A Study Design to Measure Health Benefits from Coal Power Plant Closures

Adrienne Ohler, Associate Extension Professor, Division of Applied Social Science, University of Missouri
adrienne.ohler@missouri.edu

Selected Paper prepared for presentation at the 2023 Agricultural & Applied Economics Association Annual Meeting, Washington DC; July 23-25, 2023

copies.

Title: *A Study Design to Measure Health Benefits from Coal Power Plant Closures*

Author: Adrienne Ohler, Associate Extension Professor, Division of Applied Social Science, University of Missouri, 222 Mumford Hall, Columbia, MO 65211

Abstract:
In the last decade, significant infrastructure changes occurred in the U.S. electricity sector. Between 2010-2018, 546 coal-fired power units retired, while natural gas and renewable developments grew. These infrastructure changes allow for a natural experiment examining the impact of power-plant emissions on population health, and thus can be used to study the health impacts of coal emissions on respiratory diseases. Using survey data from the National Center for Health Statistics, this paper proposes a study design to examine health benefits and estimate changes in health disparities from the closure of U.S. coal power plants. Respondent data from the National Health Interview Survey on respiratory health outcomes can be merged with the location of power plants and climatic wind direction. We provide a method to merge restricted access data and estimate a statistical model of health outcomes that allows data to remain deidentified while comparing across disparity dimensions (e.g. income, race), wind direction, and coal power plant closures. The findings contribute to the fields of environmental health economics, and energy economics by quantifying the benefit of improved health from coal plant closures across racial and income subgroups.

I. Introduction

In the last decade, significant infrastructure changes occurred in the U.S. electricity sector. The shale gas revolution led to the North American bulk power system adding 180 GW of natural gas–fired generation in 2009-2019 (NRECA, 2019). Improvements in network operations encouraged renewable developments, while 546 coal power plant units retired (Johnson & Chau, 2019). Between 2000-2018, particulate matter ($PM_{2.5}$) concentrations from electricity fell 89% (Hernandez-Cortes et al., 2022). These infrastructure changes allow for a natural experiment examining the impact of power-plant emissions on population health, and thus can be used to study the health impacts of coal emissions on childhood and adult asthma, and respiratory and pulmonary diseases.

Using survey data from the National Center for Health Statistics, this paper proposes a study design to examine health benefits and estimate changes in health disparities from the closure of U.S. coal power plants. Respondent data from the National Health Interview Survey (NHIS) on respiratory health outcomes can be merged with the location of power plants provided by the Energy Information Administration (EIA). Following a health economics framework, we provide a method to merge restricted access data and estimate a statistical model of health outcomes that allows data to remain deidentified but allows for comparisons across disparity dimensions (e.g. income, race), wind direction, and coal power plant closures. The difference-in-difference model captures the reduced health risk from a power plant closure relative to upwind residents (Abadie, 2005; Angrist & Pischke, 2008; Bertrand et al., 2004; Gertler et al., 2011). The estimated reduction in relative risk from the power plant closure can then be used to estimate the healthcare cost savings from power plant closures using Medical Expenditure Panel Survey data. Results for the study

design can then provide policymakers an estimate of health benefits from coal power plants closure, in terms of the reduction in healthcare costs from plant closure across income and race.

II. Background Literature

In the epidemiological literature, coal plants have been linked to all-cause and premature mortality, respiratory disease and lung cancer, cardiovascular disease, poorer child health, and higher infant mortality (Gupta & Spears, 2017). Causal methods have been used to show that either expansion or closures of coal-fired and fossil fuel plants have wide ranging health effects with closures reducing premature mortality (Burney, 2020; Fraenkel et al., 2022); improving pregnancy and birth outcomes (DeCicca & Malak, 2020; Keil et al., 2021; Wilkie et al., 2023; Yang & Chou, 2018); reducing absenteeism in school (Komisarow & Pakhtigian, 2022); decreasing ER visits (Komisarow & Pakhtigian, 2022); and reducing medical expenditures and improving resident health (Jia & Luo, 2023). Expansions of coal use in India have also led to an increase in respiratory disease (Gupta & Spears, 2017).

Disparities in ambient air pollution exposure has also been well-documented (Cushing et al., 2023; Goforth & Nock, 2022; Hajat et al., 2015; Henneman et al., 2023; Hernandez-Cortes et al., 2022; Jbaily et al., 2022; Spiller et al., 2021). Evidence suggests that minority and low-income populations are exposed to physical, biological, and chemical environmental hazards more frequently, and these byproducts of energy consumption are not evenly distributed across populations (Abel & White, 2011; Ash & Boyce, 2018; Clark et al., 2014; Miranda et al., 2011; Ou et al., 2008; Payne-Sturges & Gee, 2006; Tessum et al., 2019). Exposure rates to fine particulate matter from electricity generation is highest for Blacks, and low-income households (Thind et al., 2019). Notably, Cushing et al. (2023) and Lane et al. (2022) find that historical red-lining led to more fossil fuel powerplants be sited in neighborhoods deemed “hazardous” during the Great

Depression era of the 1930s by the Federal Housing Administration, leading to greater pollution exposure for Black communities. Hajat et al. (2015) and Jbaily et al. (2022) also finds that low-income populations are exposed to higher levels of particulate matter ($PM_{2.5}$). Henneman et al., (2023) and Hernandez-Cortes et al. (2022) have modelled the reduction in pollution exposure driven by the retirement of fossil fuel electric generators and showed that since 2000, the Black-White $PM_{2.5}$ disparity has narrowed by 93%, and this narrowing can be attributed largely to changes in the electric generation sector with only small effects due to people changing residences. However, they find relatively small disparities of $PM_{2.5}$ exposure by income over time. Henneman et al. (2023) notes that most of the exposure reduction after 2010 is attributable to retirements, whereas prior to 2010 reductions were attribute to scrubber installations. In the early 2000s, Black populations in the South and North Central U.S. were exposed to $PM_{2.5}$ at higher rates compared to average population exposure, and these disparities decreased with falling emissions. However, Black populations in the North Central region continue to experience higher rates of $PM_{2.5}$ exposure than the population average.[1]

Several studies have examined exposure disparities due to decarbonization policies, although the findings have been mixed. For example, Gallagher and Holloway (2022) find that decarbonization can lead to reductions in relative exposure for Black and Asian populations depending on the policy scenarios. However, Richmond-Bryant et al. (2020) find that historical changes in $PM_{2.5}$ led to reductions in premature mortality rates, but shifted the proportional burden from particulate matter from White to non-white subgroups. Mejia-Duwan et al. (2023) show that changes in air pollution from vehicle electrification is not equally distributed with $PM_{2.5}$, NO_X,

[1] The North Central region includes the states Ohio, Michigan, Indiana, Wisconsin, Illinois, Minnesota, Iowa, Missouri, Kansas, Nebraska, and North and South Dakota. The South region includes states as far west as Texas, and as far North as Maryland.

and SO_2 emission reductions disproportionately occurring in the least disadvantaged communities. Finally, Thomson et al. (2018) find that coal mining counties, White and rural counites will experience some of the best improvements in air quality and specifically $PM_{2.5}$ from the adoption of the Clean Air Act.

Yet, the impact of these pollution reductions do suggest a benefit in terms of improved health to Black and low-income communities. Using a difference-in-difference approach, Casey et al. (2018) show a decline in pre-term birth from coal and oil power plant retirements, showing a larger reduction among non-Hispanic black and Asian mothers relative to non-Hispanic white and Hispanic mothers. Using atmospheric chemistry modelling, Qiu et al. (2022) find that mortality rates declined due to increase wind power driven by renewable portfolio standards, and people living in high PM areas experienced a greater health benefit relative to the average. Black and low-income households experienced a larger benefit and Hispanic residents experienced a smaller than average benefit.

