ADDAI, NANA BOAKYEWAA. Ph.D. Estimating Avoidance Behavior from Human Mobility. (2023)
Directed by Dr. Martin Andersen. 99 pp.

The cost of damages caused by air pollution runs into billions of dollars annually. In measuring costs, researchers hardly ever include the cost of changes in behavior as a result of air pollution. Because these costs represent a loss in surplus and avoidance behavior affects the severity of health effects caused by air pollution, estimates of the effects of air pollution that do not take avoidance behavior into account are most likely biased downwards. To obtain unbiased cost estimates, some measure of avoidance behavior is needed in cost calculations for the effects of air pollution. It is important for policy makers to know the full costs of air pollution so that they can make decisions on whether to focus environmental policy on air pollution reduction or on encouraging behavior that reduces exposure to air pollution.

In this paper, I measure avoidance behavior by studying changes in human mobility patterns in response to changes in reported air quality indexes. Using data on people's mobility patterns from SafeGraph and historical air quality data from the EPA, I measure how much people change the time they spend at home, the time they spend away from home, and the distance they travel from home in response to the EPA's air quality index.

Employing a control function model to correct for selection and/or endogeneity, I find that people are on average expected to spend 8 more minutes at home, spend 10 less minutes away from home, and travel 385 meters less away from home on days of unhealthy air quality than on days of good air quality. This provides evidence of avoidance behavior in response to poor air quality.

ESTIMATING AVOIDANCE BEHAVIOR FROM HUMAN MOBILITY

by

Nana Boakyewaa Addai

A Dissertation
Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro

2023

Approved by

Dr. Martin Andersen
Committee Chair

DEDICATION

I dedicate this dissertation to every single person who has contributed in any way to my academic success: to my family, to all my teachers, to my friends and colleagues, and most importantly, to God, who has been my guide.

APPROVAL PAGE

This dissertation written by Nana Boakyewaa Addai has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Chair ____________________________
Dr. Martin Andersen

Committee Members ____________________________
Dr. Stephen Holland

____
Dr. Marie Hull

October 24, 2023
Date of Acceptance by Committee

June 20, 2023
Date of Final Oral Examination

ACKNOWLEDGEMENTS

I would like to express my profound gratitude to the following individuals and institutions whose unwavering support, guidance, and encouragement have played a pivotal role in the successful completion of this dissertation:

To Dr. Martin Andersen, I am grateful for your mentorship, patience, and expertise throughout this research process. Your guidance has been instrumental in shaping my ideas and refining this work. You showed so much interest in my progress that I had no option but to give my best.

To Dr. Stephen Holland and Dr. Marie Hull, I extend my sincerest thanks for your valuable insights and constructive feedback that helped chart the direction of my work. I will always have good memories of brainstorming sessions with Dr. Holland and encouraging emails from Dr. Hull.

To the faculty and staff at the Economics Department, I am grateful for all the support you provided throughout these four years. To Jess Saunders (who is no longer with the department), Dr. Anne Royalty and Dr. Dora Gicheva, I sincerely do not think I could have survived having a baby in the middle of the program without your support. Thank you, thank you.

To my husband, Godfred Sabbih, who always gave a listening ear to my rants about my woes as a graduate student even though he was going through similar troubles (maybe worse), I am especially grateful. You were present through it all and were my major source of motivation.

To my daughter, Lois, who (mostly) very graciously slept long enough for me to get my work done, I am thankful. I hope you do not feel pressured to pursue a graduate degree with all the research work we have done together.

To my family, who have been a solid support system throughout my academic journey, I am very grateful. You set me up for success, and I hope I have made you proud.

I am thankful to my friends and fellow graduate students who provided both academic and emotional support. The time we spent together made each challenging moment more manageable and each achievement more enjoyable.

I acknowledge the University of North Carolina at Chapel Hill's Information Technology Services for the services they provided through their Research Computing cluster.

Finally, I am grateful to the University of North Carolina at Greensboro for providing the necessary resources and a conducive environment for academic growth.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

Air quality in the United States has improved significantly over the past few decades. According to the Environmental Protection Agency (EPA), total emissions of the six principal air pollutants (carbon monoxide, lead, nitrogen oxides, ground-level ozone, particulate matter, and sulfur oxides) fell by 78% between 1970 and 2020 (US EPA, 2016). However, compared to countries like New Zealand, Sweden, Australia, and Canada, the United States still has much to do with regard to pollution management or reduction (McCarthy, 2020). As pollution control has increasing marginal costs, the cost of reducing pollution in a country like the United States, where pollution levels have been brought down significantly, becomes more and more expensive with each additional effort aimed at reducing pollution. Despite the level of spending on air pollution, exposure to pollution in the US is still quite high. The American Lung Association reported in a 2021 report that four in ten Americans live where pollution levels frequently make the air too dangerous to breathe (American Lung Association, 2021).

The amount of spending on air pollution mitigation pales is negligible compared to the cost of the effects of air pollution. The effects of air pollution are very costly, as they directly affect human wellbeing and environmental conditions (Santana et al., 2020; Chen & Chen, 2021; Delucchi et al., 2002; Gouveia & Fletcher, 2000). Between 2010 and 2020, mortality from exposure to ozone, and from exposure to PM.2.5 among adults were expected to increase by about 65% and 44%, respectively (US EPA, 2011). Adverse health conditions are not the only negative effect of air pollution. Other effects of exposure to air pollution can be seen in reduced visibility, restricted activity, and increased social misbehavior (Lu, 2020; Hyslop, 2009; Yan et al., 2019).

It is important to know the full costs of air pollution so that governments (federal and state) can make policies whose benefits will exceed those costs. It is hard to measure the full cost of air pollution as some of the effects of air pollution are not as easily observable and quantifiable as health effects. One such effect of air pollution is the change in behavior to mitigate exposure to air pollution. These changes in behavior are important of themselves as they represent a loss in utility because people give up enjoyable activities in order to avoid being exposed to pollution. Behavioral changes are also important because they indirectly affect the cost of health effects of pollution by reducing the probability of getting sick from exposure to pollutants. This type of behavior is usually referred to as avoidance behavior or averting behavior. Avoidance behavior with regards to air pollution constitutes any behavior that enables a person to stay clear of or protect themselves in environments that could potentially expose them to air pollutants. It may include staying indoors on days with high levels of air pollution, avoiding highly polluted areas, using protective coverings such as face masks, and in the longer term, migrating from highly polluted areas to less polluted areas. Estimates of the cost of the health effects (and the total costs) of air pollution that do not take into account the cost of these behavioral adjustments are biased downwards because these actions affect the frequency and severity with which people's health are affected by pollution.

The estimates from this study measure the causal effects of air quality and mobility. Measures of air quality have been found to be endogenous in the existing body of research about the relationship between air quality and human mobility. To control for this endogeneity, I employ a control function model which corrects the endogeneity problem by modelling the endogeneity in the error term. Though similar in many ways to a two-stage least squares estimation, the control function method is more advantageous as it is compatible with more model specifications (including non-linear models) and allow for heterogeneous effects. Thus, the control function method allows me to test the robustness of my results to different model specifications without changing the underlying statistical method. To the best of my knowledge, this is the first research work that employs the use of the control function method in correcting for endogeneity in the relationship between air quality and human mobility.

This study finds evidence of avoidance behavior in the United States as shown by a 1.4 percentage point increase in time spent at home, a 4.2 percentage point decrease in time spent away from home, and a 4.5 percent reduction in distance traveled from home by a device on average on days of unhealthy air quality. These translate to an 8-minute increase in time spent at home, a 10-minute reduction in time spent away from home and a 385-meter reduction in distance traveled from home on days that the EPA reports unhealthy air quality compared to days on which good air quality is reported. The study uses observed data from across the entire country and thus, findings are more reliable and generalizable than others in the literature. Results from heterogeneity analyses done as part of the study show variations in response to air pollution among different demographic groups. Minority communities, for example, are shown to be more responsive to changes in air quality than non-minority communities.

Knowledge of the extent to which avoidance measures are effective in reducing exposure to air pollution is important to policy makers, as it could serve as the basis for the development of policies geared towards making individuals play more active roles in reducing the adverse effects of pollution. If avoidance behavior reduces exposure to pollution, then it might be prudent to encourage people to practice these behaviors while the government finds more efficient ways to reduce pollution levels. It must be noted, however, that avoidance behavior has many costs aside from the cost of implementing information systems to make people aware of the levels of pollution.

As there has been a general failure to limit pollution (in the form of greenhouse gases, for instance) globally, finding a way to measure these costs is important for governments to make decisions on whether to make environmental policy more focused on mitigation and adaptation practices. Avoidance behavior, which is the focus of this paper, is a type of mitigation practice. As such, the findings from this paper contribute valuable knowledge to the decision-making process in matters concerning environmental policy.

Also, the finding that different groups respond differently to air pollution should emphasize to policymakers the importance of considering vulnerable populations when making policies related to encouraging avoidance behavior.

CHAPTER II: LITERATURE REVIEW

In this chapter, I discuss the available literature on the topic of air pollution (air quality) and its relationship with human mobility.

Air Pollution

The World Health Organization defines air pollution as the contamination of the environment by any physical, biological or chemical agent that modifies the natural characteristics of the atmosphere (World Health Organization, 2023). Relatedly, air quality refers to the degree to which air is clean enough for humans and the environment – good air quality means the air is free of harmful substances and is safe for humans and the environment The major reason why air pollution is a source of great concern is its effect on human wellbeing. Some effects on humans include poor physical and mental health, reduced worker productivity, and even poor stock market performance((M. Neidell & Pestel, 2023; S. Chen & Zhang, 2021; Chang et al., 2019; X. Zhang et al., 2017; He et al., 2023; Heyes et al., 2016). Air pollution is reported to be the biggest risk for early death, causing over six million deaths each year (Environmental Defense Fund, 2023). Studies have found that poor air quality is the major cause of many chronic respiratory and cardiovascular diseases in humans (Heutel & Ruhm, 2016; Deryugina et al., 2019; Jiang et al., 2016). To reduce the likelihood of contracting such serious health problems, human beings can make choices to control their exposure to pollutants. Whether or not they make such choices is another topic of interest to researchers. Given the importance of managing pollution in any country, there is an abundance of studies on topics related to air pollution, its measurement, and its management. The literature on the topic of the effects of pollution in relation to human health thus follows three main strands: the health costs of air pollution, human response to air pollution, and the relationship between those responses and health outcomes. I provide a selective discussion of the available literature in each strand in the following subsections.

The Health Costs of Air Pollution

The main reason why air pollution is such a concern in the world today is because of its adverse effects on human health. Exposure to air pollutants is known to affect the human respiratory, cardiovascular, nervous, urinary, and digestive systems (Kampa & Castanas, 2008). Even more alarming is the finding that exposure to air pollutants during pregnancy could affect the developing fetus in-utero (Glinianaia et al., 2004). Most studies of the effects of air pollution on health focus on respiratory and cardiovascular diseases because most air pollutants are inhaled (coming into direct contact with the respiratory system), and have the ability to bind to hemoglobin, reducing its ability to carry oxygen to the heart, and thus affecting the cardiovascular system (Kampa & Castanas, 2008). Studies have shown that diseases like asthma, pneumonia, COPD and lung cancer are exacerbated due to exposure to air pollutants such as ozone, nitrogen oxides, and particulate matter (Kurt et al., 2016; Heutel & Ruhm, 2016; Laumbach & Kipen, 2012). Aside from physical health, air pollution has also been found to adversely affect humans' mental health and affect cognitive performance (X. Zhang et al., 2017; Gatto et al., 2014; Freire et al., 2010; Clayton, 2021). Studies in the literature have found a positive relationship between air pollution and poor mental health outcomes such as depression (Borroni et al., 2022). Air pollution has even been found to increase suicide rates (Persico & Marcotte, 2022; Braithwaite et al., 2019; Xie et al., 2023).

The health effects of air pollution vary among different groups of people. Differences exist as a result of factors such as age, type of pollutant, and exposure levels. Exposure to pollution has been found to be more dangerous among children and the elderly than among younger adults

(Stone, 2000; Makri & Stilianakis, 2008; Domingo & Rovira, 2020). The type of pollutant a person is exposed to largely defines how that person's health will be affected (Pope et al., 2019; Rashidi et al., 2023; Kim et al., 2020). Furthermore, studies have found a greater effect of exposure to pollutants on the prevalence of cardiovascular diseases among the elderly (M. Liu et al., 2019), and on the prevalence of respiratory diseases among children since children are known to have weaker respiratory systems (Environmental Health (NSW), 2013). Researchers have also found that aside from children and the elderly, minority and low-income communities are particularly vulnerable to the adverse health outcomes of poor air quality (Fairburn et al., 2019; Thind et al., 2019). These differences in the magnitude of the effects of air pollution on human health among different groups are important in studying the cost of air pollution among large populations.

Measures of the health consequences of air pollution in the literature include mortality rates and number of hospital days. Since health outcomes are often observable, most researchers measure the effects of air pollution by studying changes in people's usage of healthcare facilities during or immediately after exposure to bad air. The most common measures in the literature are emergency room visits and days of hospitalization due to conditions related to bad air quality (Neidell, 2009; Moretti & Neidell, 2011; Janke, 2014). The hypothesized mechanism is that when people are exposed to some air pollutant, certain symptoms are triggered which necessitate visits to a nearby emergency room. If the symptoms are severe, they are admitted in the hospital. Thus, studying changes in the number of emergency room visits and admissions in healthcare facilities close to a place that experienced a bad air quality day or days provides a way of looking at how the bad air condition affected individuals' health on that day and perhaps the few days that followed. Since complications from conditions caused or worsened by bad air quality can get fatal, another measure of the health effects of air pollution in the literature is mortality rate from diseases known to be caused or worsened by air pollution (Dockery et al., 1993; H. R. Anderson, 2009; Di et al., 2017)

Since the most reported effects of air pollution are adverse health effects as a result of exposure, the major component of the direct cost of air pollution to humans is presented in the literature as the money value of healthcare costs related to conditions connected to exposure to polluted air. Once a researcher knows the number of hospital visits due to complications from asthma, for example, it is fairly straightforward (assuming the data is available) to get the average cost of asthma care per day in a hospital and aggregate that amount across the total number of cases to get an estimate of the total cost of hospitalizations due to asthma. However, especially in a country like the United States where healthcare costs vary widely among states, hospitals, or even patients (Kurani et al., 2021; Lagasse, 2016; Gliadkovskaya, 2021; Cummings, 2015), making generalizations from studies on hospital costs in one area may not be ideal.

Avoidance Behavior in Response to Air Pollution

Given the high costs of the effects of exposure to air pollution, it is expected that any rational individual will take non-trivial actions to avoid or reduce exposure to air pollution so as to avoid paying those costs. Aside from health effects, poor air quality also affects how much individuals enjoy doing activities outdoors. For example, taking a walk in the park with reduced visibility as a result of smog pollution would be less enjoyable than taking a walk on a clear day. When information about air quality is made available to the public, it is with the expectation that people who receive such information will take it into account in deciding their day-to-day activities. People may respond to information about air quality in one (or both) of two ways: reduce activity that exposes them to air pollutants or avoid activities that produce such pollutants. Activities that people take to reduce exposure to air pollution is what is referred to as avoidance behavior in this paper. More precisely, avoidance behavior in response to air pollution refers to any actions or measures a person may take to reduce or avoid exposure to air pollutants that have the potential to cause or aggravate certain health conditions, or which have the potential to reduce the utility individuals derive from undertaking certain activities. Since the main aim of avoidance behavior is to provide individuals with the highest amount of personal utility, given the prevailing levels of air pollution, it can be qualified as rational behavior.

Methods of avoidance vary across individuals across different pollutants. Individuals may respond differently to the same levels of pollution based on their age, lifestyle, and existing health conditions. Young children, elderly adults, and people with existing respiratory conditions like asthma, for instance, may take more measures to avoid exposure to pollution than healthy young adults (Ward and Beatty, 2015). Individuals from minority and low-income groups, even though they are more vulnerable to the effects of air pollution, may not have much freedom to practice as much avoidance behavior as they would want (or compared to other groups) due to social and economic constraints (Downey & Hawkins, 2008). In the same way, measures to avoid chemical air pollutants may differ considerably from measures to avoid particulate matter. Thus, there are different measures of avoidance to suit individual preferences. Some avoidance measures people take include staying indoors on high-pollution days, installing air filtering systems in their homes, wearing face masks, and in the longer run, migrating to places with lower pollution levels.

There is some evidence in the literature to support the notion that people exhibit avoidance behavior in response to bad air quality. In cities that publicly report air quality, alerts are issued on days when pollution levels go above thresholds that are believed to be safe for humans. In a 2017 study using administrative cycling data and government data on air pollution weather and bushfires, Saberian et al. found a 14 – 35% reduction in cycling when air quality alerts were issued in Sydney, Australia (Saberian et al., 2017). In a similar study using administrative data on zoo and observatory attendance in Southern California, Neidell found that smog alerts significantly reduced attendance at the two outdoor facilities (M. Neidell, 2009). Other studies on the evidence of avoidance behavior have found that elevated air pollution levels are associated with higher online searches for and purchases of face masks and air filters (T. Liu et al., 2018; Zhang & Mu, 2018), a reduction in attendance of outdoor sporting activities (Yoo, 2021), and a general reduction in the amount of time spent doing vigorous outdoor activities (Ward & Beatty, 2016).

