# The Airborne Toxic Event: The Effects of Socioeconomic Characteristics on Ambient Air Pollution and the Decision to Over Pollute

Harrison Cho

The University of North Carolina at Chapel Hill

**Acknowledgements:**

Where to begin.

I would not have created this without the assistance of many brilliant folks. They say it takes a village; it took the joint brain power of several Masters degrees, a handful of PhDs, and an untold number of Undergraduate degrees to get to this point. First, I would like to thank my esteemed advisor, Dr. Andrew Yates, for his guidance, persistence, and cogency. Without your direction, I would have floundered until the very end. Thank you for having confidence in my work. Second, I would like to thank my cohort leader and trusted confidante, Dr. Jane Fruehwirth. Your compassion, empathy, and wisdom know no bounds. Third, I would like to thank Matthew Jansen and Philip McDaniel, the wizards, the dynamic duo, the dream team of the Davis second floor Research Hub. Gentlemen, your skills made a hypothesis actually testable.

I would be remiss to not acknowledge the immense, loving support of my family and friends. I am the product of your hard work and patience. It took many years to get to this moment, but now this thesis stands as a testament to your collective efforts. Thank you for giving a rats, as they say.

**Abstract**

As policy solutions for climate change are adopted at an increasing rate, it is integral to prioritize environmental equity strategies, those that allow socioeconomic minorities access to proper resources and decision making power. One argument for this change follows from the inequitable distribution of pollution throughout the United States, a topic whose history is steeped in complex networks of historical discrimination. In order to account for these past trends, this paper utilizes a dynamic panel model to test whether socioeconomic characteristics affect changes in air pollution within the United States between 2010 and 2017. Utilizing stationary electric utilities throughout the contiguous United States, I tabulate changes in three air pollutants with weighted proportions of socioeconomic characteristics in nearby census tract. As a significant and prevalent source of pollution throughout the United States, power plants are an important and generalizable determinant of ambient air quality. To mitigate heterogeneity resulting from time sensitive and location sensitive trends, I incorporate spatially and temporally lagged variables for pollutant measures and socioeconomic characteristics, respectively. Utilizing GMM to measure estimates, I find that race, wealth, education, and age are significant determinants of air quality. Although subsequent improvements can be incorporated for the model's internal validity, historical inequity is still an important determinant for pollution distribution in the present.

## Introduction

The short and long term impacts of over pollution are multiplicative. Specifically, individuals exposed to the negative externalities of over pollution must bear increasing financial and personal uncertainty driven by negative health outcomes (Kim et al., 2007). This uncertainty is a burgeoning sign of disparity that has huge implications over a variety of policy issues including, but not limited to; healthcare accessibility, infrastructure degradation, and housing discrimination. These three issues have been accentuated by industrialization in America, a gradual movement that has altered the sustainability of natural and man-made environments (Judy, 2018). Gradually, we are coming to grips with the scale and magnitude of industrialization's negative byproducts, seeking to minimize deleterious effects before long term trends become irreversible. To adequately adjust for pollution variability over time, models predicting the effects of pollution have incorporated discrimination as an explanatory variable, illustrating disparate pollution patterns amongst numerous underserved communities. As such, this study seeks to develop a coherent methodology addressing environmental equity for power plant pollution at a national level. Measuring the effects of air pollution, I utilize a methodology that analyzes the relationship between pollution levels and demographic characteristics.

The effects of ambient air pollution, or the combination of air pollutants arising from a mixture of emission sources, are well documented (Curtis et al., 2006). As empirical strategies measuring ambient air pollution have evolved, increased awareness surrounding sensitivity to ambient air pollution has motivated regulators to adopt policy stressing the allocative efficiency of pollution levels (Judy, 2018). Common air pollutants include a mixture of nitrogen oxide ($NO_x$), sulfur dioxide ($SO_2$), fine particulate matter ($PM_{2.5}, PM_{10}$), and Ozone ( $O_3$). Although it is difficult to quantify the isolated effects of these common pollutants, the literature shows significant correlations between ambient air quality, morbidity rates, and mortality rates (Kim et al., 2007). The aforementioned pollutants are respiratory irritants and vascoconstrictors, as a result, respiratory and cardiovascular diseases are highly associated with poor ambient air quality (Curtis et al., 2006). Estimates associated with morbidity and mortality tend to vary, however, higher pollution concentrations tend to be a driving force behind diminished health outcomes and higher medical costs; higher rates of hospitalization and chronic illness tend are associated with low ambient air quality (Wu et al., 2020). In addition, Olstrup et al's(2018) panel study show that approximately 20% of an increase in life expectancy was explained by changes in concentrations of nitrogen oxide.

Without proper accountability measures and resources, dangerous pollution levels become enmeshed within a series of already compounding problems that affect America's disadvantaged populations. This is to say, those without previous access to the appropriate and sufficient resources are left with no recourse. Thus, the distribution of pollution and human exposure has revealed a flawed, inequitable system that weighs heavy on seemingly marginalized sub-sections of the US population. Publicized failings pertaining to Flint, Michigan's water crisis and even North Carolina's hog farms illustrate extreme disparities facilitated by environmental factors. Beyond these influential examples, pollution's more subtle effects are masked by a series of heterogeneous regulatory and production environments that fail to address systemic issues decisively. Ultimately, policy makers can address these grievances by adopting attitudes promoting environmental equity, attitudes about environmental externalities ensuring that "the spatial distribution of environmental risk is indeed equitable among different racial and socioeconomic groups (Mennis and Jordan, 2005)." Prioritizing equitable regulations that control for past bias, the US can take steps to eliminate structural differences between pollution outcomes.

Achieving environmental equity requires a sophisticated understanding of pollution's many drivers. The factors governing the decision to over pollute are distorted by variations within population, regulatory, and institutional factors. Additionally, the highly localized distribution of pollution exemplifies a need for aggregating population characteristics and measuring pollution abatement. Without adopting a unifying framework to measure pollution's effects on a broad, comprehensive scale, those who are overly affected by pollution will continue to bear an undue burden.

Utilizing panel methods, this paper models and evaluates the pollution levels amongst socioeconomically marginalized groups. Specifically, this paper estimates whether power plants over pollute in areas with higher concentrations of disadvantaged socioeconomic groups. The effects of pollution on these groups have been somewhat ambiguous over time; shifting methodologies and strategies have clouded the extent of pollution's measurable effects amongst underserved socioeconomic groups. This is compounded by statistical bias arising from multiple sources including, but not limited to: proximity pollution epicenter due to reduced cost, spatial autocorrelation of demographic characteristics, and esoteric characteristics amongst individual geographic units.. However, this paper seeks to eliminate bias traditionally seen in environmental equity research using real pollution data and an instrumental variable approach. Utilizing emissions data from Holland et al's (2018) study on power plant pollution and socioeconomic demographic data from the American Community Survey, this study constructs a fixed effects model measuring pollution as a function of pollution abatement over time and demographic characteristics. Applying geographically weighted proportions adopted from Brooks and Sethi

(1997), these groups are aggregated and spatially weighted to allot characteristics closer to power plants greater weight. I employ theoretical and empirical models utilized in Baryshnikova (2010) to account for gradual abatement in pollution. In addition, I utilize the General Method of Moments framework to account for correlation between pollution in lagged and present periods. To control for endogeneity resulting from the proximity of demographic groups to power plants, spatially lagged instruments are created. Although these variables are incorporated to eliminate correlation between demographic characteristics and the error term, there is an intuitive explanation for their inclusion. Overall, these instruments can help account for unexplained variation in pollution due to the autocorrelation of demographic characteristics. As distance increases, pollution characteristics should remain relatively similar as pollution concentration decreases. Given this relationship, we are able to better approximate change while controlling for simultaneity that may affect the internal validity of the model. Beyond these parameters, the empirical model uses a variety of controls to estimate if power plants choose to over pollute in areas with higher proportions of socioeconomically disadvantaged groups.