The environmental health literature has yet to thoroughly explore the trade-off between economic activities that produce valued goods and services, such as fossil fuel produced electricity and the health costs and negative externalities created through its production. Economic research has shown minority groups generally receive a greater share of pollution risks but fewer employment benefits from living near US industrial facilities (Ash & Boyce, 2018); and Blacks and Hispanics bear a greater pollution burden from exposure to particulate matter $PM_{2.5}$ than non-Hispanic Whites, relative to their level of consumption of goods and services (Tessum et al., 2019). In this paper, we aim to contribute to this literature by assessing the energy-health trade-off from coal power plant closures. Understanding and quantifying this tradeoff between electricity production, public health costs, and health disparities is relevant to public and population health

experts, epidemiologists, economists, environmental health scientists, energy planners, regulators, and policymakers.

III. Data

The proposed research design relies on linking power plant location with survey respondent data from the National Center for Health Statistics' National Health Interview Survey (NHIS). The NHIS is a repeated cross-sectional nationwide survey useful for this proposal in terms of sample size, demographic data, and health outcome metrics. For this proposal, the major strength of the NHIS survey is the ability to examine demographic and socioeconomic characteristics across health outcomes for a representative sample of the U.S. Between 2009-2018, the core survey has remained largely unchanged, capturing household demographics, adult and child health status, and health behaviors.

Table 1. shows statistics for respiratory health outcomes by race and income subgroups. NHIS contains participant responses to medical conditions including general health status, and we focus on the following respiratory health outcomes: chronic bronchitis, chronic obstructive pulmonary disease (COPD), emphysema, lung cancer, if they had a cold in the past 2 weeks, if they have ever been told they have asthma, if they had an asthma attack in the previous year, and if they visited the ER for asthma in the previous year. Respondents do not answer all questions, and table 1 provides the number of respondents for each health outcome, and by subgroup based on publicly available data (Blewett et al., 2022). From this sample, we calculate unweighted statistics for each health outcomes to demonstrate the sample size before matching with powerplants. These statistics highlight the health disparities between Black and White respondents with Black adults showing higher rates of fair or poor health, chronic bronchitis, lung cancer, having a cold, and asthma. When comparing across incomes, respondents in a household below

the federal poverty line (FPL) have worse rates for all measured health outcomes. These results are consistent when we use survey weights to calculate population estimates, as provided in Appendix A.

Table 1 Descriptive Statistics for Adult Health Outcomes NHIS 2009-2018

Respondents		Full Sample	Black	White	Above FPL	Below FPL
General Health Status	Respondents N=	955,782	138,123	730,316	721,396	141,438
Fair/Poor	N	100,014	19,851	72,380	63,270	25,475
	(%)	(10.5%)	(14.4%)	(9.9%)	(8.8%)	(18.0%)
Chronic bronchitis past year	Respondents N=	312,288	45,537	242,192	242,585	48,567
Yes	N	13,374	2,208	10,575	9,278	3,226
	(%)	(4.3%)	(4.9%)	(4.4%)	(3.8%)	(6.6%)
COPD	Respondents N=	224,367	30,998	175,842	176,863	33,988
Yes	N	8,607	921	7,395	6,074	1,988
	(%)	(3.8%)	(3.0%)	(4.2%)	(3.4%)	(5.9%)
Emphysema	Respondents N=	312,269	45,541	242,168	242,580	48,574
Yes	N	6,086	660	5,196	4,082	1,531
	(%)	(2.0%)	(1.5%)	(2.2%)	(1.7%)	(3.2%)
Lung Cancer	Respondents N=	30,017	2,519	26,466	24,387	3,230
Yes	N	954	124	794	710	164
	(%)	(3.2%)	(4.9%)	(3.0%)	(2.9%)	(5.1%)
Had a cold in the past 2 weeks	Respondents N=	427,586	64,504	327,070	329,461	70,373
Yes	N	52,084	8,186	39,981	38,797	10,539
	(%)	(12.2%)	(12.7%)	(12.2%)	(11.8%)	(15.0%)
Ever been told had asthma	Respondents N=	427,479	64,490	326,977	329,404	70,352
Yes	N	56,203	10,617	41,550	41,301	11,774
	(%)	(13.2%)	(16.5%)	(12.7%)	(12.5%)	(16.7%)
Asthma attack in last year	Respondents N=	56,126	41,490	10,607	41,254	11,757
Yes	N	17,980	3,616	13,091	12,492	4,532
	(%)	(32.0%)	(34.1%)	(31.6%)	(30.3%)	(38.6%)
ER visit for asthma in last year	Respondents N=	52,710	9,902	38,999	38,773	11,090
Yes	N	5,522	1,778	3,343	3,289	1,913
	(%)	(10.5%)	(18.0%)	(8.6%)	(3.5%)	(17.3%)

For control variables, we also collect data on determinants of respiratory health such as biological factors (age, BMI, sex), socio-economic status (income, marital status, race, ethnicity, education, access to paid sick leave, owns their own home), access to healthcare access (has a usual place when sick, has delayed medical care in the last 12 months,), health impacting behaviors (smoker status, alcohol frequency, moderate exercise frequency), and occupation and environment (Dodd & Mazurek, 2016; Lee et al., 2006; Robinson et al., 2011; Senthilselvan et al., 2020). Table 2 provides the estimated population rates for fair or poor health status across categories of each covariate. These variables account for variation in health status across socio-economic factors, occupation type, access to healthcare, and health-related behaviors. We see that rates of fair or poor health are higher across socio-economic factors: females, renters, some college or less for education, and being widowed, or divorced or separated. Construction, production, transportation, or installation and repair occupations have higher rates of fair or poor health as do workers in construction, mining, and manufacturing industries. Having access to paid sick leave is associated with lower rates of fair or poor health, while delaying medical treatment due to concerns about cost is associated with higher rates of fair or poor health. Finally, respondents that report smoking, or a lack of moderate activity exercising exhibit higher rates of fair or poor health.

Several studies have used the NHIS data to examine air pollution and health outcomes such as childhood respiratory allergies (Parker et al., 2009), mortality (Pope et al., 2018; Pope III et al., 2019), asthma (Nachman & Parker, 2012), all cancer mortality (Coleman et al., 2020), and heart disease (Parker et al., 2018) using the Environmental Protection Agency's Air Quality System for $PM_{2.5}$, SO_2, and NO_2 to measure exposure. Because of its sampling scheme, NHIS has the framework to merge data to the participants' residential geocodes developed by the Dept.

Table 2 Rates of Fair/Poor Health Status across Covariates

	General Health Status is Fair/Poor			General Health Status is Fair/Poor			General Health Status is Fair/Poor	
	Yes	P-value		Yes	P-value		Yes	P-value
Full Population	10.06%		**Occupation Type**			**Health-Related Behaviors**		
Socio-Economic Factors			Construction Occupation		<0.01	Ever smoked 100 cigarettes		<0.01
Male	9.44%	<0.01	No	11.68%		No	10.42%	
Female	10.65%		Yes	19.97%		Yes	19.05%	
Home Ownership		<0.01	Construction Industry		<0.01	Current Smoker		<0.01
Own	9.08%		No	12.49%		Current	21.48%	
Rent	11.97%		Yes	17.58%		Former	17.21%	
						Never	10.42%	
Education		<0.01	**Access to Healthcare**					
HS or less	12.94%		Paid Sick Leave		<0.01	Alcohol Frequency		<0.01
Associates	11.64%		No	15.81%		daily	11.68%	
Some college	10.96%		Yes	11.32%		<weekly	14.34%	
Bachelors	5.63%					weekly	7.32%	
Masters+	4.82%		Has Usual Place for Healthcare		<0.01	never	26.97%	
			Many	13.46%				
Marital Status		<0.01	No	8.23%		Moderate Activity		<0.01
Married/Living with Partner	11.09%		Yes	11.13%		daily	9.48%	
Separated/ Divorced	20.84%					<weekly	12.16%	
Widowed	27.19%		Medical Cost Delayed Treatment in Last Year		<0.01	weekly	7.96%	
Never Married	4.84%		No	8.86%		never	22.11%	
			Yes	24.23%				

P-values compare statistics across covariates based on a Chi-squared test.

of Housing and Urban Development (HUD) which can then be merged by NCHS statisticians to power-plant locations and removed. The process to merge and deidentify data requires several steps. First, researchers must cross power-plant locations with resident locations to calculate distance and bearing from the plant with the bearing later used to determine wind direction. Next, we identify people within 30 miles of only 1 coal power plant, and censor those respondents outside of 30 miles or within 30 miles of 2 or more power plants. From this nearest plant, we retain the wind direction, whether the plant operated in the respondent's survey year, and number of years since closure. Finally, NCHS statisticians mask any identifying information including the powerplant location identifiers, and external variables not needed in the final analysis, such as initial bearing.