It is important to study how individuals respond to pollution in the short run because although most of the health conditions associated with air pollution are chronic, most of the costs associated with these conditions are incurred within days of poor air quality, when the symptoms are aggravated.

Measuring the Cost of Avoidance Behavior

Practicing avoidance behavior does not come free. Making the decision to stay indoors to avoid pollution may mean giving up other enjoyable outdoor activities or giving up part of one's income due to the inability to work. Spending money on the purchase of air filters, face masks, and other protective wear would mean giving up spending on other useful goods. Moving away from a place with high levels of pollution may mean moving away from family and friends one would otherwise prefer to live close to. In other words, there are (opportunity) costs to avoidance behavior. Because the benefits of avoidance behavior are desirable, it is important to know the "price" one has to pay to get those benefits. Moretti and Neidell in a study in Los Angeles, for example, estimated that the cost of avoidance behavior was at least $11 million per year (Moretti & Neidell, 2011). However, this estimate did not include different measures of avoidance behavior and was just a representation of cost savings from healthcare, based on the assumption that the marginal benefit of avoidance behavior must equal its marginal cost.

Unlike healthcare costs from pollution related conditions, measuring some costs of avoidance behavior is not straightforward. For example, how does one measure how much utility individuals get from being outdoors compared to staying indoors; how does one measure how much individuals value staying close to family and friends in a polluted neighborhood? It may be relatively easy to measure the monetary value of work hours lost (W. Zhang et al., 2011), or of air filter and face mask purchases (Liu et al., 2018; Zhang & Mu, 2018), but even that may not measure the full cost of avoidance behavior as individuals would have otherwise preferred not to have purchased these items. There could also be unobserved transactional costs because individuals have to search for items they do not purchase on a regular basis.

Despite the difficulties associated with measuring some of the costs of avoidance behavior, some researchers have attempted to measure these costs in a more comprehensive way. The most common measure used in these studies is individuals' willingness to pay (WTP) to avoid pollution. In a study using the BRFSS 2006 – 2010, Jones found that US adults were willing to pay $373 to avoid one day of wildfire smoke over their county of residence (Jones, 2017). In another study, Sun and coauthors showed that individuals in urban China were willing to pay 382.6RMB on average per year to reduce air pollution (Sun et al., 2016). Although these studies do not fully measure the cost of avoidance behavior, they give an indication of how much people are willing to give up, to avoid exposure to polluted air. One problem with such measures of WTP is that they are estimated from information reported in surveys, and survey data are not perfectly reliable due to measurement error (Bound et al., 2001).

How Does Avoidance Behavior Affect Health Outcomes?

The main reason why individuals practice avoidance behavior is to avoid the adverse health effects of air pollution. People get sick from air pollution when they come into direct contact with air pollutants. Thus, the reasoning behind avoidance behavior is that if people are able to avoid exposure to pollutants, they will be able to prevent, or at least reduce the severity of illnesses related to air pollution. Though this might seem like common knowledge, many studies that look at the health effects of air pollution fail to take into account these behavioral adjustments people make in order to mitigate the health effects of air pollution.

Failing to account for avoidance behavior in measuring the health effects of air pollution could result in biased estimates (M. Neidell, 2009). If one just considers how much an individual pays in hospital bills as a result of a condition related to air pollution without considering how much that individual may have spent to mitigate the severity of her condition, for example, then there is the possibility that the total cost of the illness will be underestimated. A study on the welfare impacts of pollution alerts shows that alert issuance (and the corresponding avoidance behavior in response to the alert) reduced youth respiratory expenditures in South Korea by 30% and adult cardiovascular expenditures by 23% (Anderson et al., 2022). This finding is in support of the notion that practicing avoidance behavior improves individuals' welfare. In contrast, however, Janke found that ignoring avoidance behavior in measuring the cost of asthma admissions among children in England did not result in any statistically significant underestimation of the health effect of air pollution (Janke, 2014). Perhaps these contrasting findings are as a result of differences in the populations observed in each study, or even as a result of different specification models used.

Given that there are conflicting results in studies on the effect of avoidance behavior on the health outcomes due to air pollution, it would be difficult to estimate the welfare effects of avoidance behavior. Thus, what this study seeks to do is to provide a more reliable foundation for estimating the welfare effects of avoidance behavior by providing more reliable evidence on the existence and the magnitude of avoidance behavior in response to air pollution, especially in the United States. Although studies in the literature suggest that people in the United States adjust their behavior in response to changes in air quality, most of these studies are done on a small subsection of the population (e.g., Neidell, 2009). Those studies that were conducted at the national level use survey data to arrive at their conclusions (e.g., Ward & Beatty, 2016). As mentioned above, estimates from survey data are not entirely reliable due to potential measurement error.

What this dissertation contributes to the literature then, is more reliable evidence on the existence of avoidance behavior at the national level, using observed data rather than self-reported data. The findings from this dissertation could be the basis for policy makers to make decisions on whether to shift their attention to encouraging the general public to practice avoidance measures in the short run, while more efficient methods of pollution reduction are developed.

CHAPTER III: THEORETICAL MODEL

Problem

Governments are faced with the difficult task of pollution management. Typically, governments aim to reduce pollution by regulating emissions. Pollution reduction is expensive and has increasing marginal costs. Since the main aim of pollution control is to reduce its harmful effects on human health, if there was a way to manage exposure to pollutants in order to reduce the health effects of pollution, then it would buy policymakers some valuable time to find more efficient ways of controlling pollution. For this to be possible, policymakers must have proof that individuals care enough about adverse (health) effects of pollution to be willing to avoid being exposed. This way, immediate policy would aim at encouraging avoidance behaviors among individuals while more efficient methods of pollution reduction are investigated. The problem for decisionmakers, therefore, is to find whether or not individuals exhibit behaviors that reduce their exposure to pollution.

Utility Function

Exposure to outdoor air pollution has many adverse effects especially with regards to human health. Adverse health effects of exposure to air pollution include increased risk of incidence and worsening of respiratory and cardiovascular diseases such as asthma and bronchitis. To maintain their current level of health, individuals need to avoid exposure to air pollutants. However, to avoid pollution, individuals may have to give up or postpone certain activities that are necessary for their wellbeing, or that they find enjoyable. Thus, there is a tradeoff between maintaining good health by avoiding pollution and engaging in necessary or enjoyable activities while being exposed to pollution.

I begin my analysis of this problem with a simple utility function for an individual:

$$U(H(P(A, P_o), A), A)$$

Where H is the individual's health, P is the individual's level of pollution exposure, P_o is the current level of pollution, and A is the individual's outdoor activity level, which contributes both to her exposure to pollution and to her health. The individual also derives direct utility from engaging in outdoor activities.

This individual can choose her level of outdoor activity, which in part determines her pollution exposure. This individual's health is also affected, in part, by how active she is. Thus, the individual maximizes her utility by choosing the right amount of activity to keep her in good enough health, given the ambient level of pollution.

The first-order condition for this individual's utility maximization problem is:

$$\partial F = U_H \frac{\partial H}{\partial P} \cdot \frac{\partial P}{\partial A} + U_H \frac{\partial H}{\partial A} + U_A = 0$$

To further explore this individual's utility maximization problem, I make the assumptions below based partly on previous work by Barwick and his coauthors in their work that studied the value of pollution information (Barwick et al., 2019).

Assumptions

1. Health decreases with pollution exposure at an increasing rate and increases with higher levels of activity so that $\frac{\partial H}{\partial P} \leq 0$, $\frac{\partial^2 H}{\partial P^2} \leq 0$, and $\frac{\partial H}{\partial A} \geq 0$. Also, the marginal health benefit of being active decreases with higher levels of activity ($\frac{\partial^2 H}{\partial A^2} \leq 0$), so that the individual does not engage in unreasonable levels of activity.
2. Higher levels of pollution and higher levels of activity, especially at times of high pollution increases exposure to pollution at a decreasing rate. Also, engaging in higher levels of

activity in periods of high pollution increases exposure to pollution: $\frac{\partial P}{\partial P_o} \geq 0$, $\frac{\partial P}{\partial A} \geq 0$, $\frac{\partial^2 P}{\partial A^2} \leq 0$, and $\frac{\partial^2 P}{\partial A \partial P_o} \geq 0$.

3. Health and activity levels have diminishing marginal utility ($U_{HH} \leq 0$ and $U_{AA} \leq 0$).
4. The individual's utility function is additively separable so that $\frac{\partial^2 U}{\partial A \partial P} = 0$

Proposition

Suppose an individual values good health and understands that exposure to pollution is likely to impact her health negatively. Then, under the given assumptions, higher pollution levels reduce her observed levels of activity ($\frac{dA}{dP_o} \leq 0$).

Proof

From the implicit function theorem,

$$\frac{dA}{dP_o} = -\frac{\partial F/\partial P_o}{\partial F/\partial A} \cdot$$

Solving for $\partial F/\partial P_o$ gives:

$$\frac{\partial F}{\partial P_o} = U_{HH} \frac{\partial P}{\partial P_o} \left(\frac{\partial H}{\partial P}\right)^2 \frac{\partial P}{\partial A} + U_H \frac{\partial^2 H}{\partial P^2} \frac{\partial P}{\partial P_o} \frac{\partial P}{\partial A} + U_H \frac{\partial H}{\partial P} \frac{\partial^2 P}{\partial A \partial P_o} + U_{HH} \frac{\partial H}{\partial P} \frac{\partial P}{\partial P_o} \frac{\partial H}{\partial A} + U_H \frac{\partial^2 H}{\partial A \partial P} \frac{\partial P}{\partial P_o}$$

Similarly, $\partial F/\partial A$ gives:

$$\frac{\partial F}{\partial A} = \left(U_{HH} \frac{\partial H}{\partial P} \frac{\partial P}{\partial A} + U_{HH} \frac{\partial H}{\partial A}\right) \frac{\partial H}{\partial P} \frac{\partial P}{\partial A} + U_H \frac{\partial^2 H}{\partial A \partial P} \left(\frac{\partial P}{\partial A}\right)^2 + U_H \frac{\partial H}{\partial P} \frac{\partial^2 P}{\partial A^2}$$
$$+ \left(U_{HH} \frac{\partial H}{\partial P} \frac{\partial P}{\partial A} + U_{HH} \frac{\partial H}{\partial A}\right) \frac{\partial H}{\partial A} + U_H \frac{\partial^2 H}{\partial A^2} + U_{AA}$$

$$= U_{HH}\left(\frac{\partial H}{\partial P}\right)^2\left(\frac{\partial P}{\partial A}\right)^2 + U_{HH}\frac{\partial H}{\partial A}\frac{\partial H}{\partial P}\frac{\partial P}{\partial A} + U_H\frac{\partial^2 H}{\partial A\partial P}\left(\frac{\partial P}{\partial A}\right)^2 + U_H\frac{\partial H}{\partial P}\frac{\partial^2 P}{\partial A^2}$$

$$+ U_{HH}\frac{\partial H}{\partial P}\frac{\partial P}{\partial A}\frac{\partial H}{\partial A} + U_{HH}\left(\frac{\partial H}{\partial A}\right)^2 + U_H\frac{\partial^2 H}{\partial A^2} + U_{AA}$$

From the assumptions given, $\frac{\partial F}{\partial P_o}$ and $\frac{\partial F}{\partial A}$ are both expected to be negative and as a result, $-\frac{\partial F/\partial P_o}{\partial F/\partial A}$ is negative. Thus,

$$\frac{dA}{dP} \leq 0$$

Empirical Implication

The proposition made from this utility maximization method can be tested empirically. To do that, I would have to find a way to observe individuals' activity levels on different days, and test whether or not these activity levels change with different levels of pollution. I use the word "observe" because, since it is considered rational behavior to avoid exposure to pollution, if asked to report their behaviors, individuals may fail to disclose any contradictory behaviors. Observing individuals' behaviors thus provides a means of testing the hypothesis that individuals engage in lower levels of activity on days of poor air quality compared to days of good air quality.

My model predicts that higher levels of pollution will lead to a reduction in individuals' observed levels of activity. This means that individuals who value good health and are aware of the adverse health effects of air pollution will take steps to adjust their exposure to pollutants. Thus, empirically, I expect to observe a reduction in activity levels on poor air quality days.

CHAPTER IV: DATA

Mobility Data

This paper uses data on peoples' movement patterns from SafeGraph's Social Distancing Metrics (SDM)[1]. Safegraph's SDM was live during the COVID-19 pandemic and provided daily data on device locations using pings from anonymous mobile devices in the United States(*Social Distancing Metrics | SafeGraph Docs*, 2020). SafeGraph analyzes six weeks of data during night hours (between 6pm and 7am) until they obtain sufficient evidence to assign a home location for each device, which is then mapped to a census block group (CBG). The devices are aggregated by home census block group and provide metrics such as device count (number of devices observed that day for that home CBG) and time spent at home for each CBG. The SDM is no longer supported by SafeGraph and is only available for January 2019 to April 2021.

For my analyses, I create three variables from the SDM to show total time spent at home by all devices in each CBG (Home Time), the total time in hours spent away from home (Away Time), and the total distance traveled away from home in meters (Distance from Home). The original variables from SafeGraph data record the distribution of time spent at home and away, and distance traveled from the home location based on pre-specified buckets, and the number of devices in each bucket. I create my variables by multiplying the number of devices in each bucket by the midpoint of the time or distance in the bucket. For the final distance bucket, I used the bottom value of the bucket. Any averages of these variables are calculated by dividing the totals by the number of devices in the CBG (device count).

[1] SafeGraph sold their Patterns business to Advan Research in 2022. I refer to these data as "SafeGraph" data for consistency with prior research.

It must be noted that the SDM only recorded pings from mobile devices, and so the metrics do not account for the location of a device for the entire day. As such, it is expected that the metrics "Home Time" and "Away Time" for each day may not add up to twenty-four hours. Also, the variable for time spent away from home may be an imperfect measure of avoidance behavior as individuals may spend time at pollution-safe locations other than their homes.

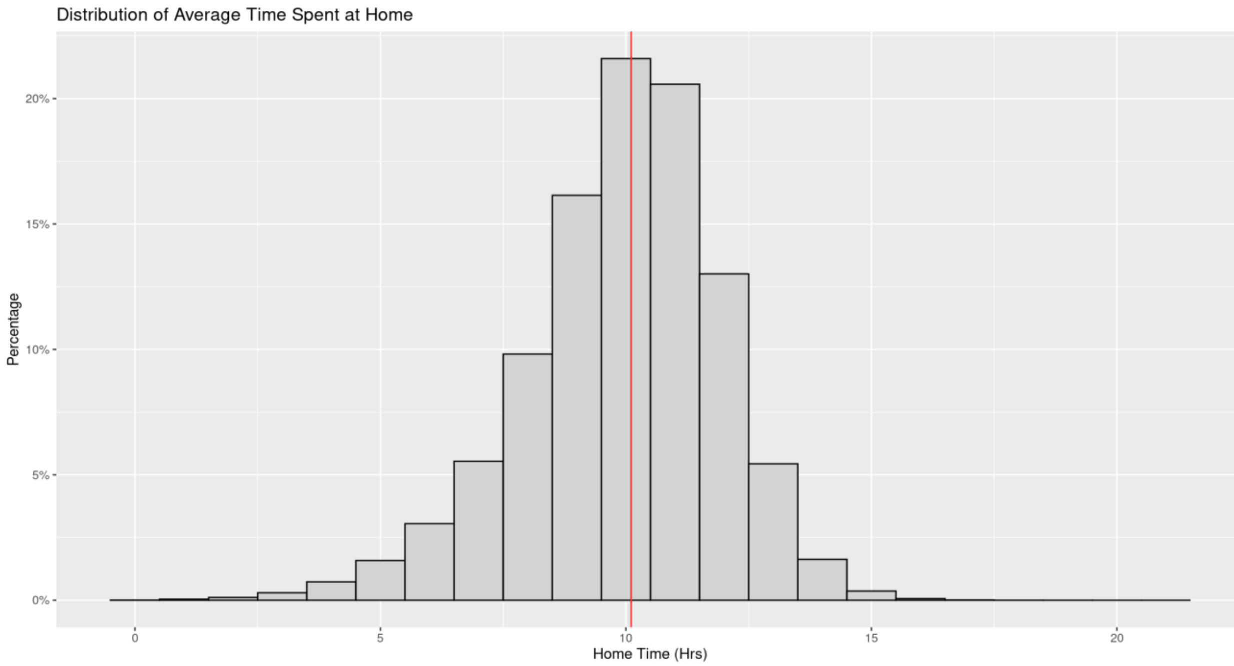
Since the nighttime location of a device is not necessarily its home (the owner of the device could be a nighttime worker), there is the potential for measurement error in all variables that use the home location of a device as a reference point for measuring mobility. This category of variables includes my created variables Home Time, Away Time, and Distance from Home. However, considering that the proportion of Americans in nighttime jobs is relatively small (Lieberman et al., 2020), and that SafeGraph computes a new home for each device at the start of each month relying on the evening and nighttime behavior of devices over the previous six weeks (Andersen, 2020), I expect potential measurement errors to be small.