In the next section, I summarize the body of econometric and theoretical literature that governs pollution mapping. Then, I cover a theoretical model, its parameters, its implications, and overall construction. Next, I discuss the data used in this study and variables of interest. Subsequently, I discuss my estimation strategy and the effects of demographic characteristics on pollution; I cover the formation of weighted demographic variables, the construction of instrumental variables, and the random trend model that allows for individual time trends. Finally, I discuss the results of my model and its implications. I provide commentary on overall improvements that can be studied in the future, and examine the implications of standard errors within this autocorrelative framework.

**Literature Review**

Industrial polluters must consider a multiplicity of factors to properly tabulate the optimal level of pollution. The physical distribution of pollution is affected by a variety of factors, including, but not limited to: the point source's location, atmospheric conditions, and topographic positioning (Mennis and Jordan, 2005). Although these factors have numerous interactions with one another, increases in distance are limit pollution concentration the greatest. Pollution's effects are more prominently experienced by those who are more proximate to its emission source, thus many environmental equity studies have thoroughly sampled populations living near emissions sources. However, these samples bias methodologies measuring the effects of pollution on any measurable characteristic. Given lower living costs and complicated historical trends of housing discrimination, many disadvantaged groups live near powerplants. Consider the first law of geography, "everything is related to everything else, but near things are more related than distant things (Chakraborty, 2011)." Overall, proximate groups share social characteristics, and are likely to lead to sampling bias, research methodologies have had difficulty eliminating endogeneity resulting from proximity. Overall, this autocorrelation poses a significant challenge in deducing causation from these interrelated, proximate populations. This study seeks to reduce autocorrelation through an instrumental variable approach. Utilizing spatial econometric techniques, I attempt to estimate procedures that allow us to draw broad causal inference. Adopting the correct procedure allows us to account for dispersion that is not seen in case studies and or regional correlative models. By looking at a broad swath of socioeconomic characteristics, we can model the relationship between pollution and demographic characteristics with increased accuracy.

Previous research on environmental equity suggests that spatial and socioeconomic characteristics also affect exposure to pollution (Liu et al., 2018; Li et al., 2019). Studies testing the effects of air pollution on mortality find that those with access to fewer resources, like income-based or housing-based resources, are more likely to suffer from higher rates of cardiovascular and respiratory morbidity (Wong et al., 2008). The population density of a geographic region can also lead to variability in pollution the effects of pollution. A paper by Holland et al. indicated that the United States' populated northeastern and southern regions tend to have higher rates of pollution. This is compounded by heterogeneous regulations and regulatory bodies governing pollution production throughout the United States. Without consistent pollution mitigation procedures across states, population demographics tend to be affected by over pollution at varying rates. Beyond geographic considerations, socioeconomic demographics tend to have strong correlations with pollution rates. A study by Li et al. found that excess emissions, emissions that are notoriously underregulated, were positively correlated with racial and ethnic minority populations. Contrary to this study, Gray and Shadbegian (2004) note that toxic emissions do not disproportionately affect these groups. Across the body of literature, a variety of studies with numerous empirical strategies tend to produce different, and sometimes conflicting results. Serial autocorrelation resulting from extreme correlation between both sources of pollution and demographic characteristics introduce endogeneity to models. Often, models measuring the relationship between pollution and demographic characteristics tend to underestimate the overall effect; this can be attributed to either a lack of proper weighting characteristics. Empirical studies have different strategies in accounting for spatial autocorrelation. Because this is one of the primary threats to validity of a model drawing causal inference, I now turn a model that measures pollution production while dealing with bias via an instrumental variable strategy.

Baryshnikova (2010) mitigates demographic-based endogeneity via a two-pronged approach. First, a fixed effects model realistically models pollution capital and production, controlling for the unique conditions of paper plants and pollution abatement targets. Second, an instrumental variable approach mitigates endogeneity resulting from the proximity of socioeconomic minorities. Baryshnikova adopts a spatially lagged instrumental variable approach developed by Gray and Shadbegian (2004). Creating temporally lagged variables is not feasible given the recency of sufficient data. Time series data describing demographic variables does not extend into periods of explicit housing segregation or zoning segregation, thus other methodologies to control for demographic-based endogeneity must be utilized. In order to control for endogeneity brought on by autocorrelation, spatially lagged instruments are created to control for demographic variables closer to the power plant. After pinpointing a paper plant's location, demographic characteristics are aggregated at 50 and 100 mile radii surrounding these plants. Intuitively, the lagged demographic instruments are correlated with populations closer to the plant, given the first law of geography. In addition, the aggregated groups at this distance have a minimal effect on a plant's pollution decision. The instrument eliminates endogeneity as pollution decreases with distance. Overall, Baryshnikova's study found that elevated pollution levels were more common amongst populations that had more children, and populations that had obtained less than a high school education. This paper establishes an important framework to account for changes in pollution, however the sample of paper plants is not robust or a generalizable source of pollution. As such, I institute Baryshnikova's framework within this study to examine a larger, more robust source of air pollution throughout the United States.

Utilizing electric utilities allows us to expand the scope of the model and institute more controls, given the distribution of power plants across the contiguous United States/

After eliminating endogeneity within demographic characteristics, additional criteria can be used to refine our model's empirical approach to location-based segregation. I broaden the characteristics considered in the econometric literature to provide a more intuitive portrait of environmental racism. First, I attempt to contextualize relationships between occupation and pollution exposure. Park and Kwan (2017) suggest a relationship between environmental outcomes and workplace segregation. Intuitively, these social characteristics encompassed in occupational variables further elucidate shared social realities amongst specific demographics. Thus, an occupation's field can be used to refine our location-based approach. In addition, this variable may allow us to develop intuitions concerning where the occupation is located, how a worker travels to that job, and other characteristics that account for where and when free time. Although location-based segregation features prominently in equity research, these additional considerations could be incorporated into the model developed by Baryshnikova. These additional controls attempt to flesh out the traditional spatial approach adopted within environmental equity studies. I expand the scope beyond simple housing discrimination to study how pollution exposure can affect access to other prominent lines of work. In sum, controlling for numerous factors that dictate the length and severity of pollution overexposure can further inform the theoretical and empirical models that this study develops.

Beyond expanding the limited scope of environmental racism, this paper also utilizes a framework that reduces error and increases generalizability. Overall, many econometric models estimating the effects of pollution on equity are biased by variability within pollution's spread and autocorrelation amongst demographic characteristics. Often, we do not have sufficient time series data to generate models that can completely control for these sources of bias. Studies must rely on utilizing spatially lagged and temporally lagged to account for this. Thus, we can improve the Generalized Method of Moments model utilized in Baryshnikova's study. Generalizability of Baryshnikova's study is lacking, and her study's results can only extend over paper plants. This study improves upon Baryshnikova's methodology utilizing several methods. First, I extend analysis over a more diffuse, and frequently tabulated source of pollution: power plants. Because air pollution has consistently diminished over the last decades, it is important to track spatial distribution using data that is temporally relevant. The frameworks governing data collection for pollution have changed, thus monitoring pollution for air pollution within the last decade provides greater precision for the dependent variable that will be measured within this study. In addition, this study uses data over a consistent time series, as opposed to interpolated data using a cubic spline method. Baryshnikova utilizes data from the decennial census, interpolating racial demographics between 1985 and 1997; no contiguous time series data is utilized. Overall, these factors expand the controls and estimations that are implemented in order to refine the environmental justice framework adopted by this study.

**Methodology**

This section will explore the theoretical underpinnings that govern the production and abatement of pollution amongst power plants in the United States. To measure and account for potential sources of bias, I outline the socioeconomic data needed to measure pollution abatement.