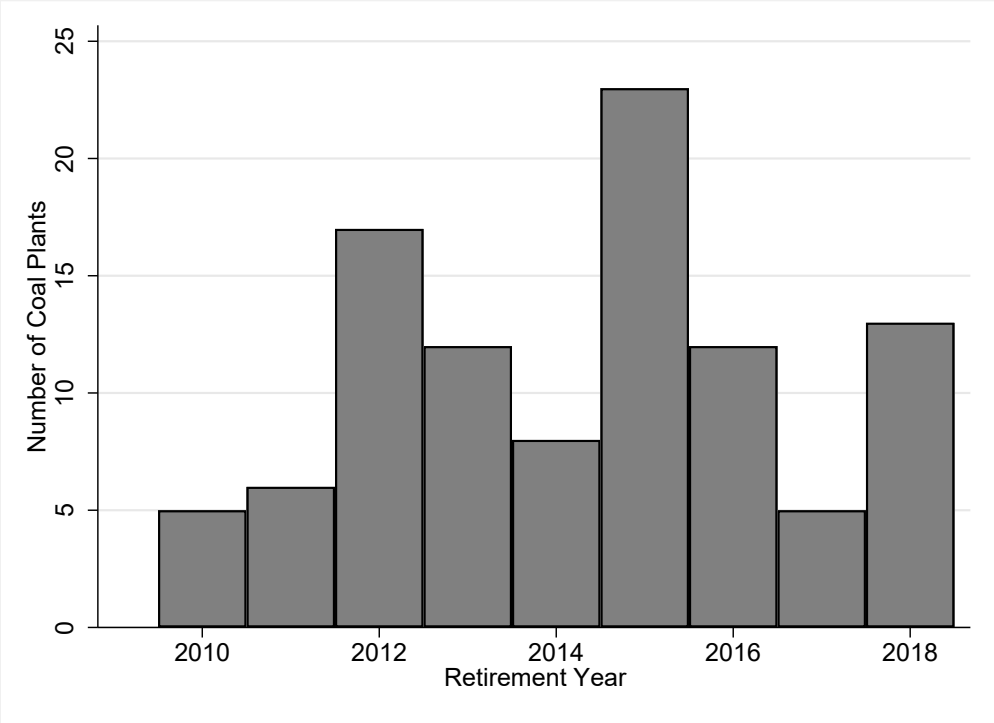
The EIA Form 860 collects generator-level information including retirements, historical production, and planned capacity. Location and retirement year information from 1,114 generators for 468 power plants were acquired for the time period 2009-2018. Inclusion criteria for generators require that the generator be operating for the entire reporting year of 2018, as opposed to being on standby or out of service for some or all of the period to ensure exposure. As a result, no power plants are excluded from this operating inclusion criteria; however, we exclude plants with partial retirements, n=38 (i.e. some generators remained operational), or with generators' retirement years that varied more than 0.5 standard deviations, n=24. Occasionally, new generators are added to existing power plants, and we exclude these plants by removing generators starting operations after 2008 and with generators' first operating years that varied more than 0.5 standard deviations, to ensure a discrete starting date. Additionally, retired plants with retirements before 2009 are excluded. The final data set consists of 382 power plants of which 101 closed with discrete closure years. A map of the excluded power plants is provided in Appendix B Figure B1.

Two cohorts of powerplants are derived: intervention and non-intervention groups. Intervention plants observed a retirement on or before 2016, and non-intervention plants retired after 2016 or are still operating as of 2018. For plants with multiple generators, retirement year was defined as the highest retirement year of all generators. Plants with several years between generator retirements (>0.5 standard deviations) and thus an unclear retirement year are excluded from the analysis. Figure 1 graphs (a) the number of retirements over the study period, and (b) the total nameplate capacity retired in our study sample. These figures highlight the three highest retirement years occurred in 2015, 2012, and 2018 with 2015 having the highest of plant retirements and 2018 having the largest amount of retirements based on nameplate capacity.

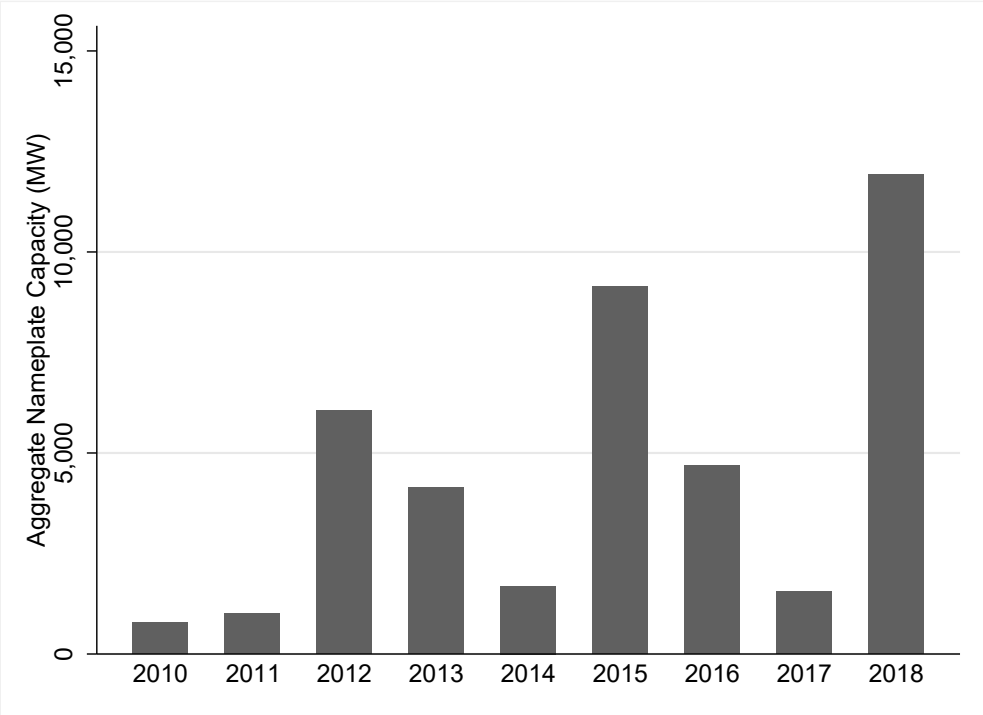
For the analysis, the intervention group includes NHIS respondents within 30 miles of a retired plant with at least 2 years of health data post-retirement. For each location, we will evaluate at least five years of NHIS data with two years before the retirement and two years after. For those respondents in the non-intervention group (ie. non-retirement plants), we will use 2013-2017 as the comparison period. A sensitivity analysis will be performed to compare the robustness of results across all possible non-intervention counterfactual periods: 2009-2013, 2010-2014, 2011-2015, 2012-2016, and 2014-2018.