Despite the potential for measurement error in my outcome variables, I still expect to obtain consistent results though the results may be less precise than using variables with no measurement errors (see Appendix A).

I use daily data for the continental USA from 2019 in my analyses because 2019 is the only year with full mobility information for all days, in which people's movement are not affected by government regulations; COVID-19 lockdown policies most likely impact movement patterns for 2020.

Figures 1-3 show the distributions of each of the three created variables in the dataset.

Figure 1. Distribution of time spent at home



It is seen from figures 1 and 2, that individuals spend about 10 hours at home, and about 3 hours away from home on average per day. As mentioned, the average number of hours spent at home and away from home do not add up to 24 hours. However, since the data accounts for more than half of the day, I expect the results to be representative of people's movement patterns. Figure 3 shows that people travel an average of about 8km away from home daily.

Figure 2. Distribution of time spent away from home

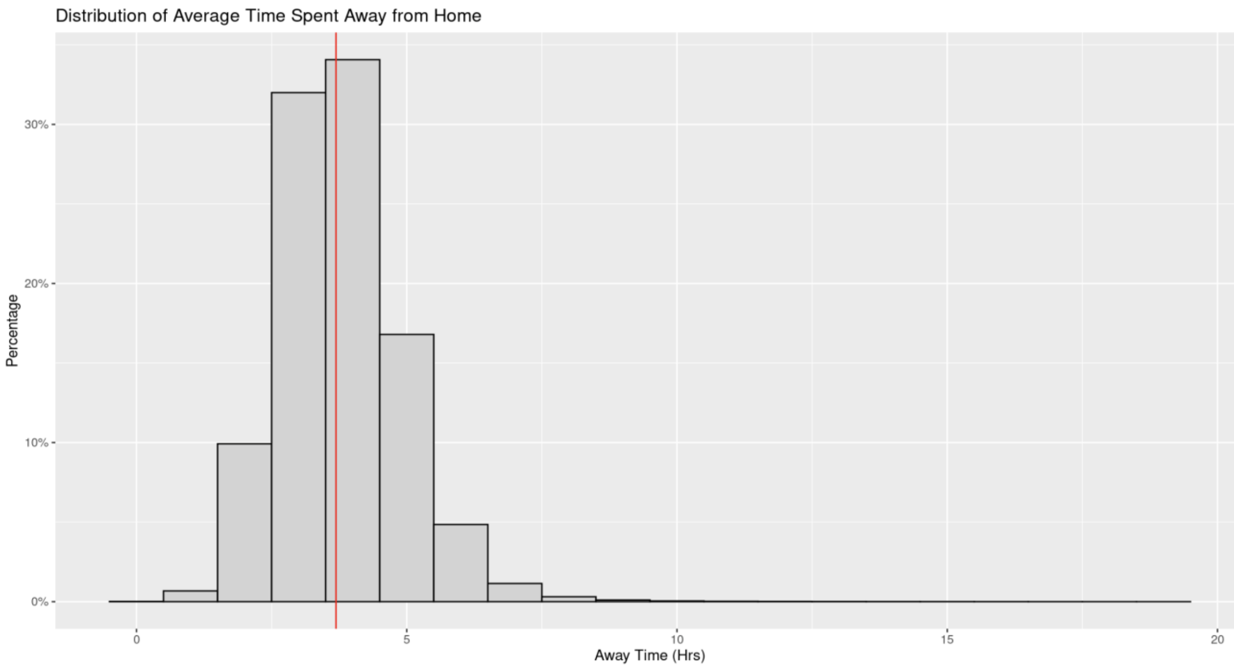
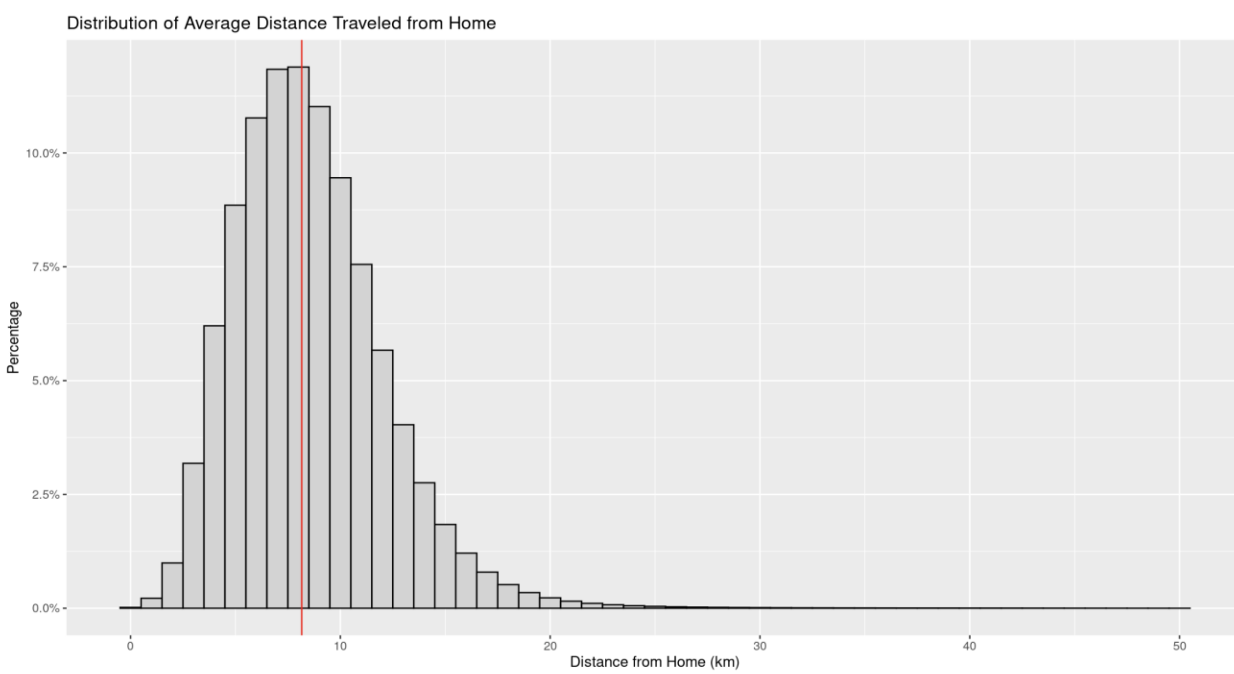


Figure 3. Distribution of distance traveled from home



To check for variations in the measures for time spent at and away from home, and for distance travelled from home across regions, I decompose my data into five regions: Northeast, Southeast, Midwest, West, and Southwest. I see a similar pattern in each region, with median time spent at home in the West about an hour greater than the median across all regions. People in the Northeastern part of the United States also appear to travel shorter distances from home on average than people from other regions. Figures for the distribution of the mobility measures for each region can be found in Appendix B.

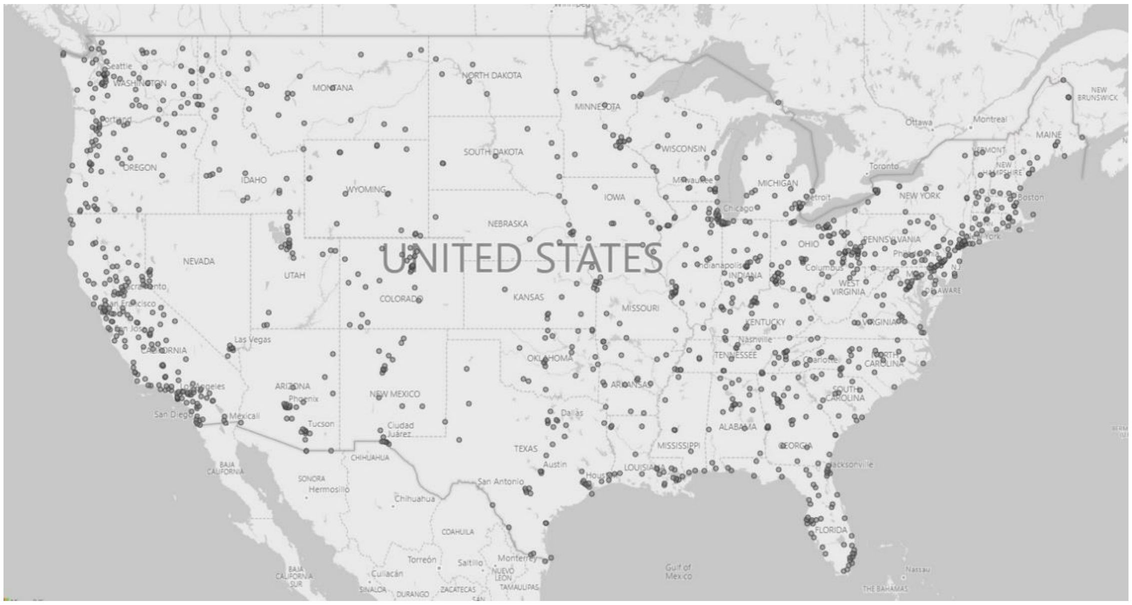
SafeGraph data also shows some variation in mobility patterns across seasons. Less mobility is seen during the colder months of the year as people are seen to spend more time at home and travel shorter distances from home than in the warm months. People are seen to travel the farthest distances from home on average in Summer (see Appendix C).

Pollution Data

Historical daily air quality indexes (AQI) are available for the seven main pollutants from all outdoor air quality monitoring sites in the United States on EPA's official website. In my analyses, I use AQI values based on levels of PM2.5 because particulate pollution can be more easily observed than other forms of pollution (may appear as haze), and is a threat throughout the year, not just in warm weather. AQI values based on levels of PM2.5 are available from 1999 to yesterday.

I use daily AQI data from 2019 to conduct my analyses based on the availability and reliability of mobility data. Since AQI data are available at the monitoring station level, and mobility data are at the CBG level, I restrict my sample to CBGs with a monitoring station within a 10-mile radius from their centroids. For CBGs with more than 1 monitor within that radius, I use the weighted mean of the air quality measurements from all stations within the CBG, with inverse distance from the centroid as weight. Figure 4 shows the locations of all monitors included in my analyses.

Figure 4. Monitor locations



The EPA reports daily AQI data on AirNow.gov on a scale that runs from 0 to 500, with higher numbers indicating worse air quality. Each category on the AQI scale has a specific color and corresponds to a different level of health concern for the public. Weather apps that report AQI follow the same color coding as the EPA. I create a variable which shows the level of concern for each day's AQI using that scale. Table 1 gives a summary of the categories of EPA's AQI scale, their representative colors, their corresponding health concern messages, and the percentage of days in my data that record each level of air quality on the scale. I show the distribution of air quality across the country and in each region and season in Appendices D and E.

AQI values below 100 are generally considered safe. Days with AQI values between 101 and 150 are considered unhealthy for sensitive groups which include young children, the elderly, and people with respiratory and cardiovascular diseases, while AQI values above 150 are considered unhealthy for the entire population. Thus, I create a variable for alert days with AQI values when the entire population would be concerned about the ambient air quality (AQI values above 150). I use this alert variable in sensitivity analyses to detect any changes in behavior in response to air quality alerts.

Table 1. Summary of AQI Scale

Daily AQI Value	Color	Level of Concern	% of Days in Data
0 - 50	Green	Good	85.47
51 - 100	Yellow	Moderate	14.43
101 - 150	Orange	Unhealthy for Sensitive Groups	0.09
151 - 200	Red	Unhealthy	0.01
201 - 300	Purple	Very Unhealthy	0.00
301 and higher	Maroon	Hazardous	0.00

As shown in table 1, PM2.5 in the United States is usually at safe levels: the share of CBGs exposed to worse than moderate air quality is less than 5% on average. Thus, it is expected that any day that experiences higher than normal levels of AQI will be accompanied by observable changes in movement behaviors. Other pollutants may, however, lead to worse air quality.

Weather Data

In order to be able to obtain reliable estimates for avoidance behavior, I include controls for weather in my analyses. It is important to include weather controls because weather elements such as temperature and humidity may affect the level of particulate pollution in the atmosphere. At the same time, these weather elements may also affect people's decisions to take part in outdoor activities. Thus, failing to include them may bias any estimates obtained from my analyses.

Data on weather conditions in the United States for 2019 are obtained from gridMet. Measurements for temperature, wind speed, precipitation, humidity and solar radiation are given at the county level. I merge weather data to the other datasets based on the assumption that weather conditions are uniform across CBGs within a county. Where maximum and minimum daily values are reported (e.g., temperature), I calculate daily averages.

Table 2 shows summary statistics for each measure of mobility, air quality and weather I use in my analyses. It can be seen from the table that individuals spend an average of about 10 hours at home and 4 hours away from home each day. Though one would normally expect time spent at home and time spent away from home each day to add up to 24 hours, the means of the time measures of mobility in the table do not add up to a full day since Safegraph only records pings from the use of applications on mobile devices rather than tracking the device throughout the day. Nevertheless, I am comfortable drawing conclusions from analyses using the data I have since it covers well over half of each day on average. Table 2 also shows that air quality in the United States is generally good (well below 50 on average) and that most of the data are clustered around the means.

I present a breakdown of my summary statistics by region and season in Appendices F and G, respectively. Tables F24 to F28 show that there was generally more movement seen in the eastern parts of the United States in 2019 than in the Midwest and the western parts. Tables G29 to G32 show more mobility during the summer than the other seasons. These statistics suggest that location and time may have an effect on mobility patterns. As such, it is important to control for such variations in movement patterns in my analyses.

Table 2. Summary statistics

Variable	Mean	Standard Deviation	25th Percentile	75th Percentile
Home Time (hrs)	9.92	1.99	8.78	11.26
Away Time (hrs)	3.79	1.12	3.02	4.44
Distance from Home (km)	8.55	3.58	6.03	10.56
Daily Air Quality Index	32.14	15.68	20.50	42.00
Temperature (°F)	58.08	18.25	44.78	72.77
Precipitation (in.)	3.04	8.29	0.00	1.50

Wind Speed (mph)	8.86	4.07	6.04	10.96
Shortwave Radiation (w/m^2)	196.38	91.12	117.70	274.70
Humidity (%)	64.72	15.02	56.90	74.80
N = 37,676,135				

**All summaries are at done at the CBG level for each day and are not population weighted*

Because of concerns for endogeneity in the air quality variable, I obtain wind speed and wind direction data from Persico and Marcotte (2022) to construct instrumental variables for air quality. The authors collected data on wind speed and wind direction from the North American Regional Reanalysis (NARR) daily reanalysis data, which reports wind conditions on a 32 by 32 kilometer grid and consist of vector pairs, one for the east-west wind direction (u-component) and one for the north-south wind direction (v-component). They converted the average u- and v-components into wind direction and wind speed after locating each wind monitor in a county and then averaged up to the county-day level, simplified to four quadrants. They defined "wind direction" as the direction the wind is blowing from (Persico &Marcotte, 2022).

Other Data

To conduct my heterogeneity analyses, I get data on the demographic characteristics drawn by Andersen et al. (2023) from the 2015 to 2019 American Community Survey. The data show the number of individuals by age, race and gender for each CBG (Andersen et al., 2023). I am able to identify CBGs that are predominantly white, those that have a larger share of children or the elderly, and those that have a larger share of females.

I conduct a set of simple bivariate regressions to determine the relationships that exist between the demographic characteristics of each CBG and the number of devices recorded for each day. I find that minority CBGs are associated with lower device counts in the data on average than predominantly white CBGs. Apart from implying that individuals in minority CBGs use a smaller number of devices, this finding could also mean that the individuals in these CBGs do not use the applications from which SafeGraph records pings. I also find that CBGs with larger shares of the elderly and females are associated with lower device counts on average. As expected, CBGs with bigger populations have higher device counts on average. I present results from these regressions in the first column of table 3. I also present the mean for each variable in column 2 of the table for some context.

Additionally, I obtain data on median income for CBGs from IPUMS NHGIS which provides population, housing, agricultural, and economic data along with shapefiles for geographic units across the United States (Manson et al., 2022).

Finally, I use data from the North American Industry Classification System (NAICS) to sort locations recorded in the monthly neighborhood patterns data from SafeGraph into indoor and outdoor locations. Locations I classify as indoor include museums, amusement arcades, casinos, and theater companies. I classify amusement and theme parks, golf courses, skiing facilities, zoos, nature parks, among others, as outdoor locations. I do this classification to be able to measure how peoples' responses to poor air quality differ in terms of the type of locations they visit.