## I. Theoretical Model

I utilize Barynshnikova's (2010) theoretical power plant model to address pollution abatement at the census tract level. Capturing abatement over time allows us to capture variability at the plant level though an autoregressive variable. The underpinnings described within this section allow us to set up this variable such that we can capture heterogeneity resulting from location-specific trends. Although there are many decisions that dictate pollution production, we condense the decision process down to several parameters. First, the model holds that pollution abatement must be measured as a gradual process. Despite the creation of pollution abatement targets for any given amount of human capital and physical capital at the plant level, only a proportion of pollution will be credibly abated. Because only a fraction of this abatement target will be attained, measuring variability in abatement's success becomes interlinked with overall production for any given year. Second, the model holds a plant's physical capital fixed over time. While measuring physical capital could be accomplished utilizing data on capital expenditures, depreciation, total assets, etc., this model does not account for malleability in levels of capital. If we were to consider variability in capital stock, the model would not adequately account for unforeseen costs. Usually, the installation of capital requires both significant investment in time and money; mastering advanced operational capabilities to competently control technology and installation costs are often costly prospects that cannot be measured in the short run. Advanced on the job training procedures, lengthy installation periods, and production learning curves are additional subtle variable costs imposed by updating technology. This is not to understate the importance of technological change in the pollution reduction process; because appropriate controls accounting for financial data are not incorporated in this model, we only consider the short term implications of pollution production.

Finally, a single agent, the plant manager, maintains authority in determining the optimal level of pollution production. As a rational agent, the plant manager's main disincentive arises from regulatory agencies that may hinder business if the manager chooses to over pollute. Regulatory agencies impose hefty fines and, if extreme enough, issue injunctions for sustained cases of over pollution. Without a regulatory presence, plant managers would always choose to maximize their profits, a function of pollution production. As a rational agent, the plant manager knows to produce until the marginal costs of production are equivalent to their marginal benefits resultant of a specific abatement target. Because regulatory agencies seek to reduce pollution over time, this model measures the level of pollution that is reduced from year to year after considering the "optimal" level of pollution to produce. In this case, regulatory bodies are critical in setting the constraints that govern these theoretical production quantities. Thus, for power plant $i$ at time $t$, the optimal level of pollution, $P_{it}^*$ , is determined using the following equation:

$$P_{it}^* = \beta_1 T_i + \beta_2 R_i + \beta_3 O_{it} + \beta_4 U_{it} \qquad \textbf{(1)}$$

To clarify notation, this model's variables can categorized within two sets: the first encompassing plant-level technology, [$T_i$ ], and the second encompassing plant-level regulation, [$R_i$, $O_{it}$, $U_{it}$]. It should be noted that quantities within this equation are all theoretical. Per the parameters previously mentioned, $T_i$ is held fixed at the plant level. Regulation, however, contains both time variant and invariant effects. Intuitively, power plants have different attributes predetermined by the populations they serve. For example, a rural power plant may not be subject to the same restrictions governing a power plant that serves a major metropolitan area. Because power plants must meet power demand for the area it serves, significant variability arises. In this model's case, variability in power demand governs the rigidity of regulations on a plant-by-plant basis. Besides power demand, other social or institutional factors determine the regulations governing power demand. Among other things, Baryshnikova notes the importance of a power plant's union representation or its district's political visibility. These theoretical quantities are fixed and often determined by geographic qualities. Overall geographic and spatial population characteristics organize complex social networks that determine the distribution of resources for power supply. These institutional fixed factor are encompassed under coefficient $R_i$.

Although one facet of regulation exists as a fixed quantity, other facets are malleable over time and place; as regulatory standards shift, pollution production should respond accordingly. Tangible temporal shifts in regulation are described in coefficient $O_{it}$. Changes in state or federal mandates would most likely cause shocks to this change. Overall, "visible" changes in regulation are borne from the complex interactions between government and business. The model acknowledges changes in theoretical quantities of regulation that could be measured via a variety of social indices, regulation considerations, or government responses. All regulation cannot be modeled, simply because power plant operations are specialized. Geography, time, and esoteric social interactions within powerplants lead to  unforeseen variability in pollution production that cannot be captured under observable regulations; $U_{it}$ attempts to account for this unexplained amount. Intuitively, not all shifts in pollution production can be explained via regulation; given larger contemporary trends determined by demographic characteristics, the last term controls for remaining variability in pollution production. Introducing socioeconomic demographics into our model, after accounting for time and distance, indicates that discrimination is a viable explanation for over pollution. Measuring changes in pollution, holding all demographics fixed, would allow us to indicate disparities as a result of a specific demographic characteristic.

## II. **Data**

To calculate power plant pollution across the United States, this study utilizes data from Holland et al's county level study of pollution damages. Sulfur Dioxide ($SO_2$), Nitrogen Oxide ($NOx$), and PM2.5 levels were recorded for 1571 individual power plants. With assistance from the US Energy Information Administration (EIA), hourly values were recorded and then averaged accordingly to create pollution levels for each year. To be included in the dataset, power plants had to maintain consecutive observations between 2010 and 2017. As such, power plants that had data for each of these three pollution measures between 2010 and 2017 were included in the study. As common pollutants, the three dependent variables affect the US population regularly. These pollutants often interact with atmospheric particulates, causing a variety of chemical reactions that can intensify the penetration of volatile compounds within the human body. High concentrations of these three common pollutants, in the long run, induce respiratory ailments that continue to increase negative externalities. Although Holland et al's study found that reductions in pollution have occurred over time, this paper specifically examines the distribution of pollution across differing socioeconomic classes. In order to deduce any effects on socioeconomically marginalized groups, I turn to a representative dataset sampled at the census tract level.

Socioeconomic data is pulled from the Bureau of Labor Statistic's American Community Survey (ACS), an annual tabulation of population demographics across the United States. Specifically, I use the ACS' 5-year tabulations to minimize error that may arise from missing values. The ACS' 5-year tabulation samples 0.2% of the US population per year. Over the course of five years, annual samples are combined to account for 1% of a representative US population. For example, work for the 2010 ACS begins in 2006 and extends over a five year period.

Although this is a time-intensive process, these rolling estimates provide a comprehensive picture. In total, lower rates of nonresponse bias and missing values provide a more complete measure of various socioeconomic characteristics weighted by the distance, between a specified radius and power plant in the numerator, and the total population of a census tract in the denominator.

$$D_i = \frac{\Sigma_k D_{ik} \left(\frac{s - d_{ik}}{s}\right)}{\Sigma_k P_{ik} \left(\frac{s - d_{ik}}{s}\right)}$$

This formula will be formally discussed in the construction of the empirical model. The ACS includes characteristics pertaining to: racial demographics, measures of income inequity, educational attainment, and housing characteristics. Table 1 is an abridged table of variables critical to the construction of my empirical model.

**Table One – Socioeconomic Variables and Definitions**

| Variable | Definition |
|---|---|
| AR(1) | Autoregressive variable accounting for pollution abatement at the plant level |
| Blue Collar | Weighted proportion of people whose jobs are either management or professional based by census tract |
| Pink Collar | Weighted proportion of people whose jobs are in the service industry or require care-intensive labor by census tract |
| Black | Weighted proportion of African Americans at the census tract level |
| Asian | Weighted proportion of Asian Americans in census tracts |
| Other | Weighted proportion of all other races in a census tract |
| Less Than 35K Income | Weighted proportion of those who make less than $35,000 dollars a year |
| Median Household Income | Median household income of the census tract |
| At Most High School Education | Weighted proportion of people who have, at most, the equivalency of a high school degree |
| Income:Poverty Ratio Less Than One | Dummy variable for that equals one when the income to poverty ratio for a census tract is less than one, indicating poverty level is higher than income |
| Child | Weighted proportion of people under the age of 14 |
| Elderly | Weighted proportion, by census tract , of people over the age of 65 |

These variables were calculated at the census tract level in order to minimize sampling error. The census designates several geographies for data collection: the census tract, the census block, and the census block group. Of these three designations, the census tract is the largest geographic unit. The comparatively larger size of the census tract over a five year span reduces measurement error resulting from sampling, nonresponse, and undercoverage bias. Because the collection methodology differs between each unit, census blocks and census block groups tend to bias the proportions of any socioeconomic characteristics at the five year level. In addition, larger geographic block result in greater proportions of heterogenous socioeconomic groups. Following the first law of geography, larger distances allow models to relax homogeneity assumptions that govern proximate groups within smaller geographic units. Higher proportions of heterogenous groups lend well to regression analysis, and allow empirical models to assess differing socioeconomic groups within a single, common unit.