For perspective on where these retirements are occurring, Figure 2 maps power plant retirements with red circles indicating retired plant locations and blue circles designating operating plants as of 2018. Each circle's radius is 30 miles. The selection criteria includes only NHIS respondents within 30 miles of only one power plant. Thus, much of the respondents in the final dataset will include individuals in the West, Midwest, and South where powerplants are spaced farther apart.

Figure 1: Coal power plant retirements 2009-2018.

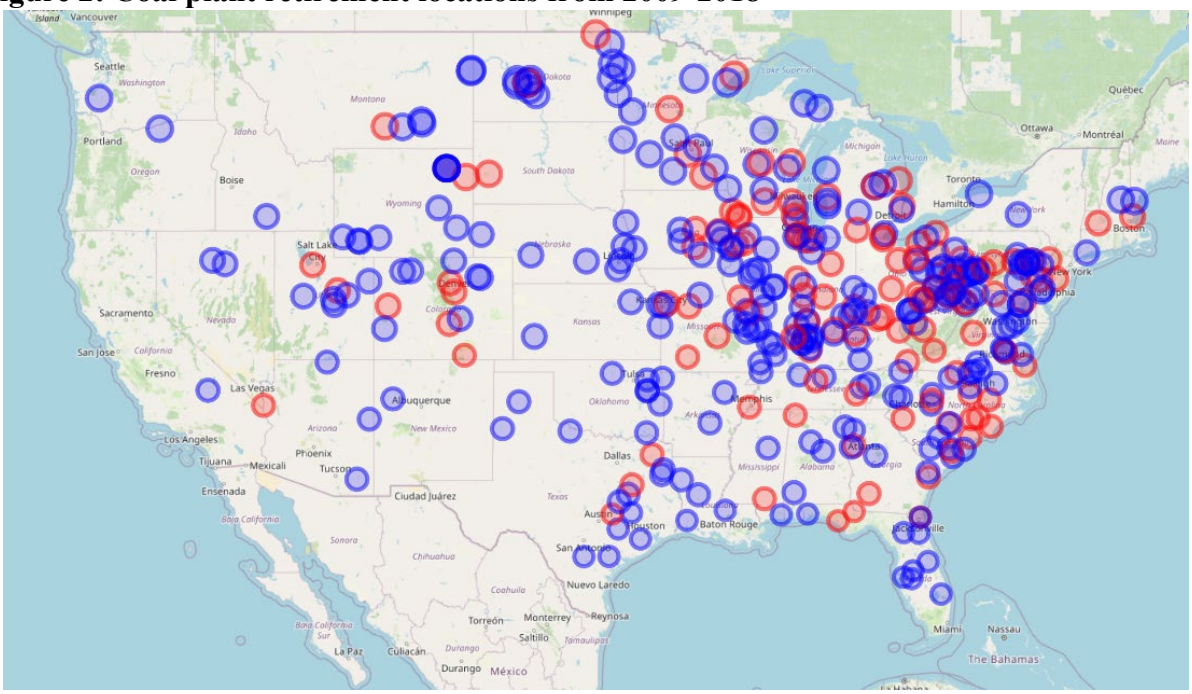


(a) Number of plant retirements by year



(b) Total retired nameplate capacity by year

Figure 2: Coal plant retirement locations from 2009-2018



Red circles indicate retired plant locations and blue circles indicate operating plants as of 2018. Each circle's radius is 30 miles.

To proxy for pollution exposure, we rely on wind direction comparing downwind to upwind households to evaluate exposed households relative to an unexposed cohort. Linking current and former coal sites to weather data is critical for accurate analysis of such proxies (Deryugina et al., 2019). Location specific wind data are available for land based stations near energy producing facilities from the National Oceanic and Atmospheric Administration. Wind direction data are collected and modelled by Visual Crossing using the nearest weather stations to interpolate data on daily wind speed, gusts, and direction for a given latitude and longitude coordinates (Visual Crossing Corportion, Reston VA). For each location, we calculate the average wind direction over the year 2018 from daily wind directions using the CIRCULAR statistics module in Stata 17 (StataCorp. 2021. College Station, TX). Upwind and downwind cohorts are determined by comparing the initial bearing of the household relative to the closest power plant,

and a 90° interval around dominant wind direction is used to define upwind and downwind locations. For this analysis, we will exclude respondents that fall outside the upwind or downwind region.

IV. Method

The difference-in-difference method measures the reduced health risk from a power plant closure by comparing a group exposed to the plant's pollutants (i.e. downwind residents) to a control group (i.e. upwind residents) before and after the closure. This approach has been commonly used in epidemiological and economics literature measuring, for example, improvements in birth outcomes from coal retirements and desulfurization (Casey et al., 2018; Fraenkel et al., 2022; Jha et al., 2019; Jia & Luo, 2023; Luechinger, 2014; Yang & Chou, 2018; Yi & Sung, 2021). To expand on this model, we propose to follow a health economics framework to assess health outcomes while incorporating health behaviors and evaluating health disparities (Brunello et al., 2016).

The model of health outcomes (i.e. fair/poor health, asthma, lung cancer) for individual *i* is based on the operations of the nearest coal power plant and is modelled econometrically through a log link model:

(1) $H_{it} \sim F(\pi_i)$

(2) $\log(\pi_i) = c + \gamma_1 W_i + \gamma_2 C_i + \gamma_3 W_i \cdot C_i + \gamma_4 T_i + \sum_{j=1}^{n} \alpha_j B_{ji} + \boldsymbol{\delta X_i} + \beta DD_i$

where H_{it} is a binary outcome of respondent i's health, assuming a Poisson distribution to model the probability π_{it} of the health outcome. The log Poisson model can then be used to estimate the incidence rate ratio (IRR), as well as the average risk difference (RD). W_i is an indicator representing the respondent's upwind/downwind direction relative to the coal power plant site, and

C_{it} is an indicator for the coal site operating or closed in the year of the health interview. The difference-in-difference component of this model is captured by the interaction of W_i and C_i: $\gamma_1 W_i + \gamma_2 C_i + \gamma_3 W_i \cdot C_i$, where the parameter γ_1 represents the impact of wind direction on health outcomes, γ_2 the effect of the closure, and γ_3 is the interaction between residents upwind and downwind-before and after closure. T_i is a discrete variable representing the number of years since closure, and controls for aggregate time effects.

Using a marginal standardization approach to calculate IRR and RD, we calculate the expected predicted probabilities to assess risk for the four cohorts, such that the probability of the health outcome for the four groups are notated:

(3) $\hat{\pi}_{11} = \frac{1}{m}\sum_i^m exp\left(c + \widehat{\gamma_1} + \widehat{\gamma_2} + \widehat{\gamma_3} + \widehat{\gamma_4} + \sum_{j=1}^n \widehat{\alpha_j} B_{ji} + \widehat{\boldsymbol{\delta}}\boldsymbol{X_i} + \hat{\beta} DD_i \right)$

(4) $\hat{\pi}_{01} = \frac{1}{m}\sum_i^m exp\left(c + \widehat{\gamma_2} + \widehat{\gamma_4} + \sum_{j=1}^n \widehat{\alpha_j} B_{ji} + \widehat{\boldsymbol{\delta}}\boldsymbol{X_i} + \hat{\beta} DD_i \right)$

(5) $\hat{\pi}_{10} = \frac{1}{m}\sum_i^m exp\left(c + \widehat{\gamma_1} + \widehat{\gamma_4} + \sum_{j=1}^n \widehat{\alpha_j} B_{ji} + \widehat{\boldsymbol{\delta}}\boldsymbol{X_i} + \hat{\beta} DD_i \right)$

(6) $\hat{\pi}_{00} = \frac{1}{m}\sum_i^m exp\left(c + \widehat{\gamma_4} + \sum_{j=1}^n \widehat{\alpha_j} B_{ji} + \widehat{\boldsymbol{\delta}}\boldsymbol{X_i} + \hat{\beta} DD_i \right)$

with $\hat{\pi}_{11}$ notating downwind after coal closures, $\hat{\pi}_{01}$ upwind after closures, $\hat{\pi}_{10}$ downwind before closure, and $\hat{\pi}_{00}$ upwind before closure. These probabilities are conditional upon confounders B_{ji}, $\mathbf{X_i}$, DD_i and time since closure T_i.