Table 3. Results from Regressions on Device Count

Variable	**Estimate** **(1)**	**Mean** **(2)**
Device Count	-	95.44
Share of Population Black	-16.2154*** (0.07)	0.18
Share of Population Hispanic	22.7076*** (0.09)	0.13

Share of Population from Other Minority Racial Groups	-6.1526*** (0.15)	0.07
Share of Population White	2.0930*** (0.05)	0.52
Share of Population Minority	-3.7360*** (0.04)	0.48
Share of Population Adult	14.4519*** (0.17)	0.63
Share of Population Children	82.2301*** (0.19)	0.21
Share of Population Elderly	-85.8662*** (0.17)	0.15
Share of population Female	-16.1395*** (0.26)	0.51
Share of Population Male	16.1395*** (0.26)	0.49
Total Population	0.0623*** (0.00)	1480

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

CHAPTER V: EMPIRICAL MODELS

Fixed Effects OLS Model

To measure the effect of air quality on people's movement patterns, I first consider a basic fixed effects[2] model using OLS:

$$\log\,(Mob_{ibt}) = a_b + c_t + \beta_1 \log\,(AQ_{bt}) + \beta_2 X_{bt} + u_{ibt} \qquad (1),$$

Ideally, Mob_{ibt} should denote a measure of mobility (the same model will be run for all three measures of mobility in the data) for individual i at CBG b on date t, but since the available data is aggregated at the CBG level, I use the device average for each measure of mobility in each CBG. AQ_{bt} denotes the AQI value reported by the EPA for CBG b on date t, X_{bt} is a vector of weather controls including temperature, precipitation and humidity, a_b and c_t represent CBG and date fixed effects respectively, and u_{ibt} is the error term. The coefficient of interest is β_1, which measures the relationship between air quality and mobility. Standard errors are clustered on CBGs in all models, and each regression is weighted by the number of devices recorded in a CBG on each day.

Since the EPA also reports levels of air quality as shown in table 1, I create a model with a categorical version of the air quality variable:

$$\log\,(Mob_{ibt}) = a_b + c_t + \beta_1 f(AQ_{bt}^{cat}) + \beta_2 X_{bt} + u_{ibt} \qquad (2),$$

where $f(AQ_{bt}^{cat})$ represents AQI bins or categories for each day.

To study the relationship between air quality and elements of the weather, I run a regression with daily AQI on the left-hand-side and weather variables on the right-hand-side:

[2] I include fixed effects to control for any seasonal and regional variations in mobility patterns. It is shown in appendices F and G that such variations exist in my data.

$$AQ_{bt} = a_b + c_t + \beta_1 X_{bt} + u_{ibt} \tag{3}$$

The coefficient β_1 in each of the fixed effects OLS models (equations 1 and 2) estimates the relationship between air quality and mobility. A positive coefficient would imply a positive relationship between air pollution and mobility (AQI is higher with higher levels of pollution), suggesting that individuals do not limit (and may even increase) their movement in response to poor air quality. Conversely, a statistically significant negative coefficient would suggest that worsening air quality is associated with reduced mobility indicating a potential "avoidance behavior" effect.

Control Function Model

To obtain unbiased estimates, it is necessary to ensure that any change in movement patterns found from my model is purely a result of changes in air quality (causal identification). Considering that most of the research on this topic find their measure of air quality to contain some measurement error (Saberian et al., 2017; Barwick et al., 2019; Graff Zivin & Neidell, 2013; Kim, 2021; Moretti & Neidell, 2011; M. J. Neidell, 2004; Persico & Marcotte, 2022), I am careful to draw conclusions on the relationship between air quality and mobility based on the estimates from my fixed effects OLS model. Additionally, to be able to interpret my estimates as causal effects, I would have to be fully convinced that my model has controlled for every possible factor that can affect both air quality and mobility. Since I am not certain my model includes all those factors and given the common findings of endogeneity in the literature, I presume the presence of endogeneity in my model.

To investigate potential endogeneity in my model (and to get consistent estimates in the case that my model is truly endogenous), I estimate a two-stage control function (CF) model using wind direction interacted with groups of counties as instruments. Within a geographic area, daily wind direction is a good instrument for air quality because it brings an endogenous source of variation in air quality (Persico & Marcotte, 2022b). I obtain wind direction data from Persico and Marcotte (2022) and follow their method of using kmeans clustering to get 200 groups of monitors that minimize the distances within each cluster and run the first stage regression of air quality on the interaction of the 200 monitor groups with four different bins of wind direction and include all the controls from my OLS model.

$$\log\,(AQ_{bt}) = a_b + c_t + \beta_1 WD_{bt} + \beta_2 X_{bt} + v_{bt} \tag{4}$$

For the second stage, I include the residuals from the first stage (v_{bt}) in the righthand-side of my model.

$$\log\,(Mob_{ibt}) = a_b + c_t + \beta_1 \log(AQ_{bt}) + \beta_2 X_{bt} + \beta_3 \hat{v}_{bt} + e_{ibt} \tag{5}$$

AQ_{bt} in equation (1) is suspected to be endogenous because of its potential correlation with the error term u_{ibt}. By including v_{bt} from equation (4) as an explanatory variable in equation (5), I obtain a new error term e_{ibt} which is uncorrelated with all other variables on the righthand-side. Thus, the air quality variable in equation (5) is no longer endogenous. In other words, v_{bt} can be thought of as a proxy for the factors in u_{ibt} that are correlated with AQ_{bt} (Wooldridge, 2015).

Even though the control function model is mathematically identical to a two-stage least squares model in the case where the endogenous explanatory variable (EEV) appears linearly, I choose to use the control function model because it can easily be adapted to fit cases where the EEV is not continuous (as is the case in my model using a categorical measure of air quality). The CF approach is useful in nonlinear models and allow for heterogeneous effects, where effects at the individual level can differ from effects at the aggregate level. The CF method is likely to be more efficient than the 2SLS method, though it may not be as robust. Additionally, the CF approach provides a simple test of the null hypothesis that AQ_{bt} is exogenous (Wooldridge, 2015).

This makes the CF approach a better choice for checking the robustness of one's results across various specification models.

The interpretations of the coefficients on the air quality variables in the CF model are similar to those of the OLS model. It must be noted that the interpretations of the estimates from both the OLS and CF models (and effectually all the models in this paper) are based on a general sense of mobility. Thus, in the case of my measures of mobility, the interpretation of the coefficients in the model using the variable "Home Time" would be the reverse of the interpretation of the coefficients in the models with the variables "Away Time" and "Distance Travelled" because higher values of "Home Time" represent less mobility.

Overall, the results from these models will contribute to our understanding of how changes in air quality affect individuals' behavior and decision-making, potentially providing insights for policymakers in developing strategies to manage pollution exposure.

CHAPTER VI: RESULTS

Main Results

I present results from the CF model and the OLS model side-by-side in table 4 for easy comparison. The first column of the table shows the relationship between air quality and the weather controls used in my regression models. Results show sizeable and statistically significant estimates of the effects of the weather variables on air quality. Columns 2, 4 and 6 of the table show the results from equation (1) for the three different measures of mobility, while columns 3, 5 and 7 show results from equation (5).

Results from the OLS model show statistically significant positive relationships between mobility and poor air quality as seen by a reduction in time spent at home and increases in time spent away from home and distance traveled from home, though the estimates are small. These results do not reflect what one would expect intuitively given the relationship between air quality and health and how much individuals value good health outcomes. The results also do not align with the findings of negative relationships between air quality and human movement in the literature. This divergence from the general consensus could be as a result of two things: (1) there is truly a positive relationship between poor air quality and mobility which researchers have failed to find over the years or, (2) estimates from the fixed effects OLS model I used are biased.

As mentioned earlier, my goal for estimating the CF regression is to investigate and correct for potential endogeneity in my OLS model.

Table 4. Regression results for measures of air quality and mobility (OLS and CF models)

Variable	Log of Daily AQI	Log of Time Spent at Home		Log of Time Spent Away from Home		Log of Distance Traveled from Home	
		OLS	CF	OLS	CF	OLS	CF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of Daily AQI		-0.0029***	-0.0010	0.0016**	-0.0007	0.0022***	-0.0111***
		(-11.43)	(-1.43)	(5.40)	(-0.90)	(8.82)	(-15.00)
Average Temperature	-0.0202***	-0.0017***	-0.0017***	0.0005***	0.0004***	0.0043***	0.0040***
	(-206.09)	(-30.84)	(-30.26)	(6.66)	(6.05)	(73.29)	(67.78)
Average Temperature Squared	0.0003***	0.0001***	0.0001***	-0.0001***	-0.0001***	-0.0001***	-0.0001***
	(310.15)	(19.37)	(17.25)	(-8.06)	(-6.85)	(-54.52)	(-43.60)
Precipitation	-0.0017***	0.0001***	0.0001***	-0.0006***	-0.0006***	-0.0008***	-0.0008***
	(-135.17)	(16.30)	(16.48)	(-73.66)	(-72.74)	(-98.94)	(-100.91)
Wind Speed	-0.0348***	0.0004***	0.0005***	-0.0003***	-0.0004***	-0.0012***	-0.0017***
	(-471.09)	(17.73)	(14.86)	(-11.67)	(-10.11)	(-45.60)	(-44.45)
Shortwave Radiation	0.0008***	-0.0001***	-0.0001***	0.0001***	0.0001***	0.0001***	0.0001***
	(177.51)	(-13.81)	(-14.28)	(40.71)	(40.40)	(55.41)	(58.20)
Humidity	0.0015***	0.0001***	0.0001***	-0.0004***	-0.0004***	-0.0003***	-0.0003***
	(61.54)	(9.44)	(9.17)	(-27.93)	(-27.67)	(-26.91)	(-25.15)
FS Residual			-0.0021***		0.0025**		0.0144***
			(-3.40)		(3.08)		(19.91)

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include CBG and date fixed effects. N = 37,622,412

While column 3 still shows a positive relationship between poor air quality and mobility, columns 5 and 7 show an inverse relationship: as seen in a decrease in average time spent away from home and a reduction in the distance travelled from home, holding weather conditions constant. The coefficients on the weather controls from the two models are almost identical in magnitude and sign. “FS Residual” in the table refers to the coefficient on the error term from the first stage regression. Its value can be essentially interpreted as a Hausman test for endogeneity: a statistically significant value shows that the OLS model is statistically different from the CF model. The high statistical significance of estimates of “FS Residual” shows proof of the presence of endogeneity in my OLS model.

As seen from table 4, the results from the CF regression do not show a clear relationship between air quality and mobility: poor air quality is seen to have a negative and insignificant relationship with both time spent at and away from home. This is most likely because rather than unit changes in the air quality index, individuals may respond more visibly to changes in the level of concern (what I call AQI scale) reported by the EPA (refer to table 1). In other words, individuals are more likely to be concerned about a change in the AQI scale from “Unsafe for Sensitive Groups” to “Unhealthy” than a change in the AQI from 149-150. For the same reasons as before, I estimate a control function model using a categorical variable for the AQI scale and show the results in table 5 along with estimates from OLS models using the same variable.

Table 5. Regression results for measures of air quality and mobility using categorical measure of air quality (OLS and CF models)

Variables	Log of Time Spent at Home		Log of Time Spent Away from Home		Log of Distance Traveled from Home	
	OLS (1)	CF (2)	OLS (3)	CF (4)	OLS (5)	CF (6)
Moderate AQ	-0.0016***	0.0002	0.0004	-0.0009**	-0.0001	-0.0025***
	(-5.21)	(0.62)	(1.33)	(-2.89)	(-0.31)	(-8.95)
AQ Unsafe for Sensitive Groups	-0.0001	0.0040**	-0.0160***	-0.0187***	-0.0091***	-0.0145***
	(-0.01)	(3.01)	(-8.93)	(-10.28)	(-4.46)	(-7.01)
Unhealthy AQ	0.0089***	0.0140***	-0.0380***	-0.0416***	-0.0385***	-0.0454***
	(3.89)	(6.05)	(-9.60)	(-10.41)	(-8.48)	(-9.96)
Average Temperature	-0.0017***	-0.0016***	0.0005***	0.0005***	0.0042***	0.0042***
	(-29.99)	(-29.64)	(6.24)	(6.06)	(72.16)	(71.61)
Average Temperature Squared	0.0001***	0.0001***	-0.0001***	-0.0001***	-0.0001***	-0.0001***
	(18.27)	(17.85)	(-7.46)	(-7.23)	(-53.51)	(-52.91)
Precipitation	0.0001***	0.0001***	-0.0006***	-0.0006***	-0.0008***	-0.0008***
	(16.90)	(17.14)	(-74.35)	(-74.56)	(-99.63)	(-99.93)
Wind Speed	0.0005***	0.0005***	-0.0004***	-0.0004***	-0.0013***	-0.0013***
	(22.50)	(23.68)	(-14.34)	(-14.97)	(-51.67)	(-53.12)
Shortwave Radiation	-0.0001***	-0.0001***	0.0001***	0.0001***	0.0001***	0.0001***
	(-15.06)	(-15.21)	(41.614	(41.73)	(56.85)	(57.06)
Humidity	0.0001***	0.0001***	-0.0004***	-0.0004***	-0.0003***	-0.0003***
	(9.21)	(9.14)	(-27.94)	(-27.94)	(-26.75)	(-26.67)
FS Residual		-0.0031***		0.0022***		0.0042***
		(-13.26)		(7.13)		(16.31)

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include CBG and date fixed effects. N = 37,622,412

Columns 2, 4 and 6 of table 5 show a negative relationship between mobility and poor air quality. We see a 1.4 percentage point (8-minute) increase in time spent at home, a 4.2 percentage point (10-minute) reduction in time spent away from home, and a 4.5 percentage point (385-meter) reduction in the distance traveled from home on a day that records unhealthy air quality compared to a day that records good air quality. The results from the CF models also show a progressive reduction in mobility with worse air quality: there is less mobility on average on days that record air quality that is unsafe for sensitive groups than on days that record moderate air quality, and even less on days that record unhealthy air quality. All estimates from the CF models are statistically significant at least at the 5% level.

Sensitivity Analyses

The results presented above in table 5 estimate daily variation of mobility within each CBG. The fixed effects included essentially control for constant characteristics within each CBG that do not vary over time. Since most legislation (for administrative purposes) is done at the state level, I explore within state variation over time by estimating another set of regressions with state and day fixed effects and present the results in table 6. I find that on days of unhealthy air quality, there is a 3-minute increase in time spent at home, a 5-minute decrease in time spent away from home, and a 331-meter decrease in distance traveled from home compared to days of good air quality. These estimates are smaller than those obtained from the models with CBG fixed effects.

Table 6. Results from CF model using state by day fixed effects

Variable	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)
Moderate AQ	-0.0005 (-1.44)	-0.0014*** (-4.26)	-0.0012*** (-3.82)

AQ Unsafe for Sensitive Groups	0.0035* (2.43)	-0.0152*** (-7.51)	-0.0189*** (-8.39)
Unhealthy AQ	0.0050* (2.07)	-0.0220*** (-5.33)	-0.0388*** (-8.46)
FS Residual	0.0002 (0.58)	0.0003 (0.65)	0.0004 (1.19)

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 37,622,679

The difference between estimates from the two sets of regressions could be because some of details controlled for in the model with CBG fixed effects (columns 2, 4, and 6 of table 5) were missed in the model with state fixed effects (table 6). To investigate the validity of this supposition, I estimate the models with county fixed effects and county-cluster fixed effects and show results in appendix H33 and H34 respectively. I expect that there will be some within county variation since weather and AQI are usually reported at the county level, and my data reports wind direction at the county-cluster level. Results from these sets of regressions show a 6-minute increase in time spent at home, a 6-minute decrease in time spent away from home, a 84-meter reduction in distance traveled from home for the regressions with county fixed effects, and a 4-minute increase in time spent at home, a 5-minute decrease in time spent away from home, and a 166-meter decrease in distance traveled from home for the on days of unhealthy air quality compared to days of good air quality for the regression with county-cluster fixed effects.

It should be noted that the estimates for "FS Residual" in table 6 are not statistically significant. From my interpretation of "FS Residual", this means air quality is not endogenous in the model with state by day fixed effects. In other words, I could have obtained unbiased results by using the OLS model. Further analyses are required to draw a conclusion as to why endogeneity does not seem to be a problem in the model with state by day fixed effects.

To explore other ways my results will respond to changes in my main model, I run several regressions with different versions of the model. First, I run a set of regressions without weights. Results from these regressions (table H35 in the appendix) show a slight decrease in estimates for all three measures of mobility. This is most likely because some CBGs record many more devices than others, and thus responses from those CBGs were influencing the results. Including weights thus evened out those differences in effects. The small size of the differences between this set of regressions and the original set does not give reason to worry about excluding weights in the original set of regressions.

To ensure that my results are not being overly influenced by outliers, I run regressions with versions of my data without the top and bottom 1% and 5%. I run these regressions only for distance traveled because the time measures of mobility are bounded (between 0 and 24 hours) and thus are less likely to contain outliers. Outliers in the distance traveled measure could show up in the data if, for example, the owner of a device went on a long trip. Results from this set of regressions (shown in table H36 in the appendix) show a 348- and a 310-meter reduction in distance travelled from home on days of unhealthy air quality compared to days of good air quality for the 99%- and 95%-complete data set versions of the regressions, respectively. These estimates are not too far off from the estimate obtained in my original analysis, so I do not worry much about outliers in the data.

Finally, I run a model in which I collapse the "Unsafe for Sensitive Groups" and "Unhealthy" categories of the air quality variable into one category which represents air quality that is considered unsafe for any group. This enables me to measure the extent to which individuals respond to poor air quality in general, and it facilitates my heterogeneity analyses to come. Results from this model (presented in table 7) show similar estimates to the estimates for

"Unsafe for Sensitive Groups" from the main model. This is to be expected as the number of days that record "Unhealthy" air quality in my data set is not enough to influence the results if both categories of air quality are combined.