## III. **Empirical Methodology**

Like Baryshnikova (2010), I use an autoregressive model with fixed effects to difference out plant-level heterogeneity. We model pollution as the result of plant-specific fixed effects, unexplained variation in pollution over time and place, pollution abatement targets, and demographic characteristics. The dependent variable, $\Delta P_{it}$, will be measured as the change in the percentage of a pollutant between two time periods, t and t-1, on a per-plant basis, i. Before accounting for fixed effects, this model can be initially described utilizing the following equation:

$$P_{it} = \alpha_i + \rho P_{it-1} + \eta D_{it-1} + \varepsilon_{it} \quad \textbf{(2)}$$

In this equation, $\alpha_i$ represents the fixed events for a specific power plant. As in other models, we assume that this fixed effect is time invariant. $\varepsilon_{it}$ accounts for unexplained variation in pollution for each power plant at a specific time period.

Coefficient $\rho$ is derived from the gradual adjustment assumption enumerated in the theoretical model. The model assumes that the optimal levels of pollution and actual levels of pollution produced are mismatched. Thus, to predict the amount of pollution that must be abated in the subsequent time period $t+1$, actual pollution for plant $i$ in time $t$, $P_{it}$, is differenced with the optimal level of pollution for the same time period, $P_{it}^*$. Only a fraction of this target is met, thus coefficient $\gamma$ is used to represent the proportion of abatement that is realistically achievable. This is modeled in equation (3):

$$Abate_{t+1} = \gamma(P_{it} - P_{it}^*) \quad \textbf{(3)}$$

By reordering the equation, we can model the actual level of pollution in subsequent time period $t+1$ as the difference between the actual level of pollution and abatement in the subsequent period. This is shown in equation (4):

$$P_{it+1} = P_{it} - Abate_{t+1} \qquad \textbf{(4)}$$

We refer back to the theoretical model at this point. Equations (2), (3), and (4) have like terms that correspond with the dependent variable of the theoretical model. The theoretical model and equation four are combined to represent the optimal level of pollution with abatement in the next time period. This model accounts for technology and regulation parameters enumerated within the model. This is modeled in equation (5):

$$P_{it+1} = \gamma(\beta_1 T_i + \beta_2 R_i) + (1-\gamma) P_{it} + \gamma\beta_3 O_{it} + \gamma\beta_4 U_{it} \qquad \textbf{(5)}$$

From the aforementioned parameters, the first term is time invariant and can be differenced out of the equation. In addition, the penultimate and final terms can be combined such that they represent all time-variant trends in regulation. An expanded form of the empirical model for pollution is represented in equation (5). After differencing out the fixed effects, represented by the first term, we are left with equation (6):

$$\Delta P_{it\,+1} = (1 - \gamma)\Delta P_{it} + \gamma\beta_3(\Delta O_{it}) + \gamma\beta_4(\Delta U_{it}) \qquad \textbf{(6)}$$

When no observable variables determine the optimal level of pollution, we assume that $O_{it}$ and $U_{it}$ combine and represent the error term. By differencing out the fixed effects at the plant level, an autoregressive variable representing gradual adjustment to pollution can be incorporated with socioeconomic controls and instrumental variables. This results in the equation (7), where $(1 - \gamma) = \rho$. Hence, rho represents the adjustment of pollution based on the relationship between pollution levels in previous time periods. It also accounts for the predicted percentage of pollution that is abated in the subsequent period

$$\Delta P_{it\,+1} = \rho\Delta P_{it} + \Delta\varepsilon_{it} \qquad \textbf{(7)}$$

Intuitively, one would expect the sign to be positive; given lower pollution trends directed by stricter abatement regulations throughout the 2010s decade, one would expect the $\rho$ to increase over time. Utilizing the reduced regulation parameters in equation (7), we are able to implement demographic characteristics and lag all values. By lagging these values, we are able to establish our GMM model and introduce the effects demographic discrimination have on pollution reduction. The result of introducing a matrix of demographic characteristics and lagging out time-specific trends results in equation (8).

$$\Delta P_{it} = \rho\Delta P_{it-1} + \eta\Delta D_{it-1} + \Delta\varepsilon_{it} \ \textbf{(8)}$$

Correlation between demographic characteristics and regulation variables, both time-dependent and time-independent, introduce some endogeneity to our model. This endogeneity arises from aspects of regulation that are easily observed, but difficult to objectively quantify.

For example, some aspects of time-independent regulation may be influenced by the geography surrounding a power plant. Geographic regions with high political activity, such as participation in unions or environmental groups, or other relationships that are difficult to quantify, may have significant unseen interactions with socioeconomic demographics. In addition, it is commonly held that demographic trends tend to influence labor market outcomes. Long run outcomes associated with growth and productivity due to demographic change reflect structural changes that are easy to observe but difficult to measure. Wealth, education, and racial characteristics are only several of these characteristics, and are difficult to predict within our empirical model. Although we attempt to difference out time-independent effects and include instrumental variables within the model, correlations between these demographics and regulation make it difficult to accurately quantify change in our response. Thus, to quantify environmental racism, we must adopt a streamlined approach. If statistically significant differences arise in pollution levels between specific demographic characteristics, we can acknowledge that consistent over pollution is de facto evidence of discrimination. By simplifying the criteria for discrimination, we can relax some assumptions about the endogeneity resulting from unquantifiable characteristics. Without complex interactions that may arise from a plant operator's intent, or lack thereof, we can better account for a causal relationship between pollution and demographic characteristics.

In the final empirical model, $\Delta D_{it-1}$ represents the percentage change in demographic characteristics. We must transform demographic characteristics such that they provide equal weight depending on their distance from a specific powerplant and the population they contain. Demographic characteristics are derived from estimates at the census tract level. Table One in the Data section illustrates several important lifestyle and mobility population characteristics the

ACS tracks for any given census tract. They will serve as central determinants in my model mapping pollution abatement and environmental equity. Demographic characteristics are aggregated and weighted at specific radii, denoted by s. Specifically, characteristics are grouped for 25, 50, 75, and 100 mile radii surrounding power plants. I used the distance between the centroid of the census tract, k, and the power plant, i, to assign specific radii. Utilizing GIS software, I tabulated the distances between centroids of 2010 census tracts and the coordinate locations of each power plant. The following equation from Brooks and Sethi (1997) calculates spatially weighted measures for each demographic characteristic:

$$D_i = \frac{\Sigma_k D_{ik} \left(\frac{s-d_{ik}}{s}\right)}{\Sigma_k P_{ik} \left(\frac{s-d_{ik}}{s}\right)} \qquad \textbf{(9)}$$

The demographic index for plant $i$, $D_i$, is a ratio between demographic characteristics and the total population. The radius is specified such that distance between plants, $d_{ik}$, can better aggregate at a specified level. Furthermore, the interpretation of $D_i * 100$ establishes a weighted percentage for demographic characteristics. This provides the percentage of the population surrounding power plant *i* that has the corresponding demographic. Weighted percentages privilege groups that are more directly exposed to the toxins emitted by powerplants. Primarily in the case of excess emissions, which often go underregulated, the pollution values contained in the dependent variable may be considerably lower than actual pollution measurements.

After fixed effects are differenced out at the plant level, however, the inclusion of closely proximate, homogenous demographic characteristics introduces serial autocorrelation, and therefore, endogeneity. Spatially lagged demographic variables, are added to the regression in order to account for correlation between the our demographics and the error term. Similar to the methodology creating weighted demographic characteristics, temporally lagged-instrumental variables for demographics aggregated at 50 and 100 mile radii are added to the regression.

Spatially lagged instruments, also constructed in a manner similar to demographic characteristics, are created for population demographic characteristics between 25 and 50, and 50 and 100 miles around powerplants. Intuitively, characteristics are correlated, even if you move farther away from any given power plant. Characteristics that are farther should not explain unknown variation in pollution. Thus, as long as pollution decreases with distance, we can assume that these instruments are exogenous. Utilizing the Generalized Method of Moments framework, we fit both instrumental variables to a matrix in order to control for exogeneity caused by the relationship between population demographics closer to power plants. Given that pollution decreases with despite correlation between covariates, we can allow for individual specific trends that result from numerous power plants over time and place. Following a second difference in our final estimated equation, we can transform our model to eliminate time dependent effects.

