices and migration. *Harvard Environmental Economics Program, Discussion Paper 16-69*.

Tange, O. (2011). GNU parallel - the command-line power tool. *;login: The USENIX Magazine*, *36*(1), 42–47.

Tessum, C. W., Hill, J. D., & Marshall, J. D. (2017). InMAP: A model for air pollution interventions. *PLOS ONE*, *12*(4), e0176131. Retrieved from https://doi.org/10.1371/journal.pone.0176131 doi: 10.1371/journal.pone.0176131

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, *64*(5), 416–424.

United States Congress (90th). (1968, February). The air quality act of 1967. *Journal of the Air Pollution Control Association*, *18*(2), 62–71. Retrieved from https://doi.org/10.1080/00022470.1968.10469096 doi: 10.1080/00022470.1968.10469096

United States Senate, Committee on Public Works, Staff Report. (1963). *A study of pollution—air* (Tech. Rep.). 88th, Congress, 1st Session.

US Census Bureau. (2016a). *Cartographic boundary shapefiles.* Retrieved 2020-10-15, from https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2016.html

US Census Bureau. (2016b). *Tiger/lines shapefiles.* Retrieved 2020-07-14, from https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2016.html

U.S. Environmental Protection Agency. (2013a). *Clean Air Act, Section 110 (42 u.s.c. 7401)* (Tech. Rep.). Retrieved from https://www.govinfo.gov/content/pkg/USCODE-2013-title42/html/USCODE-2013-title42-chap85-subchapI-partA-sec7410.htm

U.S. Environmental Protection Agency. (2013b). *Clean Air Act, Section 126 (42 u.s.c. 7426)* (Tech. Rep.). Retrieved from https://www.govinfo.gov/content/pkg/USCODE-2013-title42/html/USCODE-2013-title42-chap85-subchapI-partA-sec7426.htm

U.S. Environmental Protection Agency. (2017). *Green book national area and county-level multi-pollutant information.* Retrieved 2020-06-25, from https://www.epa.gov/green-book/green-book-national-area-and-county-level-multi-pollutant-information

U.S. Environmental Protection Agency. (2018, 07). *2014 National Emissions Inventory, v2, technical support document* (Tech. Rep.). Retrieved from https://www.epa.gov/sites/production/files/2018-07/documents/nei2014v2_tsd_05jul2018.pdf

U.S. Environmental Protection Agency. (2020, 03). *Overview of the Cross-State Air Pollution Rule (CSAPR).* Retrieved 2020-07-14, from https://www.epa.gov/csapr/overview-cross-state-air-pollution-rule-csapr

U.S. Government Accountability Office. (2011, 05). *Information on tall smokestacks and their contribution to interstate transport of air pollution (gao-11-473)* (Tech. Rep.). Institution.

U.S. House of Representatives. (2019). *Stemming Warming and Augmenting Pay Act.* (https://www.congress.gov/bill/116th-congress/house-bill/4058)

Volcovici, V. (2018, 01). Delaware to sue EPA over upwind air pollution. *Reuters*. Retrieved 2020-07-14, from https://www.reuters.com/article/us-usa-epa-smog/delaware-to-sue-epa-over-upwind-air-pollution-idUSKBN1ER1IN

Walker, W. R. (2013). The transitional costs of sectoral reallocation: Evidence from the Clean Air Act and the workforce. *The Quarterly Journal of Economics*, *128*(4), 1787–1835.

Wang, Y., Bechle, M. J., Kim, S.-Y., Adams, P. J., Pandis, S. N., Pope, C. A., Robinson, A. L., Sheppard, L., Szpiro, A. A., & Marshall, J. D. (2020). Spatial decomposition analysis of NO2 and PM2.5 air pollution in the united states. *Atmospheric Environment*, *241*. Retrieved from https://doi.org/10.1016/j.atmosenv.2020.117470 doi: 10.1016/j.atmosenv.2020.117470

Williams III, R. C., Gordon, H., Burtraw, D., Carbone, J. C., & Morgenstern, R. D. (2015). The initial incidence of a carbon tax across income groups. *National Tax Journal*, *68*(1), 195.

Woerman, M. (2020). Market size and market power: Evidence from the Texas electricity market. *HAAS Working paper 298*.

Wolverton, A. (2009). Effects of socio-economic and input-related factors on polluting plants' location decisions. *The B.E. Journal of Economic Analysis & Policy*, *9*. doi: 10.2202/1935-1682.2083

World Bank. (2020). *State and trends of carbon pricing 2020* (Tech. Rep.). Retrieved from https://openknowledge.worldbank.org/handle/10986/33809

Yamazaki, A. (2017). Jobs and climate policy: Evidence from British Columbia's revenue-neutral carbon tax. *Journal of Environmental Economics and Management*, *83*, 197–216.

Yip, C. M. (2018). On the labor market consequences of environmental taxes. *Journal of Environmental Economics and Management*, *89*, 136–152.

Yuan, M., Rausch, S., Caron, J., Paltsev, S., & Reilly, J. (2019). The MIT U.S. regional energy policy (USREP) model: The base model and revisions. *Joint Program Technical Note TN*, *18*.

Zivin, J. G., Liu, T., Song, Y., Tang, Q., & Zhang, P. (2020, November). The unintended impacts of agricultural fires: Human capital in China. *Journal of Development Economics*, *147*, 102560. Retrieved from https://doi.org/10.1016/j.jdeveco.2020.102560 doi: 10.1016/j.jdeveco.2020.102560

Zou, E. Y. (2021). Unwatched pollution: The effect of intermittent monitoring on air quality. *American Economic Review*, *111*(7), 2101–2126.