Table 7. Results from Model which Combines "Unsafe for Sensitive Groups" and "Unhealthy" Categories in One Category

Variables	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)
Moderate AQ	0.0002 (0.61) [0.0003]	-0.0009** (-2.88) [0.0003]	-0.0025*** (-8.93) [0.0003]
Unsafe AQ	0.0049*** (3.99) [0.0012]	-0.0210*** (-12.28) [0.0017]	-0.0175*** (-8.98) [0.0019]
FS Residual	-0.0031*** (-13.25) [0.0002]	0.0022*** (7.11) [0.0003]	0.0042*** (16.28) [0.0003]

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 37,622,679

Model Allowing for Non-Linearity in AQI Variable

My linear model suggests that the relationship between air quality and mobility is log-linear. However, it could be that the relationship is nonlinear such that the direction of the change in mobility is different for different levels of air quality (for example, mobility could reduce at some high values of air quality and increase at very high values of air quality.

To address the concern that the relationship between air quality and mobility is nonlinear, I estimate a model that allows for nonlinearity. I do this by including an interaction term between the air quality variable and interaction term that equals 1 if the air quality index is 100 or more. To make the presence of any nonlinearity clear, I center the AQI variable around 100. I show this in equation (6):

$$\log(Mob_{ibt}) = a_b + c_t + \beta_1 AQC_{bt} + \beta_2 AQ100_{bt} + \beta_3 AQC_{bt} * AQ100_{bt} + \beta_4 X_{bt} + \beta_5 \hat{v}_{bt} + e_{ibt} \qquad (6),$$

where AQC_{bt} is an AQ variable centered around 100 and $AQ100_{bt}$ is an indicator variable that equals 1 if the air quality index in CBG b on day t reaches 100.

I show results from this regression in table 8. It can be seen from the results that the relationship between air quality and mobility changes at the 100 AQI mark such that there is a greater reduction in mobility when the air quality index exceeds 100. The results in table 8 is similar to the findings in my main model and justify the use of the categorical version of the AQI variable when measuring the relationship between air quality and mobility.

Table 8. Results from model allowing for nonlinearity in the relationship between air quality and mobility

Variable	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)
Log of Centered AQI	-0.0010 (-1.45) [0.0007]	-0.0007 (-0.83) [0.0008]	-0.0111*** (-14.93) [0.0007]
AQI > 100	0.0007 (0.44) [0.0016]	-0.0137*** (-5.33) [0.0026]	-0.0088** (-3.07) [0.0029]
Log of Centered AQI*(AQI > 100)	0.0238*** (4.01) [0.0059]	-0.0381*** (-3.94) [0.0097]	-0.0375*** (-3.40) [0.0110]
FS Residual	-0.0021*** (-3.44) [0.0006]	0.0026** (3.19) [0.0008]	0.0145*** (20.02) [0.0007]

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 37,622,679

Heterogeneity Analyses

After confirming that my results are robust to variations on my main model, I check for differences in the effect of air quality on mobility across demographic groups. This is important because, should governments make policies based on avoidance behavior, it is imperative to ensure that those policies take into account the possibility that some groups may be more vulnerable to the effects of air pollution and engage in more avoidance behavior. I check for differences in my results based on characteristics such as age, race, income levels, and the type of locations people visit.

To investigate whether poor air quality influences the type of location individuals visit, I estimate another set of regressions separately for indoor locations and outdoor locations based on NAICS classification. I classify locations with NAICS codes starting with I obtained similar results (see table I37 in the appendix) for both locations, suggesting that changes in mobility as a result of changes in air quality is not determined by where people plan to go but rather whether or not they choose to leave their homes. Further analysis is required to reach a conclusion on this finding.

From the regressions measuring differences in response to air pollution by age (results shown in table 9), I find statistically significant reductions in mobility on days of moderate air quality for CBGs with larger shares of children and the elderly compared to those with a larger share of adult population. On days of unsafe air quality, the response is generally the opposite, though the estimates are either very small or statistically insignificant. This appears to be an unexpected finding at first glance, as it would be expected that children (or their parents) and the elderly would be more concerned about unsafe air quality than moderate air quality because of their susceptibility to respiratory and cardiovascular diseases. To explain these unexpected results, there is the possibility that since sensitive groups are more aware of the dangers of living in polluted areas, they are more likely to avoid such areas. Hence, these results might be suggesting that children and the elderly live in relatively less polluted communities and as such, hardly experience days of unhealthy air quality. However, since this proposition has not been empirically tested, I refrain from drawing any conclusions based on it.

Table 9. Age Differences in the Response to Changes in Air Quality

Variable	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)
Moderate AQ	-0.0094*** (-5.09)	0.0085*** (5.20)	0.0185*** (9.67)
Unsafe AQ	0.0121. (1.92)	-0.0373*** (-4.77)	-0.0340* (-2.51)
Moderate AQ*Share of Population Elderly	0.0275*** (6.18)	-0.0160*** (-3.94)	-0.0561*** (-11.68)
Unsafe AQ*Share of Population Elderly	-0.0014 (-0.08)	0.0336 (1.44)	0.0138 (0.41)
Moderate AQ*Share of Population Children	0.0257*** (4.43)	-0.0313*** (-6.47)	-0.0584*** (-10.64)
Unsafe AQ*Share of Population Children	-0.0276 (-1.50)	0.0513* (2.26)	0.0574 (1.64)
FS Residual	-0.0032*** (-13.78)	0.0021*** (7.33)	0.0043*** (16.70)

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 37,546,688

In my next set of regressions (results shown in table 10), I find that minority racial groups are significantly more likely to reduce their mobility with higher levels of pollution. Specifically, individuals living in predominantly Black, Hispanic, Asian, and Pacific CBGs, and CBGs in which majority of the population are from other minority groups, are expected to spend about 4 minutes less away from home, and travel 110 meters less away from home on days of unsafe air quality compared to individuals living in predominantly white CBGs. This indicates that individuals living in these communities are more likely to cancel or postpone an outing due to poor air quality than the average American.

Table 10. Racial differences in the response to air pollution

Variable	Time Spent at Home	Time Spent Away from Home	Distance Traveled from Home
Moderate AQ	0.0017***	0.0022***	-0.0004
	(5.02)	(6.08)	(-1.18)
Unsafe AQ	0.0036*	-0.0110***	-0.0103***
	(1.98)	(-4.75)	(-3.53)
Moderate AQ*Minority CBG	-0.0031**	-0.0065***	-0.0044***
	(-5.31)	(-12.27)	(-8.08)
Unsafe AQ* Minority CBG	0.0020	-0.0161***	-0.0129***
	(0.83)	(-5.20)	(-3.35)
FS Residual	-0.0031***	0.0020***	0.0042***
	(-13.57)	(7.14)	(16.18)

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 37,622,679

I conduct further analysis by looking at each racial group separately. I find a bigger increase in time spent at home for Black CBGs, and a bigger decrease in time spent away from home and distance traveled from home for both Black and Hispanic CBGs. This suggests that though minority communities are in general more likely to respond to pollution than non-minority communities, Black and Hispanic communities are much more inclined to change their mobility patterns. This is in line with the finding that respiratory diseases such as asthma and lung cancer disproportionately affect African American and Hispanic communities (Memoli, 2020; The NHLBI Working Group, 1995). Results for the more detailed racial analysis can be found in appendix I38.

Next, I run a model to measure differences in the response to air quality based on income. I create an indicator variable to separate low-income and high-income CBGs based on whether the median income in a CBG is lower or higher than the national median income for 2019 and show results in table 11. Results from a model interacting income status with a variable indicating whether a CBG experienced unsafe air quality on a day show that people living in low-income CBGs are expected to be less mobile on days that the air quality is considered unsafe compared to people living in high-income CBGs.

Table 11. Differences in Response to Changes in Air Quality Based on Differences in Median Income

Variables	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)
Unsafe AQ	0.0037 (1.57) [0.0024]	-0.0063* (-2.21) [0.0028]	-0.0013 (-0.41) [0.0031]
FS Residual	-0.0016*** (-3.49) [0.0005]	0.0028*** (4.59) [0.0006]	0.0035*** (6.83) [0.0005]
Unsafe*Low income	0.0018 (0.58) [0.0031]	-0.0146*** (-3.61) [0.0041]	-0.0146** (-3.25) [0.0045]

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 7,948,257

Finally, I run a model to measure the differences in the response to air quality by looking at CBGs whose median air quality for 2019 was above or below the median AQI for the entire data set. I call CBGs with AQI above the median "dirty" CBGs and those below the median "clean" CBGs. I find that individuals living in "clean" CBGs are less likely to reduce their mobility when the air quality is considered unsafe. Results for this set of regressions are presented in table 12.

Table 12. Differences in Response to Changes in Air Quality Based on Whether a CBG is Considered “Dirty” or “Clean”

Variables	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)
Unsafe AQ	0.0064*** (4.57) [0.0014]	-0.0250*** (-12.65) [0.0020]	-0.0246*** (-9.61) [0.0026]
FS Residual	-0.0031*** (-12.33) [0.0003]	0.0019*** (6.41) [0.0003]	0.0032*** (8.65) [0.0004]
Unsafe AQ*Clean CBG	-0.0057* (-2.22) [0.0026]	0.0160*** (4.38) [0.0036]	0.0405*** (8.87) [0.0046]

Significance codes: ‘’ 0.05, ‘**’ 0.01, ‘***’ 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 37,622,679

The differences found in how various groups respond to air pollution demonstrate how important it is for policy makers to ensure that policies geared towards reducing exposure to air pollution are inclusive and equitable. For example, parents with young children might take more action to reduce their children’s exposure to pollution at the expense of their own convenience, or that individuals in low-income and minority communities are more likely to lose significant parts of their incomes from staying away from work due to pollution.

CHAPTER VII: EXTENSION TO RESULTS

Multinomial Logit Model

To obtain measures of mobility compatible with the models above, I create three variables that measure time spent at home, time spent away from home and distance traveled from home by aggregating several sets of time and distance bins in my mobility dataset. Though aggregation makes it possible to carry out my analyses conveniently, this convenience comes at the expense of dealing with potentially inaccurate conclusions. Though my main model deals with any such possibilities and I expect my results to have a causal interpretation, it would be better, if it were possible and if it made more sense, to limit the level of contentiousness of my results by using the original variables in the data.

In tables 13, 14, and 15, I show the weighted mean (WM) and the weighted standard deviation (WSD) for the share of devices in each time bin and the daily air quality for each CBG.

Table 13. Summary statistics for home bin variables

	Share of Devices		**Daily AQI**	
Time in Hours	**WM** **(1)**	**WSD** **(2)**	**WM** **(3)**	**WSD** **(4)**
Under 1	0.24	0.05	31.97	12.48
1 to 6	0.14	0.02	31.92	12.51
6 to 12	0.17	0.03	31.87	12.56
12 to 18	0.20	0.04	31.88	12.60
Over 18	0.24	0.06	31.93	12.79

Table 14. Summary statistics for away bin variables

	Share of Devices		Daily AQI	
Time in Hours	**WM** (1)	**WSD** (2)	**WM** (1)	**WSD** (2)
Under 1	0.46	0.06	0.46	0.06
1 to 6	0.25	0.03	0.25	0.03
6 to 12	0.19	0.05	0.19	0.05
12 to 18	0.05	0.01	0.05	0.01
Over 18	0.04	0.02	0.04	0.02

Table 15. Summary statistics for distance bin variables

	Share of Devices		Daily AQI	
Distance in Meters	**WM** (1)	**WSD** (2)	**WM** (3)	**WSD** (4)
No Travel	0.33	0.05	31.93	12.69
1 to 1,000	0.09	0.02	31.81	12.57
1,001 to 2,000	0.05	0.01	31.87	12.64
2,001 to 8,000	0.23	0.04	31.87	12.49
8,001 to 16,000	0.12	0.03	31.95	12.58
16,001 to 50,000	0.10	0.04	32.08	12.76
Over 50,000	0.09	0.04	31.91	12.33

Traditionally, an ordered logit model would be employed in the case of an ordinal dependent variable such as one where there are ordered time or distance bins. Though the original binned mobility variables are in this order, the ordered logit model is not the best model for this kind of data. This is because the main assumption for the ordered logit model to be valid is the parallel regression assumption where the relationship between the independent and dependent variables is the same across all levels of the dependent variable. This is usually true in the case where we know that one category of the dependent variable is greater than the other, but we do not know by how much such as in the case of a Likert scale. In the case of my binned variables, the relationship between the dependent and independent variables are likely to differ at each category of the dependent variables. This violates the parallel regression assumption and makes the use of an ordered logit model invalid in this case.

In such a case where the ordered logit model cannot be used, the multinomial logit model is employed. As such, I employ a multinomial logistic regression model to estimate the relationship between air quality and mobility using the original binned variables as represented in equation (6) below:

$$\log\left(\frac{\pi_j}{\pi_k}\right) = a_b + c_t + \beta_{1j}\log(AQ_{bt}) + \beta_{2j}X_{bt} + \beta_{3j}\hat{v}_{bt} + e_{jbt} \tag{6}$$

where π_j is the probability that a device falls in a distance or time bin other than the baseline, and π_k is the probability that a device falls in the baseline bin. In all my regressions, I use the bin that represents the most time spent at home or least distance traveled from home as the baseline. I include the same set of controls as in my main model. Equation (6) is comparable to equation (5) in that it is a second stage regression in a two-stage model. The first stage in this model is identical to the one in the main model and it has the same function of handling potential measurement error in the AQ variable. I show results for distance traveled, home time, and away time in tables 16, 17 and 18, respectively.

Table 16. Results from multinomial logit regression on binned distance variables

Variables	1 to 1000 meters	1001 to 2000 meters	2001 to 8000 meters	8001 to 16000 meters	16001 to 50000 meters	Above 50000 meters
Moderate AQ	0.0053** (2.85)	0.0211*** (8.93)	-0.0173*** (-8.63)	-0.0274*** (-12.16)	0.0151*** (4.83)	-0.0819*** (-27.75)
USG AQ	-0.1018*** (-4.52)	-0.0642* (-2.11)	-0.1198*** (-2.47)	-0.1334*** (-4.48)	-0.1605*** (-3.97)	-0.2339*** (-9.73)
Unhealthy AQ	-0.2459*** (-6.82)	-0.1942** (-2.73)	-0.1747* (-2.47)	-0.2885*** (-4.14)	0.0681 (0.83)	0.0843 (0.9859)
N = 2,220,638						
Fixed Effects:						
CBG	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Time (Date)	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 2,220,638

Results in table 16 suggest that devices are generally less likely to fall in the longer distance bins on days when air quality is deemed unsafe for sensitive groups than on good air quality days. For days of unhealthy air quality, the likelihood of a device falling in the longest distance bins increases. This seems to suggest that, rather than stay at home, some individuals travel far away from their homes on days of poor air quality. However, these estimates are not statistically significant and so I cannot draw any conclusions on avoidance behavior based off them. The results for days of moderate air quality are not as easy to interpret as there is no clear-cut pattern.

Table 17. Results from multinomial logit regression on binned home-time variables

Variables	Less Than 1 Hour	1 to 6 Hours	6 to 12 Hours	12 to 18 Hours
Moderate AQ	-0.0228*** (-10.02)	-0.0104*** (-5.35)	-0.0043. (-1.89)	-0.0178*** (-9.17)

USG AQ	-0.1144*** (-6.92)	-0.0244 (-1.47)	-0.0008 (-0.03)	-0.0551* (-2.20)
Unhealthy AQ	-0.2188*** (-6.91)	-0.1856*** (-4.68)	-0.2565*** (-5.13)	-0.2282*** (-5.16)

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 1,586,170

Table 18. Results from multinomial logit regression on binned away-time variables

Variables	1 to 6 Hours	6 to 12 Hours	12 to 18 Hours	18 to 24 Hours
Moderate AQ	0.0001 (0.10)	-0.0207*** (-8.52)	-0.0013 (-5.98)	-0.0051. (-1.80)
USG AQ	-0.1088*** (-6.03)	-0.0708* (-2.18)	-0.1137*** (-5.98)	0.2157*** (-8.83)
Unhealthy AQ	-0.0966. (-1.91)	-0.2626*** (-3.46)	0.0330 (0.68)	0.1290 (1.60)

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 1,586,170

Results from the regressions using time bins (shown in tables 17 and 18) also show that devices are more likely to fall in the bins that represent more time spent at home or less time spent away from home on days of poor air quality compared to days of good air quality. In other words, on a day of poor air quality, devices are more likely to reduce time spent away from home or increase time spent at home compared to a day of good air quality. Reference categories for the time variables are the time bins that represent the least time spent outside the home (less than 1 hour for away time and more than 18 hours for home time). As such, in this case, estimates for both the home time and away time bins have similar interpretations.

Though these results generally suggest a reduction in mobility on days of poor air quality, the estimates are difficult to interpret individually, and one can only make sense of them by inferring a general conclusion from looking at several estimates. This beats the purpose of running the model with the original binned variables as having to make general conclusions from the estimates is akin to aggregating the variables in the first place. Thus, this provides some justification for using the model with the variables derived from aggregating the original variables.