**Estimations**

Table two details the effects of socioeconomic demographics on ambient air pollutants sulfur dioxide, oxides of nitrogen, and PM2.5. I utilize GMM to estimate the values of abatement coefficients in addition to the demographic characteristics. Each GMM specification implements temporally lagged variables for pollution measures and spatially lagged variables for demographic characteristics. Standard errors are clustered at the plant level, and robust against arbitrary heteroskedasticity and autocorrelation. For each specification, I consider demographic characteristics within 25 miles of any given power plant between the years 2010 and 2017.

**Table Two: Changes in Pollutants by Concentration of Demographic Characteristics**

| *Variables* | $SO_2$ | $NO_x$ | PM2.5 |
|---|---|---|---|
| AR(1) | 0.56*** | 0.61*** | 0.81*** |
| | (0.13) | (0.14) | (0.13) |
| Blue Collar | -7.88*** | -6.16*** | -4.31*** |
| | (1.10) | (1.01) | (0.59) |
| Pink Collar | -2.17** | -1.80* | -1.21** |
| | (1.07) | (1.04) | (0.55) |
| Black | 10.02*** | 5.76*** | 4.31*** |
| | (1.93) | (1.63) | (1.27) |
| Asian | -7.74*** | -3.97** | -3.56*** |
| | (1.77) | (1.62) | (1.05) |
| Other | -2.96*** | -3.56*** | -1.71*** |
| | (0.65) | (0.62) | (0.34) |
| Less than 35K Income | 1.29* | 2.08*** | 1.12*** |
| | (0.67) | (0.60) | (0.34) |
| Median Household Income | 0.07*** | 0.08*** | 0.04*** |
| | (0.01) | (0.01) | (0.01) |
| At Most HS Education | 3.82*** | 3.07*** | 1.89*** |
| | (1.26) | (1.17) | (0.66) |
| Income:Poverty Ratio <1 | 1.80** | 1.05 | 0.20 |
| | (0.80) | (0.72) | (0.44) |
| Child | -0.58 | -1.36 | 0.89 |
| | (1.49) | (1.45) | (0.80) |
| Elderly | -6.81*** | -7.29*** | -2.37*** |
| | (1.53) | (1.47) | (0.84) |
| Constant | 0.03*** | 0.03*** | 0.01 |
| | (0.01) | (0.01) | (0.01) |
| Cragg – Donald Wald F Statistic | 4.113 | 3.819 | 3.660 |
| Hansen p-value | 0.0424 | 0.0487 | 0.0064 |
| R Squared | -3.73 | -3.46 | -2.29 |

Notes:
i) Weighted regressions, standard errors clustered at Plant level
ii) * p$<$0.10, ** p$<$0.05, *** p$<$0.01
iii) GMM specification contains spatially lagged instruments constructed from data within 50-100 miles of a plant; temporally lagged instruments are constructed utilizing the second lag of demographic characteristics.

To reduce redundancy, I discuss variables that are both statistically significant and whose first stage instruments are valid. Following this, I discuss potential sources of error within my estimation strategy in greater detail. The autoregressive variable, AR(1), within the scope of the random effects model, captures additional variation. The AR(1) estimate for sulfur dioxide is significant. Utilizing the theoretical model, we interpret this abatement coefficient as the pollutant abatement rate within 25 miles of the power plant. Recall that $\rho = 1 - \gamma$; 1 – AR(1) indicates that approximately 44% of total pollution was abated between each time period. The collective amount of pollution that is abated at the 25 mile radius indicates that rates of sulfur dioxide, a particularly harmful air pollutant, can still linger throughout this gradual process.

Table three illustrates that some of the instruments, individually, contribute to the explanatory power of the model. Valid instruments for Blacks, Asians, those with less than $35,000 in income, children, and the elderly are indicative of improvement to the model's internal validity. All other variables are individually exogenous, but weak. Despite individual instrument validity, the model lacks joint validity. The Wald F-statistic, 4.113, and the Hansen p-value, 0.0424, are both below their respective thresholds. As such, the GMM specification's predictive power may suffer from a variety of problems, predominately measurement error. The construction of the instruments do not adequately capture variability in the model simply because they were not made correctly. As such, the statistically significant effects of the GMM specification within the 25 mile radius may need to be evaluated further in subsequent studies.

The GMM specification for sulfur dioxide contains numerous significant effects; all demographic characteristics, excluding the weighted proportion of children, experience significant changes in sulfur dioxide. Both the proportion of blue collar workers and the proportion of pink collar workers in a census tract experience decreases in pollution.

**Table Three: First Stage GMM F-Statistics for Sulfur Dioxide**

| | Overidentification | | Weak |
|---|---|---|---|
| | SW Chi-sq( 1) | P-val | SW F( 1, 1423) |
| Blue Collar | 69.66 | 0 | 5.78 |
| Pink Collar | 71.64 | 0 | 5.95 |
| Black | 174.73 | 0 | 14.51 |
| Asian | 354.23 | 0 | 29.42 |
| Other | 66.45 | 0 | 5.52 |
| Less than 35K Income | 177.4 | 0 | 14.73 |
| Median Household Income | 107.97 | 0 | 8.97 |
| At Most HS Education | 119.5 | 0 | 9.92 |
| Income:Poverty Ratio <1 | 90.72 | 0 | 7.53 |
| Child | 197.34 | 0 | 16.39 |
| Elderly | 170.4 | 0 | 14.15 |

A unit increase in the proportion of blue collar workers results in a 7.88% decrease in sulfur dioxide levels, while a unit increase in the proportion of pink collar workers results in a 2.17% decrease in sulfur dioxide levels. Because the spatially lagged instruments for these variables are not valid in the first stage, however, it is difficult to opine on the validity of this effect. At face value, ceteris paribus, white collar workers are exposed to higher proportions of sulfur dioxide. The result is surprising, but may be explained by two important considerations. First, the occupations for white collar workers were made up of only two categories: professionals and management. Because very few people represent this category, outliers may have the ability to introduce significant variability to our model and overpredict for our estimates. Second, many managers and professionals work in urban environments. Disregarding pollution resulting from vehicle emissions, these professionals would be exposed to greater power plant pollution given proportional demand for electricity. Thus, urban environments may contribute to the unexpected sign for our worker dummies.

Both Black Americans and Asian Americans have relevant first stage instruments. Table three's individual F-statistics and p-values indicate that the first stage instruments positively contribute to the overall explanatory power of our model. As such, their statistical significance bears some special consideration within the context of our GMM specification. A unit increase in the proportion of Black Americans coincides with a massive 10.02% increase in the total levels of sulfur dioxide. Compared to White Americans, this exponential increase in pollution corroborates the dangerous and lasting effects of discrimination in the present. After controlling for temporal gaps and autocorrelation of demographic characteristics, it appears that power plants significantly over pollute in areas with higher concentrations of Blacks. On the other hand, Asian Americans experience significant decreases, as much as 7.74% for every unit increase in the weighted proportion of Asian Americans. As a racial group, Asian Americans boast the highest average income in the United States. Recently, extreme income inequality has allowed Asian Americans to rank above even White Americans in upper percentiles of the country's wealth distribution. Compared to White Americans in the present, Asian Americans benefit from decreased exposure to sulfur dioxide. Finally, Americans of other races, ceteris paribus, experience a 2.96% decrease in sulfur dioxide levels. This variable's instruments are not valid in the first stage. Furthermore, the coding for this variable is imprecise. Because the ACS dataset did not provide statistics on other large racial and ethnic groups like Hispanics, all racial variables besides Blacks, Whites, and Asians were aggregated within the other category. Given the generalizability of this variable, I cannot make specific claims over other racial or ethnic groups. After aggregating, however, it appears that other races experience better air quality compared to White Americans.