The IRR compares the relative difference in health between two populations, for example representing the ratio of asthma cases for downwind residents relative to upwind residents. As such, $IRR_{wind|open} = \frac{\hat{\pi}_{10}}{\hat{\pi}_{00}}$ captures the incidence rate ratio of downwind residents relative to upwind residents when the power plant is operating; and $IRR_{wind|closed} = \frac{\hat{\pi}_{11}}{\hat{\pi}_{01}}$ captures the incidence rate ratio of downwind residents relative to upwind residents given a closed power plant.

Similarly, the $IRR_{closure|downwind} = \frac{\hat{\pi}_{11}}{\hat{\pi}_{10}}$ and $IRR_{closure|upwind} = \frac{\hat{\pi}_{01}}{\hat{\pi}_{00}}$ capture the incidence rate ratio from closure for downwind and upwind residents, respectively.

We will assess two measures of interactive risk: multiplicative and additive. First, multiplicative interaction occurs when $\frac{\hat{\pi}_{11} \cdot \hat{\pi}_{00}}{\hat{\pi}_{10} \cdot \hat{\pi}_{01}}$ >1, and can be thought of as the relative risk from closure for downwind residents W_i=1 versus the relative risk from closure for upwind resident W_i=0. This interaction term can also be interpreted as the relative risk for downwind resident from closure, C_i=1, versus the relative risk for downwind residents when the plant is open, C_i=0. Second, the excess risk due to interaction (RERI) is an additive interaction measure is calculated as $RERI = {(\hat{\pi}_{11} - \hat{\pi}_{10} - \hat{\pi}_{01} + \hat{\pi}_{00})}/{\hat{\pi}_{00}}$. In epidemiological literature, RERI is known as an interaction contrast ratio that is an additive interaction of risk (VanderWeele, 2009; VanderWeele & Knol, 2014). RERI gives the direction of the additive interaction, with RERI >0 indicating a closure would have a larger effect in the downwind group, and RERI <0 implies the closure would have a larger effect on the upwind group.

From the Poisson regression model, we will also assess the average risk difference (RD). Notated as $RD_{wind|open} = \hat{\pi}_{10} - \hat{\pi}_{00}$, this metric quantifies the average RD for the study population comparing downwind to upwind resident when the power plant is operating; $RD_{wind|closed} = \hat{\pi}_{11} - \hat{\pi}_{01}$ estimates the risk difference for downwind residents after the closure of a power plant; $RD_{closure|downwind} = \hat{\pi}_{11} - \hat{\pi}_{10}$ is the risk difference of the closure for downwind residents; and $RD_{closure|upwind} = \hat{\pi}_{01} - \hat{\pi}_{00}$ is the risk difference of the closure for upwind residents. Following Güdemann et al., (2022), we will assess the conditional average treatment effect on the treated (ATT) as measured as the risk difference $ATT = \hat{\pi}_{11} - \hat{\pi}_{01}$, given confounders B_{ji}, $\mathbf{X_i}$, and DD_i. Weighting can be used to account for the nature of the data as

repeated cross-sectional (Abadie, 2005). An alternative metric for ATT is the difference-in-difference estimates (DiD) measured as: $DiD = \hat{\pi}_{11} - \hat{\pi}_{01} - (\hat{\pi}_{10} - \hat{\pi}_{00})$, if significant differences between upwind and downwind cohorts exist (Athey & Imbens, 2006; Güdemann et al., 2022). The number needed to treat is simply NNT = 1/ATT, used in the benefit calculation.

Given that the outcome variables are binary, we will estimate this model using a Poisson pseudo maximum likelihood and examine the parallel trends assumption given the nonlinear model (Roth & Sant'Anna, 2023). A negative binomial model will be used as a robustness check if problems of overdispersion arise. The Poisson regression provides collapsible measures of risk measuring the incidence rate ratio rather than an odds ratio, and can be used to calculate, for example, the number of fewer asthma cases caused by a plant closure allowing for the benefit estimation (Colnet et al., 2023; Sagiv et al., 2005; Siddika et al., 2019; Spiegelman & VanderWeele, 2017). The estimated reduction in relative risk from the power plant closure will be used to estimate the healthcare cost savings from power plant closures using Medical Expenditure Panel Survey data for disease specific health costs. If convergence is not achieved, a logistic regression will be used to estimate relative risk and risk difference using sampling weights (Bieler et al., 2010).

To address confounding from health behaviors and socio-economic factors, we will include a vector of health related behaviors (e.g. smoking, drinking, physical activity) such that B_{ji} is behavior *j*. Additionally, $\mathbf{X1_{it}}$ represents a vector of appropriate controls for respiratory health outcomes, such as gender, age, BMI, education, marital status, occupation, access to healthcare, and region of the country, to address confounding and increase precision. Finally, DD_i represent a vector of disparity dimensions which captures socio-economic differences (e.g. income, race).

We propose to extend the model in equation (2) to examine health disparities by interacting the disparity dimension with the difference and difference component:

$$(7)\ \log(\pi_i) = c + \gamma_1 W_i + \gamma_2 C_i + \gamma_3 W_i \cdot C_i + \gamma_4 T_i + \sum_{j=1}^{n} \alpha_j B_{ji} + \boldsymbol{\delta X_i} + \beta DD_i + \theta_1 W_i \cdot DD_i + \theta_2 C_i \cdot DD_i + \theta_3 W_i \cdot C_i \cdot DD_i$$

The parameters $\gamma_1 + \theta_1$ represents the impact of wind direction on health outcomes for a given disparity dimension. For example, if the disparity dimension is race, then $\gamma_1 + \theta_1$ represents the wind effect on health for Black relative to White residents. Similarly, $\gamma_2 + \theta_2$ represents the closure effect on health for Black relative to White residents, and $\gamma_3 + \theta_3$ is interaction between residents upwind and downwind-before and after closure for Black residents relative to White. Similar analysis will be performed for other disparity dimensions such as income or gender with a clearly established reference group. Inference and assumptions of this triple difference model will be assessed following methods proposed by (Olden & Møen, 2022).

Performing a sensitivity analysis, we will test the robustness of the model to account for plant random effects, and clustered standard errors by power plant.

V. Benefits Estimation

The estimated reduction in risk from the power plant closure can then be used to estimate the cost savings from power plant closures using Medical Expenditure Panel Survey (MEPS) data. MEPS is administered by the Agency for Healthcare Research and Quality and is the only national source of data measuring how Americans pay for medical care. The household component summary tables provide expenditure estimates that are nationally representative. Table 3 shows the 2019 mean expenditures by race, and poverty status for persons with COPD, asthma, or other respiratory conditions. Expenditures trend upward for poorer households, yet near poor households (i.e. households with income between 100%-125% of the federal poverty line (FPL)) experience the highest average expenditure per person receiving care.

Table 3 Mean expenditure per person with care for COPD, asthma, and other respiratory conditions by income and race in 2019

	Mean Expenditure	95% Confidence Interval	
All persons	$2,153	$1888	$2419
Household Income			
Poor <=100% FPL	$2830	$2207	$3454
Near poor >100% FPL and <=125%FPL	$3347	$1089	$5606
Low income >125% FPL and <=200%FPL	$2049	$1412	$2686
Middle income >200% FPL and <=400%FPL	$2072	$1555	$2590
High income >400% FPL	$1844	$1536	$2151
Race			
White	$2231	$1884	$2578
Black	$2412	$1493	$3331
Hispanic	$1819	$1127	$2510

Income definitions defined by ARHQ based on Federal Poverty Line (FPL). Classification by race and ethnicity is based on self-reported family members.