Models with Lagged Air Quality

To test if air quality in the past few days affects how individuals respond to current air quality, I run two models that include lags of the air quality index for each day. In the first model, I include five lags of an air quality "alert" indicator variable which equals 1 if air quality for a day exceeds 100 to see how far out in time air quality affects mobility. I include only the fifth lag in the second model to show how the air quality index published five days before affects the relationship between air quality and mobility on a certain day. The difference between the two models is that the former includes five lags of the "alert" variable, while the latter includes only the fifth lag of the continuous AQI variable. While one answers the question of how alerts issued days before affect behaviors on a certain day, the other answers the question of how much influence air quality from five previous days has on present day behavior. I show results in tables 19 and 20, respectively.

Table 19 shows a reduction in the magnitudes of estimates for time spent at home, time spent away from home, and distance traveled from home on days when the air quality index is 100 or more. Specifically, we see about a 2-minute increase in time spent at home, a 3-minute reduction in time spent away from home, and a 91-meter reduction in distance traveled from home. These results show that air quality conditions from previous days affect how people change their mobility patterns in response to contemporaneous air quality.

It can also be seen from table 19 that air quality alerts (assuming that alerts are issued on the days that air quality index exceeds 100) generally continue to affect mobility up to three days after they are issued.

Table 19. Results from model with 5 lags of AQI

Variable	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)
AQI > 100			
Contemporaneous	0.0041*** (4.22) [0.0010]	-0.0138*** (-9.58) [0.0014]	-0.0106*** (-6.28) [0.0017]
Lag 1	0.0035*** (4.4526) [0.0008]	-0.0149*** (-11.5149) [0.0013]	-0.0128*** (-8.6432) [0.0015]
Lag 2	0.0031*** (3.9589) [0.0008]	-0.0099*** (-6.6783) [0.0015]	-0.0077*** (-4.8372) [0.0016]
Lag 3	-0.0016* (-1.9921) [0.0008]	-0.0052*** (-3.9191) [0.0013]	-0.0037* (-2.4948) [0.0015]
Lag 4	-0.0009 (-1.1920) [0.0008]	-0.0093*** (-7.3964) [0.0013]	-0.0069*** (-4.8161) [0.0014]
Lag 5	-0.0054*** (-5.0672) [0.0011]	-0.0086*** (-5.9752) [0.0014]	-0.0003 (-0.1993) [0.0016]
FS Residual			
Contemporaneous	-0.0033*** (-21.0436) [0.0002]	0.0004* (1.9903) [0.0002]	0.0043*** (22.6101) [0.0002]

Variable	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)
Lag 1	0.0004*** (3.8435) [0.0001]	-0.0022*** (-14.7264) [0.0001]	-0.0018*** (-11.7463) [0.0002]
Lag 2	-0.0017*** (-15.1121) [0.0001]	0.0009*** (6.2060) [0.0001]	0.0015*** (10.2823) [0.0001]
Lag 3	-0.0010*** (-9.0994) [0.0001]	0.0013*** (8.5815) [0.0001]	0.0019*** (13.3518) [0.0001]
Lag 4	-0.0007*** (-6.8717) [0.0001]	0.0003* (2.1402) [0.0001]	0.0001 (0.4677) [0.0001]
Lag 5	-0.0003. (-1.8573) [0.0001]	0.0012*** (6.9252) [0.0002]	0.0005** (2.9081) [0.0002]

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 37,434,728

Table 20 shows that the response to air quality decreases significantly when the air quality index from five days before is taken into consideration. This could be because when individuals experience high pollution on some day, they are more likely to put measures in place to reduce exposure to pollution without affecting their activity levels. Experiencing poor air quality on a day may motivate an individual to switch to activities that reduce exposure to pollution or to engage in activities in more pollution-safe environment over the next few days. For example, the individual may use a treadmill in a gym, rather than going running in a park on the days following days of high pollution.

Table 20. Results from model with 5th lag of AQI

Variable	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)
Log of Daily AQI	0.0000. (-1.7853) [0.0000]	-0.0001*** (-6.1678) [0.0000]	-0.0001*** (-8.6497) [0.0000]
5^{th} Lag of AQI	0.0000 (0.4612) [0.0000]	-0.0001*** (-4.4223) [0.0000]	-0.0001*** (-9.6163) [0.0000]
FS Residual	-0.0025*** (-7.0642) [0.0004]	0.0031*** (6.7473) [0.0000]	0.0073*** (18.1678) [0.0004]
5^{th} Lag of FS Residual	-0.0009* (-2.3776) [0.0004]	0.0033*** (6.8932) [0.0001]	0.0041*** (9.8548) [0.0004]

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls, 5^{th} lag of weather controls and CBG and date fixed effects.

N = 37,434,728

Poisson Model

One final extension to my analysis is running a model without logging my outcome variables. Even though the mobility variables I use are bounded between 1 and 200 (my data does not include air quality indices of more than 200) and there is little likelihood for outliers, I still use a log transformation because of the very important advantage of having an elasticity interpretation. To assure any readers who might argue that a log transformation was not necessary in this case because of the general absence of skewness in the outcome variables (as shown in figures 1-3), I run a fixed effects Poisson model that is essentially the non-logged version of my main model. I present my results in table 21.

Results shown in table 21 shows that a change in the level of air quality from good to moderate increases the odds of reducing mobility by about 1%. We see a progression in the magnitudes of the estimates as air quality worsens. These results are similar to those from the model with logged mobility variables both in magnitude and direction, thus I feel confident that I do not lose much information about the relationship between air quality and mobility by log transforming the variables in my main model.

Table 21. Results from Poisson model

Variables	Time Spent at Home (1)	Time Spent Away from Home (2)	Distance Traveled from Home (3)
Moderate AQ	0.0006*** (3.6367) [0.0002]	-0.0006* (-2.3479) [0.0003]	-0.0019*** (-4.5758) [0.0004]
AQ Unsafe for Sensitive Groups	0.0027** (2.8270) [0.0010]	-0.0202*** (-12.0966) [0.0017]	-0.0165*** (-6.0808) [0.0027]
Unhealthy AQ	0.0116*** (5.7477) [0.0020]	-0.0428*** (-11.3212) [0.0038]	-0.0481*** (-10.3722) [0.0046]
FS Residual	0.0031*** (-19.2762) [0.0002]	0.0017*** (6.4208) [0.0003]	0.0039*** (10.6037) [0.0004]

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects. N = 37,434,728

Regression Discontinuity Model

Because of the adverse health effects of air pollution on human health, different counties have different programs that monitor air pollution and issue alerts on days that the air quality in that region goes beyond approved thresholds. Depending on the county, alerts go out either when the air quality exceeds 100 - code orange, or when it exceeds 150 - code red (*Action Days | AirNow.Gov*, 2023).

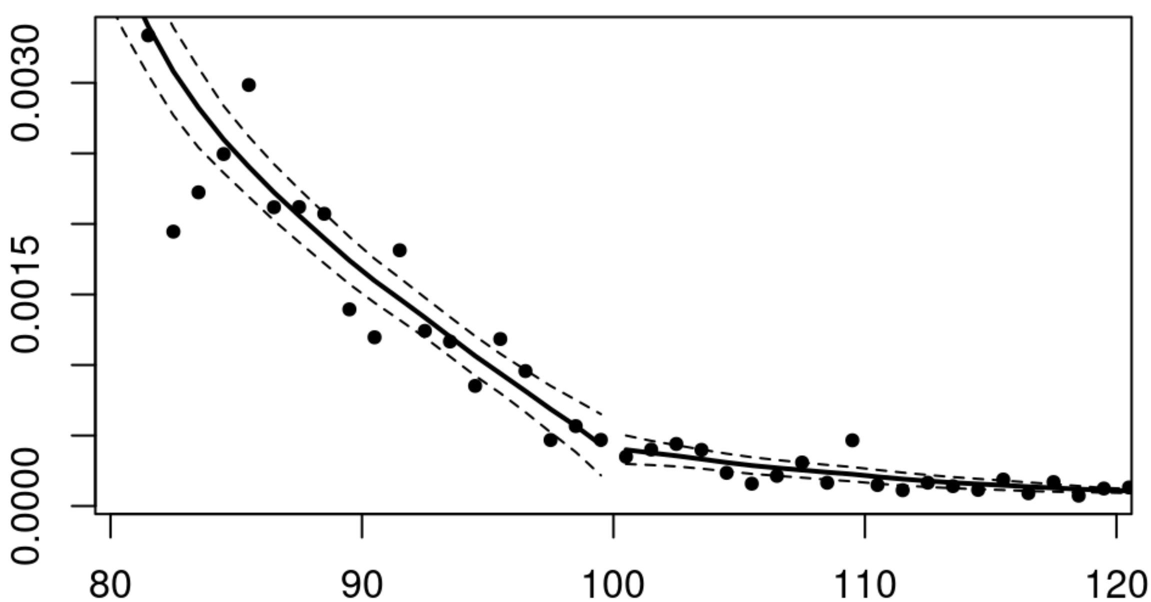
Since these air quality alerts are announced countywide, it is expected that there will be some response to this information. Some researchers have studied the relationship between air quality and human behavior by measuring the response to the issuance of air quality alerts (e.g. (T. Liu et al., 2018; M. Neidell, 2009; Saberian et al., 2017)).

I extend my analyses to measure any such effects. Since the alert is triggered by the AQI exceeding some threshold, I can estimate the causal effect of air quality on mobility by comparing individuals' mobility just above and below the threshold. In other words, if the alert is issued when the AQI exceeds 100 for instance, then I can compare mobility of individuals on days when the AQI is between 80 and 100, and days when the AQI is between 100 and 120. Essentially, what I seek to find is whether the issuance of alerts causes changes in individuals' mobility behaviors.

Using the regression discontinuity design (RDD) requires that some identifying assumptions are not violated. The first assumption is that there are no other points close to the cutoff where alerts are issued. For example, it must be established that counties (or whichever level of government that issues alerts in a place) are not issuing alerts when the AQI is 98 or 99. There must be a clean cutoff at 100. Another assumption is that alerts are only issued at the cutoff, and not some other cutoff of say 75 or 150. The last identifying assumption in the RDD is that the treatment has an effect only on the outcome variable and not on the non-outcome covariates. In the case of my analysis, this assumption would require the issuance of alerts to affect only peoples' mobility, and not weather conditions.

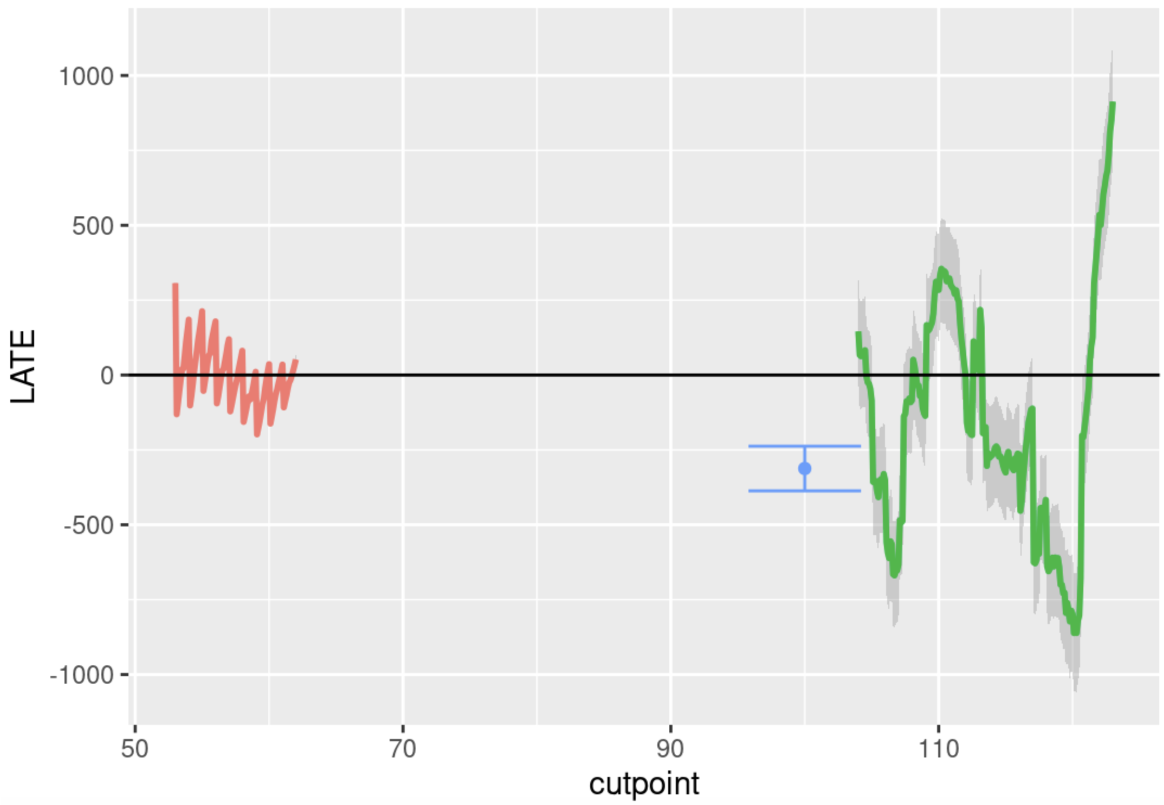
Following Thoemmes et al. (2017), I check for the validity of these assumptions in my model (Thoemmes et al., 2017). I first conduct the McCrary's test to check for any discontinuities in the AQI variable. The test for discontinuity came out significant with a z statistic of 6.0 and a p value of close to 0. This indicates the presence of some bunching around the cutoff as show in figure 5 and a violation of the first assumption.

Figure 5. Result from McCrary's Test for Discontinuity at the Cutoff



Next, I conduct placebo tests using the built-in function of the rddtools package. I show the case for the distance traveled variable (my analyses produced similar results for the two time variables) in figure 6 with effects at the cutoff displayed in blue, effects at lower cutoffs in red, and effects at higher cutoffs in green. The y-axis in the figure measures the local average treatment effect (LATE) at the cutoff, while the x-axis displays AQI. It can be seen in the figure that cutoffs below 100 had effects that were close to zero, suggesting no response at those cutoffs. Many of the cutoffs greater than 100 yield mostly larger sized effects with some in the opposite direction, and their confidence intervals largely do not cover zero. This shows significant responses (mostly negative) at cutoffs greater than 100 and is a violation of the second assumption.

Figure 6. Results from Placebo Tests Using RDD Package



Lastly, I conduct tests on how the independent variables in my model respond to the issuance of an alert at an AQI of 100 and find significant effects on all five tested variables. I show a summary of the results in table 22. These results show a violation of the third assumption.

Table 22. Results from nonoutcome covariates tests

	Temperature	**Precipitation**	**Wind Speed**	**Shortwave Ra diation**	**Humidity**
Daily AQI (LATE at the cutoff)	-8.099*** (-10.378)	0.6157*** (4.483)	1.3237*** (5.7320)	-12.705*** (-6.021)	2.298*** (3.409)
N	**7345**	**13907**	**5577**	**26044**	**16040**

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

Since my model violates all three assumptions of the regression discontinuity design, I conclude that the RDD is not the best model specification to use in my analysis. I came up with a couple of reasons as to why that is the case. First of all, counties and other reporting institutions could have the motivation to issue alerts when the AQI is very close to the cutoff because issuing more alerts could encourage policies geared towards managing air pollution. Secondly, my data covers the entire country, and different parts of the country have different cutoffs for the issuance of alerts. Perhaps the RDD would be more useful in research projects looking at areas with the same cutoff for the issuance of air quality alerts. Lastly, the tests on the independent variables may have come out positive because air quality is known to be correlated with weather variables, as shown in the first column of table 4.

Even though my model violates all three assumptions, I go ahead and estimate the relationship between air quality and mobility using the RDD to see if I would find any significant effects despite the violations. I use a regression discontinuity model as represented in equation (7):

$$\log\,(Mob_{ibt}) = \beta_1 alert_{bt} + \beta_2 AQ\_Centered_{bt} + \beta_3 X_{bt} + e_{ibt} \tag{7}$$

where $alert_{bt}$ is a variable indicating whether an alert was issued in CBG b on day t and $AQ\_Centered_{bt}$ is a variable obtained from recentering the AQ variable by subtracting the cutoff in each case. Using the same set of covariates as in my main model, I estimate the relationship between mobility and air quality alerts and present my results in table 23.

Table 23. Results from regression discontinuity model

Variable	Home Time (1)	Away time (2)	Distance (3)

Alert = 100	0.0067*** (2.79)	-0.0062. (-1.83)	-0.0063 (-1.59)
N = 207,265			
Alert = 150	0.0518 (1.44)	-0.0589 (-0.98)	0.1954** (2.70)
N = 9,051			

Significance codes: '' 0.05, '**' 0.01, '***' 0.001. T-values in parentheses*

All models include weather controls and CBG and date fixed effects

Table 23 shows a statistically significant increase in time spent at home of about 1% on days when an alert (code orange) is issued. There is a reduction in time spent away from home and distance traveled from home of comparable magnitudes, but these estimates are not significant at the 5% level. The estimates for time spent at and away from home on code red days are larger but lack statistical significance. Distance traveled from home on code red days appears to increase on average.