Those who make less than $35,000 dollars in income experience a 1.29% increase in sulfur dioxide. This increase is unsurprising, given traditional relationships between income, poverty, and environmental quality. Surprisingly, increases in median household income lead to increases in pollution levels. While this variable's spatial instrument is not valid in the first stage, the coefficient's sign may be explained in a manner similar to blue and pink collar workers. Those with higher incomes may work in more urbanized areas, where there are high levels of pollution. As such, there may be an increase in variability due to outliers in the data, and this coefficient may be overestimated. Our final income-based characteristic, the income to poverty ratio, indicates that unit increases result in a 1.8% increase in sulfur dioxide. Although this predictor's spatially lagged instruments are not valid, the positive sign is not surprising. Finally, increases in the proportion of those with at most a High School degree also increase sulfur dioxide levels by 3.82%. This variable's sign is not surprising, however we cannot definitively conclude on the legitimacy of this estimate given that our first stage instruments are not valid.

The instruments for both the child and elderly predictors are valid in the first stage. However, the effect of children is not significantly different from zero. With this in mind, the GMM specification estimates that increases in the proportions of elderly Americans leads to a 6.81% decrease in sulfur dioxide levels. This may be influenced by numerous factors, among them the wealth of elderly Americans. Generational wealth is high amongst those aged 65 or greater compared to younger Americans. Because older Americans tend to be richer, they may have access to housing that is shielded from pollution or have the means to move away from areas with higher concentrations of sulfur dioxide. The wealth gap amongst younger and older Americans may also be an important determinant of ambient air quality.

Like sulfur dioxide's GMM estimation,  the NOx estimation lacks joint validity. Tests for relevance fail to surpass an F-statistic of 10 and tests for exogeneity fail to surpass a p-value of 0.1. Although the model jointly suffers from threats to internal validity, table four illustrates that individual instruments within the first stage are valid for regression. Like sulfur dioxide GMM estimates, NOx first stage instruments are all overidentified. In addition, a collection of spatially lagged instruments pass tests for relevance; instruments for Black Americans, Asian Americans, people with income less than $35,000, people with at most a high school education, children, and the elderly have F-statistics greater than 10. After accounting for linear projections in the instrumental variable matrix, a handful of robust predictors remain within this GMM model.

**Table Four: First Stage GMM F-Statistics for NOx**

| | Overidentification | | Weak |
|---|---|---|---|
| | SW Chi-sq( 1) | P-val | SW F( 1, 1423) |
| Blue Collar | 68.73 | 0 | 5.71 |
| Pink Collar | 68.26 | 0 | 5.67 |
| Black | 213.6 | 0 | 17.74 |
| Asian | 326.66 | 0 | 27.13 |
| Other | 61.41 | 0 | 5.1 |
| Less than 35K Income | 212.23 | 0 | 17.62 |
| Median Household Income | 109.72 | 0 | 9.11 |
| At Most HS Education | 126.29 | 0 | 10.49 |
| Income:Poverty Ratio <1 | 87.06 | 0 | 7.23 |
| Child | 177.71 | 0 | 14.76 |
| Elderly | 187.7 | 0 | 15.59 |

The autoregressive variable indicates that approximately 39% of NOx is successfully abated within the first 25 miles of a power plant. This difference between the optimal level of pollution and the actual level is quite significant, and indicates that year to year changes are quite sluggish in comparison to promised targets. Because there are higher proportions of sulfur dioxide in the air compared to NOx, the magnitude of change between the two pollutants is important to consider. This does not necessarily divorce power plants from accountability, though; optimal pollution estimates and respective abatement coefficients must be held at a strict standard such that health outcomes can improve for those within the immediate vicinity of a power plant.

Both Black and Asian Americans experience significant changes in NOx. A unit increase in the proportion of Black Americans yields a 5.76% increase NOx; Asians experience a 3.97% decrease. Similar to sulfur dioxide GMM estimates, the magnitude of this change is quite significant. Consistency of significance and magnitude across both pollutants are indicative of systemic differences between different American racial groups. Although race is one characteristic in the evaluation, it is important to remember that racial categorizations are arbitrary, malleable, and ingrained within the decision making heuristics of the American psyche. I draw this distinction to reiterate that past legislation and litigation have created a chasm in opportunity based on race; lower air quality is an effect of a biased cause. NOx data substantiates, after considering spatial and temporal differences, that subsets of Americans must endure reduced access to clean air. At the power plant level, this can be attributed to a decision to intentionally over pollute because these groups do not have sufficient access to recourse.

Those with incomes lower than $35,000 also experience a 2.08% increase in NOx for every one unit change. Overall, this severe gap in proportions speaks to the destitution of American poverty. Here, I draw a distinction between income and wealth. Wealth, which is a more comprehensive estimator, is not factored into my analysis. Without considering one's net assets in my model, I cannot opine on the levels of liquidity people with reduced incomes may possess. This liquidity can credibly impact, among other things, socioeconomic mobility and access to social capital. However, I claim that reduced incomes, intuitively, make it more difficult to build liquidity. Thus, Americans with the lowest incomes must continue to experience higher concentrations of ambient air pollutants, specifically sulfur dioxide and NOx. Without sufficient resources to build one's net assets over time, low income Americans are trapped in a taxing cycle of negative health outcomes and reduced access to institutional assistance.

In NOx GMM estimations, the instrument measuring the proportion of those with at most a high education becomes relevant. Similar to sulfur dioxide estimates, those with lower education levels have reduced ambient air quality. Intuitively, the sign of the coefficient matches our expectation; lower education levels correspond to lower ambient air quality. Interestingly, Baryshnikova did not find the proportion of high school dropouts to be a significant predictor of ambient air quality. This shift in significance may occur because I incorporate more recent data and also factor in the proportion of people with a high school diploma. Barring the variable construction method, however, the statistically significant effect for education contributes to the systemic differences between living conditions. The instruments for age groupings are valid in the first stage. Like in sulfur dioxide's GMM estimation, however, only the elderly predictor produces a statistically significant effect on pollution. Compared to Sulfur Dioxide, changes in NOx levels occur with reduced magnitude between the base specification. Overall, the elderly predictor's statistically significant effects across specifications connotes improved explanatory power in measuring pollution abatement.

In the final GMM specification, the AR(1) variable for PM2.5 indicates that pollution abatement at the plant level is significantly lacking. Only 19% of PM2.5-based pollution is successfully abated. Within 25 miles of every power plant, between 19-44% of pollutants are

**Table Five: First Stage GMM F-Statistics for PM2.5**

| | Overidentification | | Weak |
|---|---|---|---|
| | SW Chi-sq( 1) | P-val | SW F( 1, 1423) |
| Blue Collar | 84.78 | 0 | 7.04 |
| Pink Collar | 71.92 | 0 | 5.97 |
| Black | 111.35 | 0 | 9.25 |
| Asian | 244.16 | 0 | 20.28 |
| Other | 70.36 | 0 | 5.84 |
| Less than 35K Income | 178.76 | 0 | 14.85 |
| Median Household Income | 119.61 | 0 | 9.93 |
| At Most HS Education | 118.62 | 0 | 9.85 |
| Income:Poverty Ratio <1 | 96.1 | 0 | 7.98 |
| Child | 180.72 | 0 | 15.01 |
| Elderly | 135.92 | 0 | 11.29 |

successfully abated. Besides discriminatory practices, low abatement indicates that power plants do not have the production capital necessary to minimize pollution. This may be attributed to the high costs of production capital at the municipal and plant level. Without intervention, high costs of capital may incentivize power plants to continue with this behavior. Power plants, and their operators, seek to profit maximize. No company will invest in significant capital expenditures, simply because pollution abatement does not provide an adequate return on investment. The IRR for pollution abatement equipment is minimal. Producing electricity above demanded levels is not an efficient method of production, and increases operational costs. Furthermore, adding pollution scrubbers or other mitigation capital is only resultant of regulation or tax incentives. Even significant decreases in the price of capital do not logically incentivize polluters to adopt if regulation is inconsistent or weak. To fully consider the extent of change in the AR(1) variable, introducing predictors for asset pricing or tax incentives may be helpful in future studies.