MEPS data will be used to calculate the healthcare cost savings from reduced emissions based on changes in health outcomes derived from our statistical model. Specifically, the number needed to treat (NNT) along with the estimated population size will be used to derive both benefits from retired plants as well as cost of having active plants remain open. We use the following formula to calculate benefits:

(8) $Benefit_{Retired} = Mean\ Expenditure \times \sum_k {}^{N_{Closed}}/_{NNT_k}$

(9) $Cost_{Active} = Mean\ Expenditure \times \sum_k {}^{N_{Active}}/_{NNT_k}$

where N_i represents the population size of residents downwind of an active or closed power plant. For the k respiratory health outcomes including chronic bronchitis, COPD, lung cancer, emphysema, and asthma, we estimate the NNT the models presented in equation (2). Summing over the health outcomes provides us with the number of respiratory cases in the population

downwind of a closed powerplant, $\sum_k {}^{N_{Closed}}/_{NNT_k}$, or active power plant, $\sum_k {}^{N_{Active}}/_{NNT_k}$. Average expenditures are then used to calculate total benefit from coal retirements, as well as total cost keeping open active power plants.

This model will then be extended to examine disparities in benefits by evaluating the NNT from models presented in equation (7), which accounts for disparities in the average treated effect on the treated by income and race. Equations (8-9) will be used across these dimensions of disparity to compare the benefits from powerplant closures across race and income.

Further, we will quantify the trade-off with energy through the formulas:

(10) $$Benefit\ per\ MW_{Retired} = {}^{Benefit_{Retired}}\big/_{(S * NP_R)}$$

(11) $$Cost\ per\ MW_{Active} = {}^{Cost_{Active}}/_{(S * NP_A)}$$

where benefit per unit of capacity is a function of total benefit divided by the mean plant size, S, in units of MW, and the number of retired plants within our NHIS merged cohort, NP_R. Similarly, cost of keeping plants open per unit of capacity is a function of total cost divided by S and the number of active matched plants NP_A.

VI. Conclusion

This research protocol proposes to examine environmental health disparities and the subsequent disparities in the economic burden of disease across two distinct dimensions: income and race. This combination will allow strong inferences into the relationship between the environment, socio-economic factors, health outcomes, and healthcare costs. The study design takes advantage of restricted access data from the National Center for Health Statistics and applies econometric inference techniques to answer environmental health science and health economic questions.

Specifically, the research protocol estimates the health and healthcare impacts from a reduction in air pollution exposure from coal power plant closures across socio-economic subgroups. These infrastructure changes proxy for changes in environmental exposures, and thus can be used to study the health impacts of childhood and adult asthma, chronic bronchitis, emphysema, COPD, and lung cancer. Furthermore, the study design adds to the fields of environmental health economics, and health disparities research by taking advantage of locational and plant closure differences to implement quasi-experimental designs methods, and merging restricted access data in a manner such that maintains data stewardship best-practices, but allows for health disparities research.

The findings will contribute to the fields of environmental health economics, and energy economics by quantifying the cost of health disparities created by unequal exposure rates across racial and income subgroups. Thomson et al. (2018) find that coal mining counties, White and rural counites experienced some of the best improvements in air quality and specifically PM2.5 from the adoption of the Clean Air Act. Studying the Northeast's Regional Greenhouse Gas Initiative, Declet-Barreto and Rosenberg (2022) find that such policies focusing on total emissions reductions have largely benefitted non-environmental justice communities. Disparate exposure to energy-related air pollution increases health inequalities for low and minority populations, increases their cost of healthcare, and leads to a struggle between energy production, and public health and health disparities.

Limitations of the proposed methods are the sample size used for the assessment. While the NHIS dataset from 2009-2018 includes 955,782 survey respondents, the inclusion criteria of respondents within 30 miles of only one powerplant will reduce the available sample size. Additionally, not all respondents are asked questions regarding respiratory health further limiting

the sample size, and causing the sample size to vary between health outcomes. Sampling weights will be used to expand the analysis to represent other individuals in the U.S. As noted above, many of the geographic areas with only one coal power plant are west of the Mississippi River or in the southeastern U.S., limiting the generalizability of the interpretation. A final limitation of the study design is the influence of the inclusion and exclusion criteria on benefit and cost estimation. Our sample only includes individuals within 1 mile of coal power plant and excludes those near multiple power plants or near partially retired plants. This suggests the benefits of closure may be larger than this method estimates, and cumulative effects from multiple plants are unmeasured.

Two possible extensions of this proposal are possible. First, NCHS has a second survey collecting health information through the National Health and Nutrition Examination Survey (NHANES). These health interviews are more intensive than NHIS but can provide a more complete analysis of health-related behaviors and occupation. The survey combines interviews with physical examinations and biological specimen collection. The biennial survey provides a nationally representative sample from about 5,000 persons, and includes demographic, socioeconomic, dietary, and health-related questions. The main research trade-off between NHIS and NHANES is sample size for information on health-related behaviors to better account for lifestyle differences. Models based on NHANES can incorporate a greater level of detail for health-related behaviors, such as smoking, drug and alcohol use, and physical activity. The examination portion provides measurements on height, weight, blood pressure, and measures of lung function. Future research building on this model can also compare counterfactuals with residents near nuclear power plants, or wind and solar developments to measure risk differences based on energy source rather than wind direction.

VII. References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *The Review of Economic Studies*, *72*(1), 1-19.

Abel, T. D., & White, J. (2011). Skewed riskscapes and gentrified inequities: environmental exposure disparities in Seattle, Washington. *American journal of public health*, *101*(S1), S246-S254.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics*. Princeton university press.

Ash, M., & Boyce, J. K. (2018). Racial disparities in pollution exposure and employment at US industrial facilities. *Proceedings of the National Academy of Sciences*, *115*(42), 10636-10641.

Athey, S., & Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, *74*(2), 431-497.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, *119*(1), 249-275.

Bieler, G. S., Brown, G. G., Williams, R. L., & Brogan, D. J. (2010). Estimating model-adjusted risks, risk differences, and risk ratios from complex survey data. *American journal of epidemiology*, *171*(5), 618-623.

Blewett, L. A., Drew, J. A. R., King, M. L., Williams, K. C. W., Ponte, N. D., & Convey, P. (2022). *IPUMS Health Surveys: National Health Interview Survey* Version 7.2). https://doi.org/https://doi.org/10.18128/D070.V7.2

Brunello, G., Fort, M., Schneeweis, N., & Winter-Ebmer, R. (2016). The Causal Effect of Education on Health: What is the Role of Health Behaviors? *Health Economics*, *25*(3), 314-336. https://doi.org/10.1002/hec.3141

Burney, J. A. (2020). The downstream air pollution impacts of the transition from coal to natural gas in the United States. *Nature Sustainability*, *3*(2), 152-160.

Casey, J. A., Karasek, D., Ogburn, E. L., Goin, D. E., Dang, K., Braveman, P. A., & Morello-Frosch, R. (2018). Retirements of coal and oil power plants in California: association with reduced preterm birth among populations nearby. *American journal of epidemiology*, *187*(8), 1586-1594.

Clark, L. P., Millet, D. B., & Marshall, J. D. (2014). National patterns in environmental injustice and inequality: outdoor NO 2 air pollution in the United States. *PloS one*, *9*(4), e94431.

Coleman, N. C., Burnett, R. T., Higbee, J. D., Lefler, J. S., Merrill, R. M., Ezzati, M., Marshall, J. D., Kim, S. Y., Bechle, M., & Robinson, A. L. (2020). Cancer mortality risk, fine particulate air pollution, and smoking in a large, representative cohort of US adults. *Cancer Causes & Control: CCC*.