The estimates from table 23 are obtained from restricting the data to only AQI values that exceed or fall short of the threshold by 20. Reducing that bandwidth to a range of 10 more or less than the threshold gives estimates that are less precise as only a small number of observations fall within that range (see table K33 in the appendix). Using the full data to estimate the model produces statistically significant estimates that are higher in magnitude than those from the regressions using restricted data (see table K31 in the appendix). However, since the whole idea of regression discontinuity is about focusing on how observations occurring just above or below the cutoff are affected by the treatment, I refrain from drawing any conclusions based on the results from the regressions using the full data.

CHAPTER VIII: CONCLUSIONS

Results from the analyses done in this paper provide evidence that individuals adjust their daily movement patterns in response to changes in air quality. I find a reduction in mobility of between 1.4% and 4.5% on days of unhealthy air quality compared to days of good air quality. I also find that people respond to changes in the published category of air quality (color codes) rather than a unit change in the air quality index. I also find that the response to poor air quality increases with the severity of pollution reported such that there is progressively less observed mobility with worse air quality. Finally, I find that individuals living in CBGs with larger shares of minority populations are more likely to respond to air pollution by reducing their mobility. All these findings point toward the conclusion that people practice avoidance behavior in response to air pollution, and that the response to pollution is heterogeneous across different demographic groups.

What sets the findings in this paper apart from those in the literature is that I use observed data on mobility patterns for the whole United States and thus provide more reliable results since observed data is less likely to be plagued with issues like recall bias. Also, my findings are more generalizable than others in the literature since the study is not focused on a small subset of the population, or on effects of changes in air quality on a particular activity such as cycling.

It is important to note that this study has limitations. First, SafeGraph data does not report the location of devices throughout the day, but only reports a location when it notices pings from mobile apps at any time. As such, people who are less active users of such apps will have less representation in the data. If it the case that Safegraph records pings from mobile apps whose users have some common characteristics, then users of the dataset should be concerned about selection (sampling) bias. Though researchers have found compelling sampling rates across the United

States when compared to census population (Squire, 2019), it is important for users of the dataset to consider this potential for sample bias when interpreting results from their analyses.

A second limitation in my analyses is that I am not able to distinguish normal travel patterns from actual changes in travel behavior; what I may see as avoidance behavior in my results could actually be normal mobility behavior which coincides with a day of poor air quality. I included CBG and date fixed effects in my models to control for any such possibilities.

Third, using an AQI measure for a single pollutant (PM2.5) may not be enough to estimate the full effect on mobility. Though most studies in the literature focus on two pollutants at most, using a measure of AQI for all pollutants may provide better estimates of behavioral responses to changes in air quality. My future research will include finding the best measure of AQI that encompasses all pollutants reported by the EPA.

Other limitations of my analyses in this dissertation include the use of aggregated data and the potential for omitted variable bias. Although it makes it easier to analyze data from different sources at different levels (such as CBG or county level), aggregation presents the possibility that one may draw inaccurate inferences or false conclusions based on a relatively small number of outliers. Also, as I stated earlier, it is likely that there are other factors that I did not include in my model which have the potential to confound my findings. Although my models take these limitations into account, it would be better, given that data is available on all factors affecting air quality and mobility and that these data are available at more granular levels, to include all possible effects to avoid omitted variable bias, and to avoid aggregation altogether. However, since it is usually not possible to obtain the perfect dataset, researchers use econometric methods to control for such limitations as I have done. Thus, I am comfortable drawing conclusions from my results.

Future research could explore individual-level mobility data and investigate additional factors that may influence the relationship between air quality and mobility. Nonetheless, the findings in this paper have policy implications as evidence of avoidance behavior suggests that individuals care enough about the effects of air pollution to avoid exposure to pollutants. Governments can thus take this evidence into account when making the decisions on the most cost-effective ways to manage air pollution: abatement to reduce the amount of pollution (which has increasing marginal costs) or encouraging avoidance behavior to reduce exposure to pollutants while finding cheaper ways to reduce pollution. Evidence on heterogeneity in the response to air pollution should prompt policymakers to be more attentive to certain demographic groups (minority populations, children and the elderly) when enacting policies targeted at avoidance behavior since these groups are more vulnerable to the effects of pollution.

REFERENCES

Action Days | AirNow.gov. (2023). https://www.airnow.gov/aqi/action-days/

American Lung Association. (2021). *American Lung Association State of the Air 2021*. https://www.lung.org/research/sota

Andersen, M. (2020). Early Evidence on Social Distancing in Response to COVID-19 in the United States. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3569368

Andersen, M., Maclean, J. C., Pesko, M. F., & Simon, K. (2023). Does paid sick leave encourage staying at home? Evidence from the United States during a pandemic. *Health Economics*, *32*(6), 1256–1283. https://doi.org/10.1002/hec.4665

Anderson, H. R. (2009). Air pollution and mortality: A history. *Atmospheric Environment*, *43*(1), 142–152. https://doi.org/10.1016/j.atmosenv.2008.09.026

Anderson, M. L., Hyun, M., & Lee, J. (2022). *Bounds, Benefits, and Bad Air: Welfare Impacts of Pollution Alerts* (Working Paper 29637; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w29637

Barwick, P. J., Li, S., Lin, L., & Zou, E. (2019). *From Fog to Smog: The Value of Pollution Information* (Working Paper 26541; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w26541

Borroni, E., Pesatori, A. C., Bollati, V., Buoli, M., & Carugno, M. (2022). Air pollution exposure and depression: A comprehensive updated systematic review and meta-analysis. *Environmental Pollution*, *292*, 118245. https://doi.org/10.1016/j.envpol.2021.118245

Bound, J., Brown, C., & Mathiowetz, N. (2001). Chapter 59—Measurement Error in Survey Data. In J. J. Heckman & E. Leamer (Eds.), *Handbook of Econometrics* (Vol. 5, pp. 3705–3843). Elsevier. https://doi.org/10.1016/S1573-4412(01)05012-7

Braithwaite, I., Zhang, S., Kirkbride, J. B., Osborn, D. P. J., & Hayes, J. F. (2019). Air Pollution (Particulate Matter) Exposure and Associations with Depression, Anxiety, Bipolar, Psychosis and Suicide Risk: A Systematic Review and Meta-Analysis. *Environmental Health Perspectives*, *127*(12), 126002. https://doi.org/10.1289/EHP4595

Chang, T. Y., Graff Zivin, J., Gross, T., & Neidell, M. (2019). The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China. *American Economic Journal: Applied Economics*, *11*(1), 151–172. https://doi.org/10.1257/app.20160436

Chen, F., & Chen, Z. (2021). Cost of economic growth: Air pollution and health expenditure. *Science of The Total Environment*, *755*, 142543. https://doi.org/10.1016/j.scitotenv.2020.142543

Chen, S., & Zhang, D. (2021). Impact of air pollution on labor productivity: Evidence from prison factory data. *China Economic Quarterly International*, *1*(2), 148–159. https://doi.org/10.1016/j.ceqi.2021.04.004

Clayton, S. (2021). Climate Change and Mental Health. *Current Environmental Health Reports*, *8*(1), 1–6. https://doi.org/10.1007/s40572-020-00303-3

Cummings, M. (2015, December 15). *Hospital prices show 'mind-boggling' variation across U.S. driving up health care costs*. YaleNews. https://news.yale.edu/2015/12/15/hospital-prices-show-mind-boggling-variation-across-us-driving-health-care-costs

Delucchi, M. A., Murphy, J. J., & McCubbin, D. R. (2002). The health and visibility cost of air pollution: A comparison of estimation methods. *Journal of Environmental Management*, *64*(2), 139–152. https://doi.org/10.1006/jema.2001.0515

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*. 105.

Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Dominici, F., & Schwartz, J. D. (2017). Air Pollution and Mortality in the Medicare Population. *New England Journal of Medicine*, *376*(26), 2513–2522. https://doi.org/10.1056/NEJMoa1702747

Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., & Speizer, F. E. (1993). An Association between Air Pollution and Mortality in Six U.S. Cities. *New England Journal of Medicine*, *329*(24), 1753–1759. https://doi.org/10.1056/NEJM199312093292401

Domingo, J. L., & Rovira, J. (2020). Effects of air pollutants on the transmission and severity of respiratory viral infections. *Environmental Research*, *187*, 109650. https://doi.org/10.1016/j.envres.2020.109650

Downey, L., & Hawkins, B. (2008). Race, Income, and Environmental Inequality in the United States. *Sociological Perspectives : SP : Official Publication of the Pacific Sociological Association*, *51*(4), 759–781. https://doi.org/10.1525/sop.2008.51.4.759

Environmental Defense Fund. (2023). *Health effects of air pollution*. Environmental Defense Fund. https://www.edf.org/health/effects-of-air-pollution

Environmental Health (NSW). (2013). *Who is affected by air pollution*. NSW Government. https://www.health.nsw.gov.au/environment/air/Pages/who-is-affected.aspx

Fairburn, J., Schüle, S. A., Dreger, S., Karla Hilz, L., & Bolte, G. (2019). Social Inequalities in Exposure to Ambient Air Pollution: A Systematic Review in the WHO European Region. *International Journal of Environmental Research and Public Health*, *16*(17), Article 17. https://doi.org/10.3390/ijerph16173127

Freire, C., Ramos, R., Puertas, R., Lopez-Espinosa, M.-J., Julvez, J., Aguilera, I., Cruz, F., Fernandez, M.-F., Sunyer, J., & Olea, N. (2010). Association of traffic-related air pollution with cognitive development in children. *Journal of Epidemiology & Community Health*, *64*(3), 223–228. https://doi.org/10.1136/jech.2008.084574

Gatto, N. M., Henderson, V. W., Hodis, H. N., St. John, J. A., Lurmann, F., Chen, J.-C., & Mack, W. J. (2014). Components of air pollution and cognitive function in middle-aged and older adults in Los Angeles. *NeuroToxicology*, *40*, 1–7. https://doi.org/10.1016/j.neuro.2013.09.004

Gliadkovskaya, A. (2021, October 8). *Hospital prices vary widely, often higher with insurance than cash, The New York Times finds*. Fierce Healthcare. https://www.fiercehealthcare.com/hospitals/hospital-prices-widely-vary-by-payer-often-higher-than-cash-price

Glinianaia, S. V., Rankin, J., Bell, R., Pless-Mulloli, T., & Howel, D. (2004). Particulate Air Pollution and Fetal Health: A Systematic Review of the Epidemiologic Evidence. *Epidemiology*, *15*(1), 36–45.

Gouveia, N., & Fletcher, T. (2000). Time series analysis of air pollution and mortality: Effects by cause, age and socioeconomic status. *Journal of Epidemiology & Community Health*, *54*(10), 750–755. https://doi.org/10.1136/jech.54.10.750

Graff Zivin, J., & Neidell, M. (2013). Environment, Health, and Human Capital. *Journal of Economic Literature*, *51*(3), 689–730. https://doi.org/10.1257/jel.51.3.689

He, Z., Zhao, Y., & Zheng, L. (2023). How does air pollution affect the stock market performance? Evidence from China. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-023-27446-8

Heutel, G., & Ruhm, C. J. (2016). Air Pollution and Procyclical Mortality. *Journal of the Association of Environmental and Resource Economists*, *3*(3), 667–706. https://doi.org/10.1086/686251

Heyes, A., Neidell, M., & Saberian, S. (2016). *The Effect of Air Pollution on Investor Behavior: Evidence from the S&P 500* (Working Paper 22753). National Bureau of Economic Research. https://doi.org/10.3386/w22753

Hyslop, N. P. (2009). Impaired visibility: The air pollution people see. *Atmospheric Environment*, *43*(1), 182–195. https://doi.org/10.1016/j.atmosenv.2008.09.067

Janke, K. (2014). Air Pollution, Avoidance Behaviour and Children's Respiratory Health: Evidence from England. *Journal of Health Economics*, *38*, 23–42.

Jiang, X.-Q., Mei, X.-D., & Feng, D. (2016). Air pollution and chronic airway diseases: What should people know and do? *Journal of Thoracic Disease*, *8*(1). https://doi.org/10.3978/j.issn.2072-1439.2015.11.50

Jones, B. A. (2017). Are we underestimating the economic costs of wildfire smoke? An investigation using the life satisfaction approach. *Journal of Forest Economics*, *27*, 80–90. https://doi.org/10.1016/j.jfe.2017.03.004

Kampa, M., & Castanas, E. (2008). Human health effects of air pollution. *Environmental Pollution*, *151*(2), 362–367. https://doi.org/10.1016/j.envpol.2007.06.012

Kim, M. J. (2021). Air Pollution, Health, and Avoidance Behavior: Evidence from South Korea. *Environmental and Resource Economics*, *79*(1), 63–91.

Kim, S.-Y., Kim, E., & Kim, W. J. (2020). Health Effects of Ozone on Respiratory Diseases. *Tuberculosis and Respiratory Diseases*, *83*(Supple 1), S6–S11. https://doi.org/10.4046/trd.2020.0154

Kurani, N., Rae, M., Pollitz, K., Amin, K., & Cox, C. (2021). Price transparency and variation in U.S. health services. *Peterson-KFF Health System Tracker*. https://www.healthsystemtracker.org/brief/price-transparency-and-variation-in-u-s-health-services/

Kurt, O. K., Zhang, J., & Pinkerton, K. E. (2016). Pulmonary Health Effects of Air Pollution. *Current Opinion in Pulmonary Medicine*, *22*(2), 138–143. https://doi.org/10.1097/MCP.0000000000000248

Lagasse, J. (2016). *Prices for healthcare vary widely across states, even nearby cities, Health Care Cost Institute says*. Healthcare Finance News. https://www.healthcarefinancenews.com/news/prices-healthcare-vary-widely-across-states-even-nearby-cities-health-care-cost-institute-says

Laumbach, R. J., & Kipen, H. M. (2012). Respiratory health effects of air pollution: Update on biomass smoke and traffic pollution. *Journal of Allergy and Clinical Immunology*, *129*(1), 3–11. https://doi.org/10.1016/j.jaci.2011.11.021

Lieberman, H. R., Agarwal, S., Caldwell, J. A., & Fulgoni, V. L., III. (2020). Demographics, sleep, and daily patterns of caffeine intake of shift workers in a nationally representative sample of the US adult population. *Sleep*, *43*(3), zsz240. https://doi.org/10.1093/sleep/zsz240

Liu, M., Xue, X., Zhou, B., Zhang, Y., Sun, B., Chen, J., & Li, X. (2019). Population susceptibility differences and effects of air pollution on cardiovascular mortality: Epidemiological evidence from a time-series study. *Environmental Science and Pollution Research*, *26*(16), 15943–15952. https://doi.org/10.1007/s11356-019-04960-2

Liu, T., He, G., & Lau, A. (2018). Avoidance Behavior against Air Pollution: Evidence from Online Search Indices for Anti-PM_{2.5} Masks and Air Filters in Chinese Cities. *Environmental Economics and Policy Studies*, *20*(2), 325–363.

Lu, J. G. (2020). Air pollution: A systematic review of its psychological, economic, and social effects. *Current Opinion in Psychology*, *32*, 52–65. https://doi.org/10.1016/j.copsyc.2019.06.024

Makri, A., & Stilianakis, N. I. (2008). Vulnerability to air pollution health effects. *International Journal of Hygiene and Environmental Health*, *211*(3), 326–336. https://doi.org/10.1016/j.ijheh.2007.06.005

Manson, S., Schroeder, D., Riper, V., Ruggles, S., & Kugler, T. (2022). *IPUMS National Historical Geographic Information System Minneapolis, MN: IPUMS. 2022* (17.0) [Dataset]. https://doi.org/10.18128/D050.V17.0

McCarthy, N. (2020). *This is the global economic cost of air pollution*. World Economic Forum. https://www.weforum.org/agenda/2020/02/the-economic-burden-of-air-pollution/

Memoli, J. S. W. (2020, August 6). *African-americans-and-respiratory-illness*. https://www.medstarhealth.org/blog/african-americans-and-respiratory-illness

Moretti, E., & Neidell, M. (2011). Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles. *Journal of Human Resources*, *46*(1), 154–175.

Neidell, M. (2009). Information, Avoidance Behavior, and Health: The Effect of Ozone on Asthma Hospitalizations. *Journal of Human Resources*, *44*(2), 450–478.

Neidell, M. J. (2004). Air Pollution, Health, and Socio-economic Status: The Effect of Outdoor Air Quality on Childhood Asthma. *Journal of Health Economics*, *23*(6), 1209–1236.