Although the PM2.5 specification's instruments lack joint validity, instruments for Asian Americans, people with incomes lower than $35,000, children, and the elderly are all valid. As noted in table ten, instruments for Black Americans, median household income levels, and those with at most a high school education almost pass the test for relevance; their F-statistics approach, but do not surpass the threshold required in the weakness test. Every unit increase in the concentration of Asian Americans induces a 3.56% decrease in the level of pollution. People with incomes less than $35,000 also experience significant changes; PM2.5 increases by 1.12% for every unit increase. Finally, elderly adults experience a decrease of 2.37% in PM2.5 for every one unit increase. These coefficient estimates are consistent in magnitude and sign across each pollutant specification.

There are several potential sources of error that may compromise the internal validity of the models. Although I have previously discussed measurement error's effects on my spatially lagged instrument variables, I have not discussed the complexities of accurately tabulating plant level characteristics by distance and the plant's nearest neighbors. Although my model incorporates a robust sample, I use a simple Euclidian distance calculation to consider all points within a certain distance. This methodology may not accurately reflect pollution distribution, thus complex spillover effects are likely at the census tract level. Although census tracts are the largest geographic unit under the census, their shape and concentration are not uniform. The density and complex borders of a census tract contribute to spillover, which may compromise the internal validity of the model. Urban areas or areas on the coast are more likely to experience these effects. Future studies may refine this geographic approach utilizing GIS software. Combined with the autoregressive controls presented within the GMM model, modeling spillover utilizing geolocation data can minimize error.

Although sampling bias is always a general concern for modeling change in pollution over time, our unit of observation, the power plant, is relatively robust. Compared to other stationary sources of pollution, such as with pulp mills, chemical plants, or food processing facilities, pollutants emitted from power plants are well monitored and tabulated. Government agencies, such as the EPA and the EIA, measure pollution at high frequencies utilizing technology directly connected to point sources. Although the dependent variables are yearly averages of common ambient air pollutants, fine grain powerplant pollution data is collected by the minute. Subsequent improvements to the panel data could use monthly or seasonal data to measure change in power plant pollution. Accounting for seasonality in pollution levels depending on geography may also need to be factored into this analysis to control for additional variability. Besides accurate pollution tabulations, coal fired power plants are an incredibly vital source of electricity in the United States; the 1571 powerplants in the contiguous United States are only a subset of the EIA's monitored 10126 powerplants across the United States. Because powerplant pollution is diffuse and significant within this study, it would be beneficial to study powerplant pollution on a greater scale. Additional data points would increase variability in pollutant data and increase the coverage of the United States, subsequently affecting the variability of demographic characteristics within defined radii. Collectively, power plants produce two billion tons of carbon-based pollution per year, the single largest proportion of pollution within the United States (NRDC, 2020). Overall, utilizing pollution data from powerplants over time allows subsequent studies to increase generalizability and incorporate a nuanced approach to understanding stationary pollution sources.

At face value, this study indicates that pollution levels are significantly biased on racial characteristics. While racial discrepancies in pollution are indicative of a lower quality of life, they speak to the structural differences in which environmental inequalities are resolved. A lack of decision making power within communities of color speaks to the necessity of procedural justice; without proper external decision making or advocacy, pollution levels may remain disproportionately high in minority communities. Ameliorating environmental racism is one of many mechanisms to bridge stark divides in general welfare across the country; systemic differences in lived experience lead to alienation and division amongst communities of color. Although alleviating environmental inequity is one way assist disenfranchised communities, it is one of many changes needed to alleviate historic, ongoing discrimination. The skewness in the scale and magnitude of over pollution within minority communities is evidence that speaks for itself; proving ones motives is not necessary if, over time, clear differences arise based on historically held distinctions. While the ultimate goal is to eliminate over pollution in its entirety, it is integral to recognize the sum of burdens that rest on minority communities. Oftentimes, the distribution of pollution is justified by the retention of political power. Actors, whether it be firms or individuals, that locate areas of reduced political or collective action, and capitalize on inadequacies to over pollute, are a significant impetus for current trends (Hamilton, 1995). In a study modeling the production capacity of hazardous waste firms, neighborhoods with low political efficacy tended to explain greater proportions of pollution. Forced to incur the health costs and act as a buffer to other communities, disenfranchised groups have been unable to procure the resources necessary to combat waste products. Moral hazard drives the decision making of communities unaffected by pollution. There is no incentive to lower the "prices" of social and physical capital when unaffected communities bear minimal costs.

Overall, there is general consensus that there are significant correlations between socioeconomic demographics and changes in pollutants. Although effects across each GMM estimation remained consistent, the model lacks joint validity. Thus, I prioritize estimates with valid first stage instruments across a majority of the pollution models. On a consistent basis, African Americans, Asian Americans, those with low incomes, those without adequate access to education, and the elderly experienced significant changes in pollution levels. Although I supplement analysis for individual regressors throughout this discussion, these socioeconomic characteristics serve as this study's primary explanation for a significant portion of pollution variation at a national level. Unsurprisingly, characteristics beyond race dictate pollution exposure; environmental equity presents a challenge to numerous groups that exist in and outside of the socioeconomic status quo. After minimizing endogeneity from temporal and spatial characteristics through instrumentation, it is apparent that power plants fail to abate significant quantities of ambient air pollutants. The unexplained variation enumerated within the theoretical model governing pollution indicates, other things equal, complex networks of socioeconomic characteristics affect this decision to over pollute. This model attempts to explain rising social inequalities while accounting for discrimination that exists in the past and at the plant level. Reduced health outcomes indicated by pollutant levels demonstrate the compounding costs of racism and poverty over time, and the effects pollution production may have on minority communities. With reduced mobility, minority population groups may continue to act as a buffer to other, more affluent socioeconomic groups.

**Conclusion**

This study presents a General Method of Moments framework that utilizes spatially lagged instrumental variables to test the decision to over pollute at the power plant level; this decision to over pollute is based on the socioeconomic composition of nearby census tracts in which the power plant resides. In total, disenfranchised groups tend to incur a greater proportion of pollution compared to more advantaged socioeconomic groups. Given current policy and regulatory trends governing pollution reduction, this study seeks to provide a comprehensive evaluation about the disparate distribution of pollution amongst minority socioeconomic groups. Although pollution has decreased over time, it is important to recognize that current pollution standards set by regulatory agencies are still not suitable for human health. Combined with prolonged exposure to ambient air pollution and apathetic oversight, over pollution is a silent threat that exacerbates the compounding effects of inequality. Utilizing aggregated pollution data at a national level, I test whether there are significant differences across socioeconomic characteristics within the last decade, and whether power plants over pollute.

GMM specifications indicate that over pollution is consistently affected by numerous factors including, but not limited to: race, income, education, and age. Although each GMM specification's instruments lack joint significance, the model provides a succinct overview of factors dictating modern pollution exposure. After controlling for simultaneity resulting from correlation between our errors and demographic characteristics, we can better account for bias in our estimates. Utilizing regression methods, such as OLS, that do not account for endogeneity may skew our estimates. In general, spatial econometric approaches evaluating environmental equity can be vastly improved through increased focus on instrumentation. Among other general recommendations, studies can more accurately account for error caused by localized effects using this GMM and instrumental variable approach. Other approaches, such as differences in differences or regression discontinuity, may not adequately guard against threats to internal validity.

This study can improve in several important facets. The first, and most critical, is that additional data from the ACS should be factored into future analysis. Naturally, constraints imposed by the collection and subsequent tabulation of socioeconomic data limits the model.

Although time trends and location based trends may increase heterogeneity over time, a larger sample that is representative of the variability within the US population is preferred. Second, seeking additional controls to limit the spillover effects of pollution may be necessary for regression analysis. Although the model counters endogeneity through instrumentation, there are many special cases that may overweight or underweight the effects of the model, given the innumerable shapes of census tracts. Census tracts that exist on national borders may be predisposed to overprediction, given the lack of surrounding neighbors that add additional information to regression. On the other hand, census tracts that are within high density areas may be predisposed to underprediction, given that weights across time and space may not be accurately distributed. Finding a methodology to interpret and then alleviate these geographic considerations may further augment the model. Beyond these considerations, utilizing other general measures of pollution, such as vehicle emissions, may also serve as an important determination of air quality. A large portion of pollutants experienced on a daily basis are resultant of traffic congestion. Finally, additional socioeconomic controls may be helpful in increasing explanatory power. Although additional variables may lead to increased multicollinearity in future models, the framework governing environmental equity is malleable. Instituting additional variables may be necessary to control for biased statistical methods that may come with additional changes to environmental equity frameworks.