Colnet, B., Josse, J., Varoquaux, G., & Scornet, E. (2023). Risk ratio, odds ratio, risk difference... Which causal measure is easier to generalize? *arXiv preprint arXiv:2303.16008*.

Cushing, L. J., Li, S., Steiger, B. B., & Casey, J. A. (2023). Historical red-lining is associated with fossil fuel power plant siting and present-day inequalities in air pollutant emissions. *Nature Energy*, *8*(1), 52-61.

DeCicca, P., & Malak, N. (2020). When good fences aren't enough: The impact of neighboring air pollution on infant health. *Journal of Environmental Economics and Management*, *102*, 102324.

Declet-Barreto, J., & Rosenberg, A. A. (2022). Environmental justice and power plant emissions in the Regional Greenhouse Gas Initiative states. *Plos one*, *17*(7), e0271026.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, *109*(12), 4178-4219.

Dodd, K. E., & Mazurek, J. M. (2016). Asthma among employed adults, by industry and occupation—21 states, 2013. *Morbidity and Mortality Weekly Report*, *65*(47), 1325-1331.

Fraenkel, R., Zivin, J. S. G., & Krumholz, S. D. (2022). *The Coal Transition and Its Implications for Health and Housing Values*.

Gallagher, C. L., & Holloway, T. (2022). US decarbonization impacts on air quality and environmental justice. *Environmental Research Letters*, *17*(11), 114018.

Gertler, P., Premand, P., & Rawlings, L. (2011). WHO-Impact evaluation in practice, Chapter 6. *Geneva: World Health Organization*.

Goforth, T., & Nock, D. (2022). Air pollution disparities and equality assessments of US national decarbonization strategies. *Nature Communications*, *13*(1), 7488.

Güdemann, L., Dennis, J. M., McGovern, A. P., Rodgers, L. R., Shields, B. M., & Bowden, J. (2022). Triangulating Instrumental Variable, confounder adjustment and Difference-in-Difference methods for comparative effectiveness research in observational data. *arXiv preprint arXiv:2202.09164*.

Gupta, A., & Spears, D. (2017). Health externalities of India's expansion of coal plants: Evidence from a national panel of 40,000 households. *Journal of Environmental Economics and Management*, *86*, 262-276.

Hajat, A., Hsia, C., & O'Neill, M. S. (2015). Socioeconomic disparities and air pollution exposure: a global review. *Current environmental health reports*, *2*, 440-450.

Henneman, L. R., Rasel, M. M., Choirat, C., Anenberg, S. C., & Zigler, C. (2023). Inequitable exposures to US coal power plant–related PM 2.5: 22 years and counting. *Environmental Health Perspectives*, *131*(3), 037005.

Hernandez-Cortes, D., Meng, K. C., & Weber, P. E. (2022). *Decomposing Trends in US Air Pollution Disparities from Electricity*.

Jbaily, A., Zhou, X., Liu, J., Lee, T.-H., Kamareddine, L., Verguet, S., & Dominici, F. (2022). Air pollution exposure disparities across US population and income groups. *Nature*, *601*(7892), 228-233.

Jha, A., Matthews, P. H., & Muller, N. Z. (2019). Does environmental policy affect income inequality? Evidence from the Clean Air Act. AEA Papers and Proceedings,

Jia, X., & Luo, X. (2023). Residents' Health Effect of Environmental Regulations in Coal-Dependent Industries: Empirical Evidence from China's Cement Industry. *Sustainability*, *15*(3), 2512.

Johnson, S., & Chau, K. (2019). *More U.S. coal-fired power plants are decommissioning as retirements continue*. Website

Keil, A. P., Buckley, J. P., & Kalkbrenner, A. E. (2021). Bayesian g-computation for estimating impacts of interventions on exposure mixtures: demonstration with metals from coal-fired power plants and birth weight. *American journal of epidemiology*, *190*(12), 2647-2657.

Komisarow, S., & Pakhtigian, E. L. (2022). Are power plant closures a breath of fresh air? Local air quality and school absences. *Journal of Environmental Economics and Management*, *112*, 102569.

Lane, H. M., Morello-Frosch, R., Marshall, J. D., & Apte, J. S. (2022). Historical redlining is associated with present-day air pollution disparities in US cities. *Environmental science & technology letters*, *9*(4), 345-350.

Lee, D. J., Fleming, L. E., LeBlanc, W. G., Arheart, K. L., Chung-Bridges, K., Christ, S. L., Caban, A. J., & Pitman, T. (2006). Occupation and lung cancer mortality in a nationally representative US Cohort: The National Health Interview Survey (NHIS). *Journal of occupational and environmental medicine*, 823-832.

Luechinger, S. (2014). Air pollution and infant mortality: a natural experiment from power plant desulfurization. *Journal of health economics*, *37*, 219-231.

Mejía-Duwan, J., Hino, M., & Mach, K. J. (2023). Emissions redistribution and environmental justice implications of California's clean vehicle rebate project. PLOS Climate, 2(5), e0000183.

Miranda, M. L., Edwards, S. E., Keating, M. H., & Paul, C. J. (2011). Making the environmental justice grade: the relative burden of air pollution exposure in the United States. *International journal of environmental research and public health*, *8*(6), 1755-1771.

Nachman, K. E., & Parker, J. D. (2012). Exposures to fine particulate air pollution and respiratory outcomes in adults using two national datasets: a cross-sectional study. *Environmental Health*, *11*(1), 25.

NRECA. (2019). *2019 Long-Term Reliability Assessment* (LTRA, Issue. https://www.nerc.com/

Olden, A., & Møen, J. (2022). The triple difference estimator. *The Econometrics Journal*, *25*(3), 531-553.

Ou, C.-Q., Hedley, A. J., Chung, R. Y., Thach, T.-Q., Chau, Y.-K., Chan, K.-P., Yang, L., Ho, S.-Y., Wong, C.-M., & Lam, T.-H. (2008). Socioeconomic disparities in air pollution-associated mortality. *Environmental Research*, *107*(2), 237-244.

Parker, J. D., Akinbami, L. J., & Woodruff, T. J. (2009). Air pollution and childhood respiratory allergies in the United States. *Environmental health perspectives*, *117*(1), 140-147.

Parker, J. D., Kravets, N., & Vaidyanathan, A. (2018). Particulate matter air pollution exposure and heart disease mortality risks by race and ethnicity in the United States: 1997 to 2009 National Health Interview Survey with mortality follow-up through 2011. *Circulation*, *137*(16), 1688-1697.

Payne-Sturges, D., & Gee, G. C. (2006). National environmental health measures for minority and low-income populations: tracking social disparities in environmental health. *Environmental Research*, *102*(2), 154-171.

Pope, C. A., Ezzati, M., Cannon, J. B., Allen, R. T., Jerrett, M., & Burnett, R. T. (2018). Mortality risk and PM 2.5 air pollution in the USA: an analysis of a national prospective cohort. *Air Quality, Atmosphere & Health*, *11*(3), 245-252.

Pope III, C. A., Lefler, J. S., Ezzati, M., Higbee, J. D., Marshall, J. D., Kim, S.-Y., Bechle, M., Gilliat, K. S., Vernon, S. E., & Robinson, A. L. (2019). Mortality risk and fine particulate air pollution in a large, representative cohort of US adults. *Environmental health perspectives*, *127*(7), 077007.

Qiu, M., Zigler, C. M., & Selin, N. E. (2022). Impacts of wind power on air quality, premature mortality, and exposure disparities in the United States. *Science Advances*, *8*(48), eabn8762.

Richmond-Bryant, J., Mikati, I., Benson, A. F., Luben, T. J., & Sacks, J. D. (2020). Disparities in distribution of particulate matter emissions from US coal-fired power plants by race and poverty status after accounting for reductions in operations between 2015 and 2017. *American Journal of Public Health*, *110*(5), 655-661.

Robinson, C. F., Sullivan, P. A., Li, J., & Walker, J. T. (2011). Occupational lung cancer in US women, 1984–1998. *American journal of industrial medicine*, *54*(2), 102-117.

Roth, J., & Sant'Anna, P. H. (2023). When is parallel trends sensitive to functional form? *Econometrica*, *91*(2), 737-747.

Sagiv, S. K., Mendola, P., Loomis, D., Herring, A. H., Neas, L. M., Savitz, D. A., & Poole, C. (2005). A time series analysis of air pollution and preterm birth in Pennsylvania, 1997–2001. *Environmental Health Perspectives*, *113*(5), 602-606.

Senthilselvan, A., Coonghe, W., & Beach, J. (2020). Respiratory health, occupation and the healthy worker effect. *Occupational Medicine*, *70*(3), 191-199.

Siddika, N., Rantala, A. K., Antikainen, H., Balogun, H., Amegah, A. K., Ryti, N. R., Kukkonen, J., Sofiev, M., Jaakkola, M. S., & Jaakkola, J. J. (2019). Synergistic effects of prenatal exposure to fine particulate matter (PM2. 5) and ozone (O3) on the risk of preterm birth: A population-based cohort study. *Environmental research*, *176*, 108549.

Spiegelman, D., & VanderWeele, T. J. (2017). Evaluating public health interventions: 6. Modeling ratios or differences? Let the data tell us. *American Journal of Public Health*, *107*(7), 1087-1091.

Spiller, E., Proville, J., Roy, A., & Muller, N. Z. (2021). Mortality risk from PM 2.5: a comparison of modeling approaches to identify disparities across racial/ethnic groups in policy outcomes. *Environmental Health Perspectives*, *129*(12), 127004.

Tessum, C. W., Apte, J. S., Goodkind, A. L., Muller, N. Z., Mullins, K. A., Paolella, D. A., Polasky, S., Springer, N. P., Thakrar, S. K., Marshall, J. D., & Hill, J. D. (2019). Inequity in consumption of goods and services adds to racial-ethnic disparities in air pollution exposure. *Proc Natl Acad Sci U S A*, *116*(13), 6001-6006. https://doi.org/10.1073/pnas.1818859116

Thind, M. P., Tessum, C. W., Azevedo, I. L., & Marshall, J. D. (2019). Fine particulate air pollution from electricity generation in the US: health impacts by race, income, and geography. *Environmental science & technology*, *53*(23), 14010-14019.

Thomson, V. E., Huelsman, K., & Ong, D. (2018). Coal-fired power plant regulatory rollback in the United States: Implications for local and regional public health. *Energy Policy*, *123*, 558-568.

VanderWeele, T. J. (2009). On the distinction between interaction and effect modification. *Epidemiology*, 863-871.

VanderWeele, T. J., & Knol, M. J. (2014). A tutorial on interaction. *Epidemiologic methods*, *3*(1), 33-72.

Wilkie, A. A., Richardson, D. B., Luben, T. J., Serre, M. L., Woods, C. G., & Daniels, J. L. (2023). Sulfur dioxide reduction at coal-fired power plants in North Carolina and associations with preterm birth among surrounding residents. *Environmental Epidemiology*, *7*(2), e241.

Yang, M., & Chou, S.-Y. (2018). The impact of environmental regulation on fetal health: Evidence from the shutdown of a coal-fired power plant located upwind of New Jersey. *Journal of Environmental Economics and Management*, *90*, 269-293.

Yi, D., & Sung, J.-H. (2021). An environmentally related policy impact analysis considering wind effect: evidence from suspending old coal-fired generators in South Korea. *Applied Economics Letters*, 1-9.

Appendix A – Population Estimates of Health Outcome by Disparity Dimensions Poverty and Race

Table A1 provides the population estimates for health outcome for White, Black, and other minority respondents, and again by poverty status. Health outcomes are worse for those below the federal poverty line across all metrics of respiratory health. Black respondents have worse rates for general health, chronic bronchitis, lung cancer, having a cold, and asthma relative to White respondents.

Table A1: Population Estimates of Health Outcome by Poverty and Race.

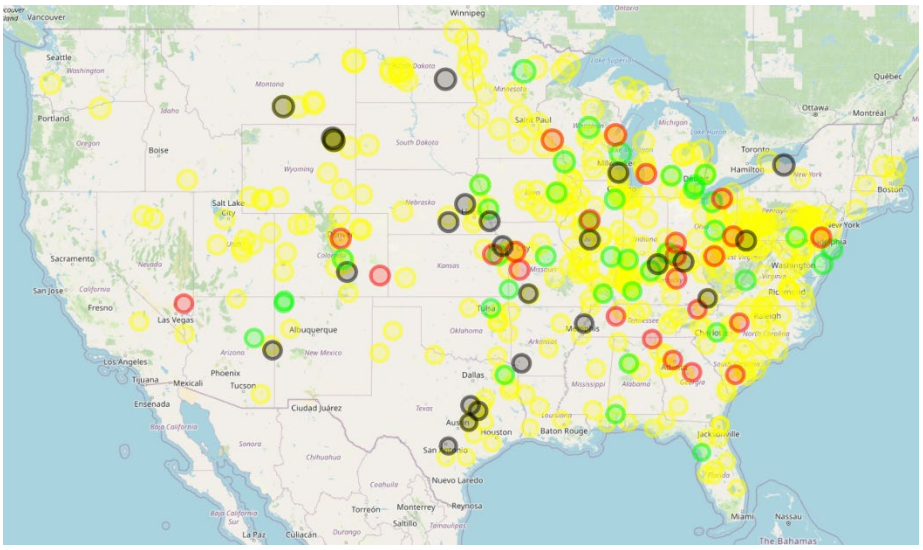
	White	Black	Other Minority	P-value	Above FPL	Below FPL	P-value
General Health Status – Fair/Poor	9.63%	13.38%	8.72%	<0.01	8.43%	18.27%	<0.01
Chronic bronchitis past year	4.37%	4.51%	2.41%	<0.01	3.83%	6.74%	<0.01
COPD	4.08%	2.7%	1.59%	<0.01	3.33%	5.9%	<0.01
Emphysema	2.15%	1.24%	0.89%	<0.01	1.68%	3.27%	<0.01
Lung Cancer	2.89%	4.66%	3.97%	<0.01	2.78%	4.98%	<0.01
Had a cold in the past 2 weeks	12.21%	12.77%	10.76%	<0.01	11.82%	15.26%	<0.01
Ever been told had asthma	12.71%	16.38%	11.16%	<0.01	12.49%	17.19%	<0.01
Asthma attack in last year	31.14%	33.44%	31.75%	<0.01	29.97%	38.14%	<0.01
ER visit for asthma in last year	8.08%	17.41%	10.79%	<0.01	8.06%	16.41%	<0.01

Source: NHIS 2009-2018 collected from IPUMS (Blewett et al., 2022).

Appendix B – Power plants excluded from analysis

Excluded plants include those with partial retirements, n=38 (i.e. some generators remained operational), or with generators' retirement years that varied more than 0.5 standard deviations, n=24. Occasionally, new generators are added to existing power plants, and we exclude these power plants by removing generators starting operations after 2008 and with generators' first operating years that varied more than 0.5 standard deviations, to ensure a discrete starting date. Figure B1 shows the 30 mile radius of included plants (yellow) and excluded plants (red, green, black) based on exclusion criteria. Most excluded plants are exclusive of other plants or their radius falls within two or more plants, which would have further excluded residents in those areas from the analysis.

Figure B1: Excluded Coal Locations



Yellow circles indicate the 30 mile radius of plants included in the analysis. Black circles indicate plants that added generation during the study period. Green circles indicate plants with only a partial retirement. Red circle plant locations indicate plants with a standard deviation in its generators' retirement years greater than 0.5.