Although future improvements can be implemented, both the unit of observation and its subsequent results have important implications in measuring environmental equity. As a significant and widespread source of accurately tabulated pollution, powerplants act as a foundation for generalizability amongst differing pollution sources. Although measuring the spatial distribution of pollution requires more advanced modeling in the future, the accessibility of this pollution data is central in elucidating environmental disparities. Pollution inequity perpetuated by racial divides are symptomatic of systemic differences in mobility, wellbeing, and protection. As such, the movement to alleviate environmental racism must be viewed through a lens of awareness, solidarity, and political action. While future models may continue to reveal gaps in other socioeconomic characteristics, this study can serve as an analysis for developing comprehensive resources to combat climate change and other structural differences in minority neighborhoods. The *res ipsa loquitor* evidence of environmental inequality is only one of the injustices that disenfranchised communities must supersede in order to live a healthy life. A larger assessment that indexes systemic differences must be initialized such that American regulators can restructure their approach to aid and investment. Until this time, minority communities will unduly bear the negative financial and health externalities traditionally associated with over pollution.

Even in an era of increasing skepticism about climate change's uncertain impacts, efforts to mitigate environmental damages have had limited success in establishing and enforcing baseline levels of pollution exposure. Air pollution and air quality are important determinants of long term health outcomes. As climate policy continues to be deregulated over time, climate change mitigation is quickly transforming into a moral imperative. The complex nature of government regulation at regional, state, and federal levels complicates investments that can have extremely beneficial effects in the long run.  Prolonged divestment within economically marginalized populations illustrates the importance of investment in both pollution abatement capital and social resources. On a very visceral level, disparities that exist in environmental quality and socioeconomic quality are an existential threat. Significant disconnects between regulatory bodies and affected citizens prevent reasonable safeguards for health, mobility, and success. To continuously damage the natural environment sets a dangerous precedent for future generations; the compounding and multifaceted negative externalities presented by a climate crisis may be insurmountable if appropriate action isn't taken to limit the byproducts of industrialization. The threats posed by reduced climate quality, specifically to those who are already in precarious socioeconomic positions, must be taken seriously. Ultimately, instituting environmental protections within vulnerable communities is a way to ensure America's collective future.

Literature Cited

Baryshnikova, N. V. (2010). Pollution abatement and environmental equity: A dynamic study. *Journal of Urban Economics*, *68*(2), 183–190. https://doi.org/10.1016/j.jue.2010.04.001

Brooks, N. et al. (1997). The Distribution of Pollution: Community Characteristics and Exposure to Air Toxics. *Journal of Environmental Economics and Management*, *32*(2), 233–250.

Chakraborty, J., Maantay, J. A., & Brender, J. D. (2011). Disproportionate Proximity to Environmental Health Hazards: Methods, Models, and Measurement. *American Journal of Public Health*, *101*(S1), S27–S36. https://doi.org/10.2105/AJPH.2010.300109

Curtis, L., Rea, W., Smith-Willis, P., Fenyves, E., & Pan, Y. (2006). Adverse health effects of outdoor air pollutants. *Environment International*, *32*(6), 815–830. https://doi.org/10.1016/j.envint.2006.03.012

Gray, W. B., & Shadbegian, R. J. (2004). 'Optimal' pollution abatement—Whose benefits matter, and how much? *Journal of Environmental Economics and Management*, *47*(3), 510–534. https://doi.org/10.1016/j.jeem.2003.01.001

Hamilton, J. T. (1995). Testing for Environmental Racism: Prejudice, Profits, Political Power? *Journal of Policy Analysis and Management*, *14*(1), 107. https://doi.org/10.2307/3325435

Holifield, R. (2001). DEFINING ENVIRONMENTAL JUSTICE AND ENVIRONMENTAL RACISM. *Urban Geography*, *22*(1), 78–90. https://doi.org/10.2747/0272-3638.22.1.78

Holland, S., Mansur, E., Muller, N., & Yates, A. (2018). Decompositions and Policy Consequences of an Extraordinary Decline in Air Pollution from Electricity Generation. *National Bureau of Economic Research*, 1–42. https://doi.org/10.3386/w25339

Hornstein, Donald. (2017). Risk Analysis and Environmental Law. In *ENEC 350 Casepack Spring 2017* (1st ed., pp. 143–148). University of North Carolina - Chapel Hill.

Judy, L. (2018). Liberty and Environmental Justice for All: An Empirical Approach to Environmental Racism,. *Wake Forest Law Review*, *53*(4), 739–766.

Kan, H., Wong, C.-M., Vichit-Vadakan, N., & Qian, Z. (2010). Short-term association between sulfur dioxide and daily mortality: The Public Health and Air Pollution in Asia (PAPA) study. *Environmental Research*, *110*(3), 258–264. https://doi.org/10.1016/j.envres.2010.01.006

Kim, J. Y., Burnett, R. T., Neas, L., Thurston, G. D., Schwartz, J., Tolbert, P. E., Brunekreef, B., Goldberg, M. S., & Romieu, I. (2007). Panel discussion review: Session two — interpretation of observed associations between multiple ambient air pollutants and health effects in epidemiologic analyses. *Journal of Exposure Science & Environmental Epidemiology*, *17*(S2), S83–S89. https://doi.org/10.1038/sj.jes.7500623

Li, Z., Konisky, D. M., & Zirogiannis, N. (2019). Racial, ethnic, and income disparities in air pollution: A study of excess emissions in Texas. *PLOS ONE*, *14*(8), e0220696. https://doi.org/10.1371/journal.pone.0220696

Liu, Q., Wang, S., Zhang, W., & Li, J. (2018). Income distribution and environmental quality in China: A spatial econometric perspective. *Journal of Cleaner Production*, *205*, 14–26. https://doi.org/10.1016/j.jclepro.2018.09.090

Lohmann, L. (n.d.). *Expanding the concept of environmental racism*. 3.

Mennis, J. L., & Jordan, L. (2005). The Distribution of Environmental Equity: Exploring Spatial Nonstationarity in Multivariate Models of Air Toxic Releases. *Annals of the Association of American Geographers*, *95*(2), 249–268. https://doi.org/10.1111/j.1467-8306.2005.00459.x

Olstrup, H., Forsberg, B., Orru, H., Spanne, M., Nguyen, H., Molnár, P., & Johansson, C. (2018). Trends in air pollutants and health impacts in three Swedish cities over the past three decades. *Atmospheric Chemistry and Physics*, *18*(21), 15705–15723. https://doi.org/10.5194/acp-18-15705-2018

Wigley, D. C., & Shrader-Frechette, K. S. (1996). *Environmental Racism and Biased Methods of Risk Assessment*. *7*(1), 55–88.

Wong, C.-M., Ou, C.-Q., Chan, K.-P., Chau, Y.-K., Thach, T.-Q., Yang, L., Chung, R. Y.-N., Thomas, G. N., Peiris, J. S. M., Wong, T.-W., Hedley, A. J., & Lam, T.-H. (2008). The Effects of Air Pollution on Mortality in Socially Deprived Urban Areas in Hong Kong, China. *Environmental Health Perspectives*, *116*(9), 1189–1194. https://doi.org/10.1289/ehp.10850

Wu, Y., Li, R., Cui, L., Meng, Y., Cheng, H., & Fu, H. (2020). The high-resolution estimation of sulfur dioxide (SO2) concentration, health effect and monetary costs in Beijing. *Chemosphere*, *241*, 125031. https://doi.org/10.1016/j.chemosphere.2019.125031

Yoo Park, & Mei-Po Kwan. (2017). Multi-Contextual Segregation and Environmental Justice Research: Toward Fine-Scale Spatiotemporal Approaches. *International Journal of Environmental Research and Public Health*, *14*(10), 1205. https://doi.org/10.3390/ijerph14101205