Neidell, M., & Pestel, N. (2023). Air pollution and worker productivity. *IZA World of Labor*. https://doi.org/10.15185/izawol.363

Persico, C., & Marcotte, D. E. (2022a). Air Quality and Suicide. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4114633

Persico, C., & Marcotte, D. E. (2022b). *Air Quality and Suicide* (SSRN Scholarly Paper 4114633). https://doi.org/10.2139/ssrn.4114633

Pope, 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., & Burnett, R. T. (2019). Mortality Risk and Fine Particulate Air Pollution in a Large, Representative Cohort of U.S. Adults. *Environmental Health Perspectives*, *127*(7), 077007. https://doi.org/10.1289/EHP4438

Rashidi, R., Khaniabadi, Y. O., Sicard, P., De Marco, A., & Anbari, K. (2023). Ambient PM2.5 and O3 pollution and health impacts in Iranian megacity. *Stochastic Environmental Research and Risk Assessment*, *37*(1), 175–184. https://doi.org/10.1007/s00477-022-02286-z

Saberian, S., Heyes, A., & Rivers, N. (2017). Alerts work! Air quality warnings and cycling. *Resource and Energy Economics*, *49*, 165–185. https://doi.org/10.1016/j.reseneeco.2017.05.004

Santana, J. C. C., Miranda, A. C., Yamamura, C. L. K., Silva Filho, S. C. da, Tambourgi, E. B., Lee Ho, L., & Berssaneti, F. T. (2020). Effects of Air Pollution on Human Health and

Costs: Current Situation in São Paulo, Brazil. *Sustainability*, *12*(12), Article 12. https://doi.org/10.3390/su12124875

Social Distancing Metrics | SafeGraph Docs. (2020). SafeGraph. https://docs.safegraph.com/changelog/april-2020-release-notes

Squire, R. F. (2019, October). *Google Colaboratory*. https://colab.research.google.com/drive/1u15afRytJMsizySFqA2EPlXSh3KTmNTQ#offline=true&sandboxMode=true

Stone, V. (2000). Environmental Air Pollution. *American Journal of Respiratory and Critical Care Medicine*, *162*(supplement_1), S44–S47. https://doi.org/10.1164/ajrccm.162.supplement_1.maic-12

Sun, C., Yuan, X., & Xu, M. (2016). The public perceptions and willingness to pay: From the perspective of the smog crisis in China. *Journal of Cleaner Production*, *112*, 1635–1644. https://doi.org/10.1016/j.jclepro.2015.04.121

The NHLBI Working Group. (1995). Respiratory Diseases Disproportionately Affecting Minorities. *CHEST*, *108*(5), 1380–1392. https://doi.org/10.1378/chest.108.5.1380

Thind, M. P. S., 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. https://doi.org/10.1021/acs.est.9b02527

Thoemmes, F., Liao, W., & Jin, Z. (2017). The Analysis of the Regression-Discontinuity Design in R. *Journal of Educational and Behavioral Statistics*, *42*(3), 341–360. https://doi.org/10.3102/1076998616680587

US EPA. (2011). *The Benefits and Costs of the Clean Air Act from 1990 to 2020: Summary Report* (p. 35).

US EPA, O. (2016, May 4). *Air Quality—National Summary* [Data and Tools]. https://www.epa.gov/air-trends/air-quality-national-summary

Ward, A. L. S., & Beatty, T. K. M. (2016). Who Responds to Air Quality Alerts? *Environmental and Resource Economics*, *65*(2), 487–511. https://doi.org/10.1007/s10640-015-9915-z

Wooldridge, J. M. (2015). Control Function Methods in Applied Econometrics. *Journal of Human Resources*, *50*(2), 420–445. https://doi.org/10.3368/jhr.50.2.420

World Health Organization. (2023). *Air pollution*. https://www.who.int/health-topics/air-pollution

Xie, H., Cao, Y., Li, J., Lyu, Y., Roberts, N., & Jia, Z. (2023). Affective disorder and brain alterations in children and adolescents exposed to outdoor air pollution. *Journal of Affective Disorders*, *331*, 413–424. https://doi.org/10.1016/j.jad.2023.03.082

Yan, L., Duarte, F., Wang, D., Zheng, S., & Ratti, C. (2019). Exploring the effect of air pollution on social activity in China using geotagged social media check-in data. *Cities*, *91*, 116–125. https://doi.org/10.1016/j.cities.2018.11.011

Yoo, G. (2021). Real-time information on air pollution and avoidance behavior: Evidence from South Korea. *Population and Environment*, *42*(3), 406–424. https://doi.org/10.1007/s11111-020-00368-0

Zhang, J., & Mu, Q. (2018). Air Pollution and Defensive Expenditures: Evidence from Particulate-Filtering Facemasks. *Journal of Environmental Economics and Management*, *92*, 517–536.

Zhang, W., Bansback, N., & Anis, A. H. (2011). Measuring and valuing productivity loss due to poor health: A critical review. *Social Science & Medicine*, *72*(2), 185–192. https://doi.org/10.1016/j.socscimed.2010.10.026

Zhang, X., Zhang, X., & Chen, X. (2017). Happiness in the air: How does a dirty sky affect mental health and subjective well-being? *Journal of Environmental Economics and Management*, *85*, 81–94. https://doi.org/10.1016/j.jeem.2017.04.001

APPENDIX A: MEASUREMENT ERROR IN THE DEPENDENT VARIABLE

Measurement error in a regression model refers to the case when the true value of a variable in the model cannot be observed or measured correctly. The difference between the actual value that is not observed and the value that is observed is the error.

Measurement error in SafeGraph data may result because of the improper allocation of a device's home. Since all the mobility variables reference the device's home, they are likely to contain some amount of error.

The presence of measurement error in an independent variable in a regression model creates an endogeneity bias because the error in that variable becomes part of the error term in the regression equation (Pischke, 2007). However, if the measurement error is in the dependent variable, following the assumptions that define the classical errors-in-variables model, then the error is not correlated with the independent variables (or the error term in the regression), and thus the model can be estimated consistently by OLS.

Consider, for example, a simple regression equation with an outcome variable y such that

$$y = \beta x + \varepsilon \qquad (1).$$

If there is measurement error in the dependent variable such that

$$\tilde{y} = y + v \qquad (2),$$

where v is the measurement error, then substituting (1) in (2), the regression equation becomes

$$\tilde{y} = \beta x + \varepsilon + v \qquad (3).$$

Since v is uncorrelated with x, β can be estimated consistently in this example. It must be noted though that the estimates from this model will be less precise because of larger standard errors.

Figure B7. Regional distribution of time spent at home

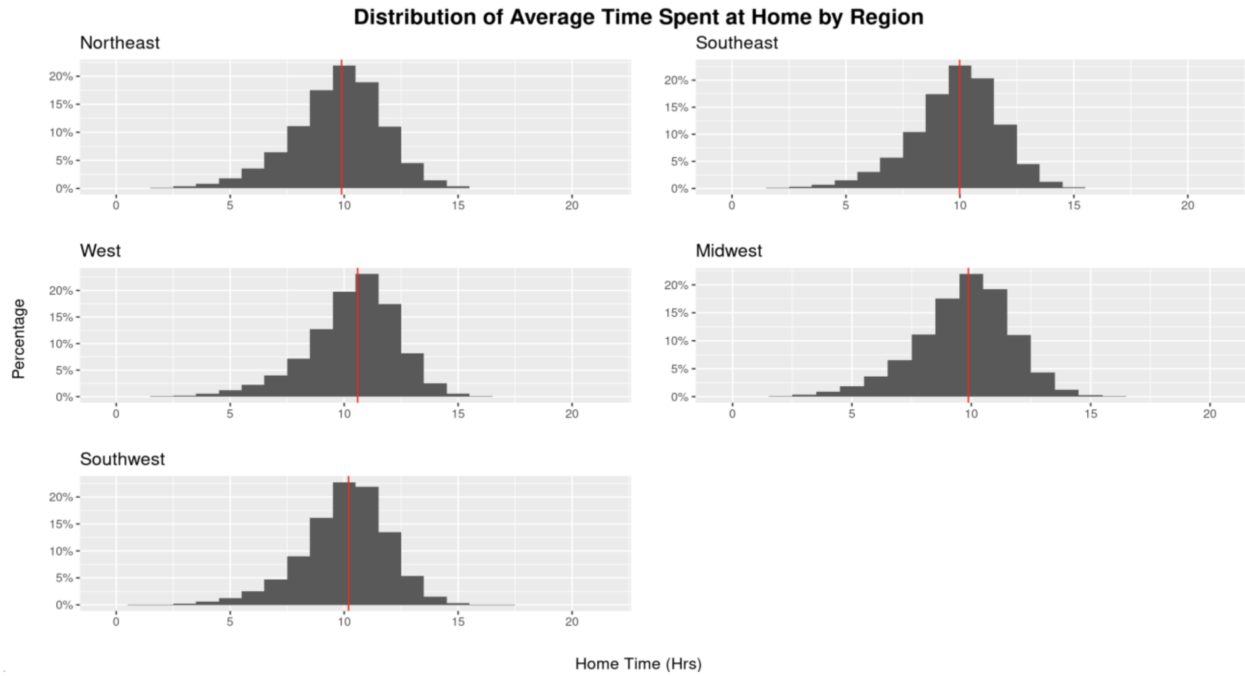


Figure B8. Regional distribution of time spent away from home

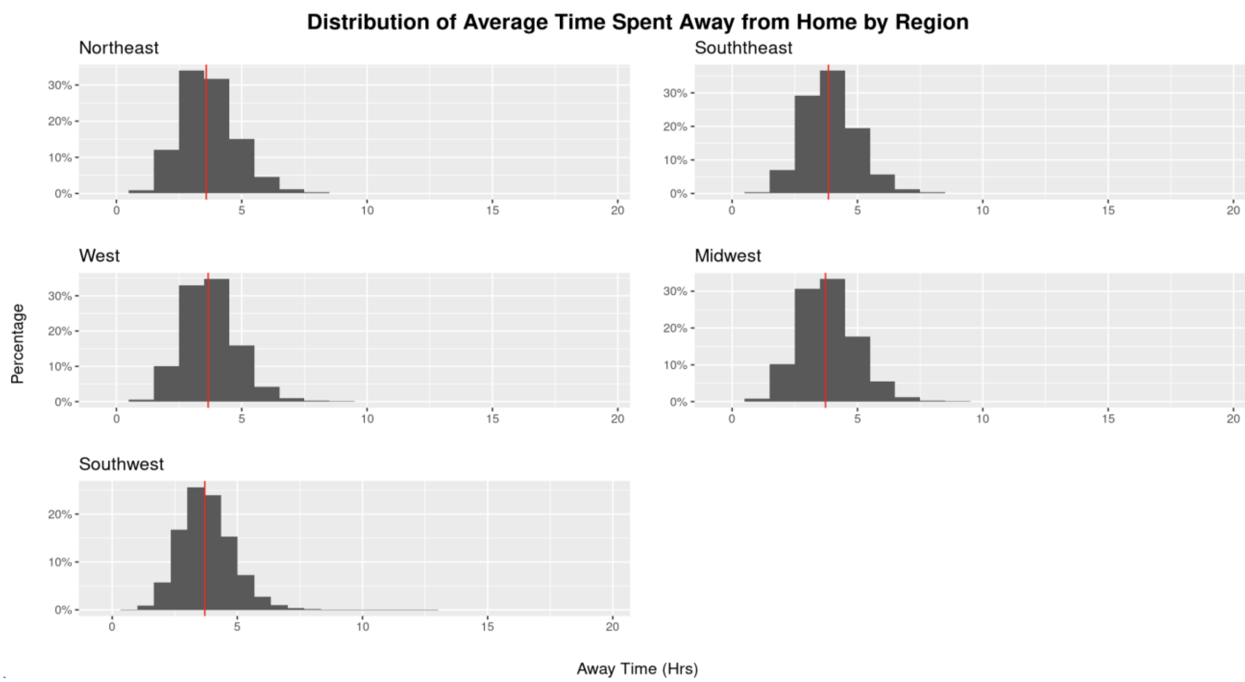
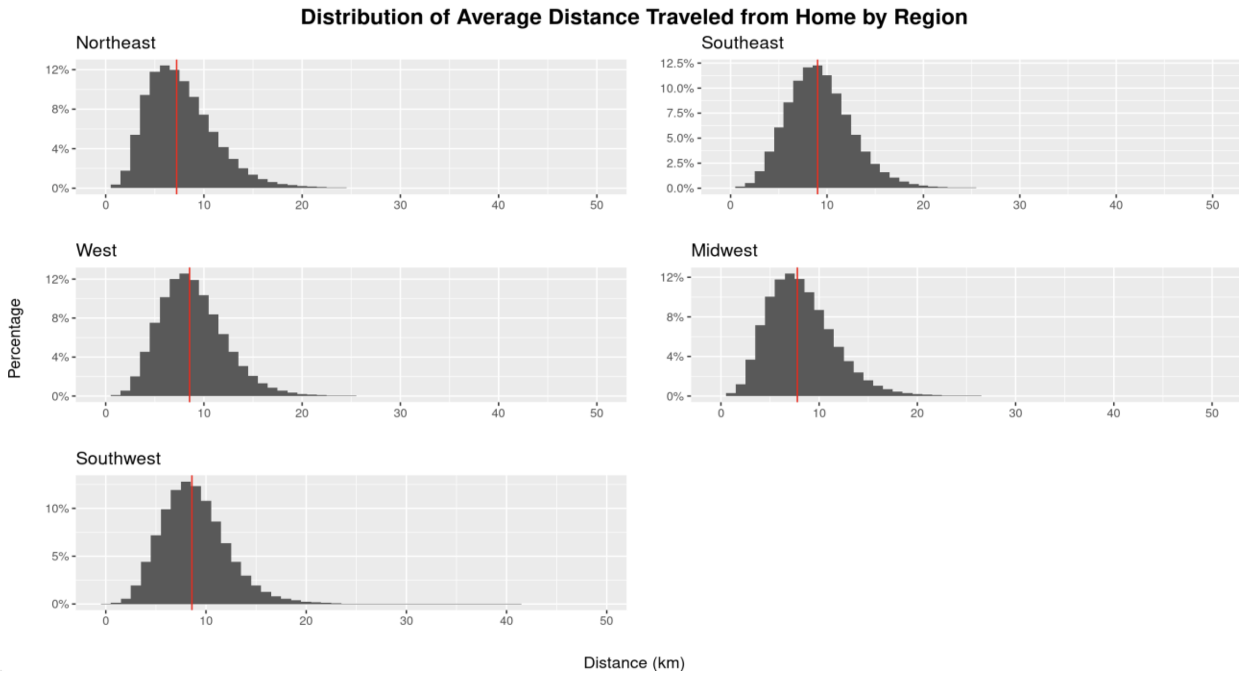


Figure B9. Regional distribution of distance traveled from home



APPENDIX C: SEASONAL DISTRIBUTION OF MOBILITY VARIABLES

Figure C10. Seasonal distribution of time spent at home

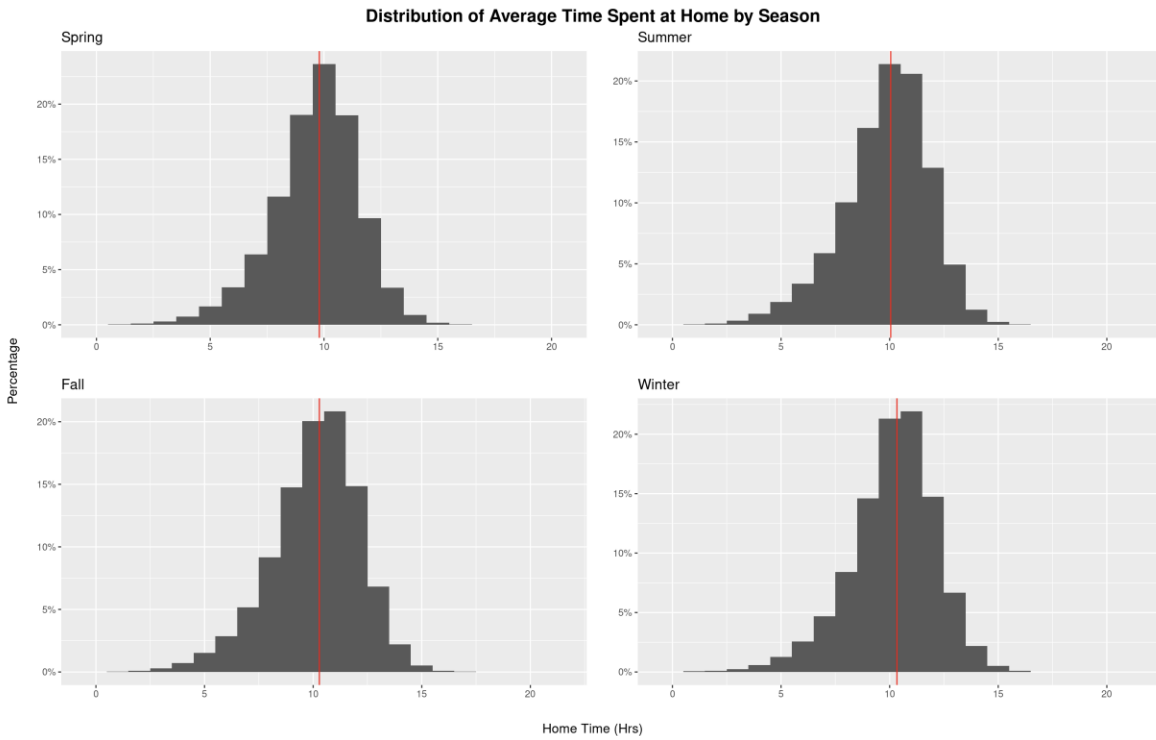


Figure C11. Seasonal distribution of time spent away from home

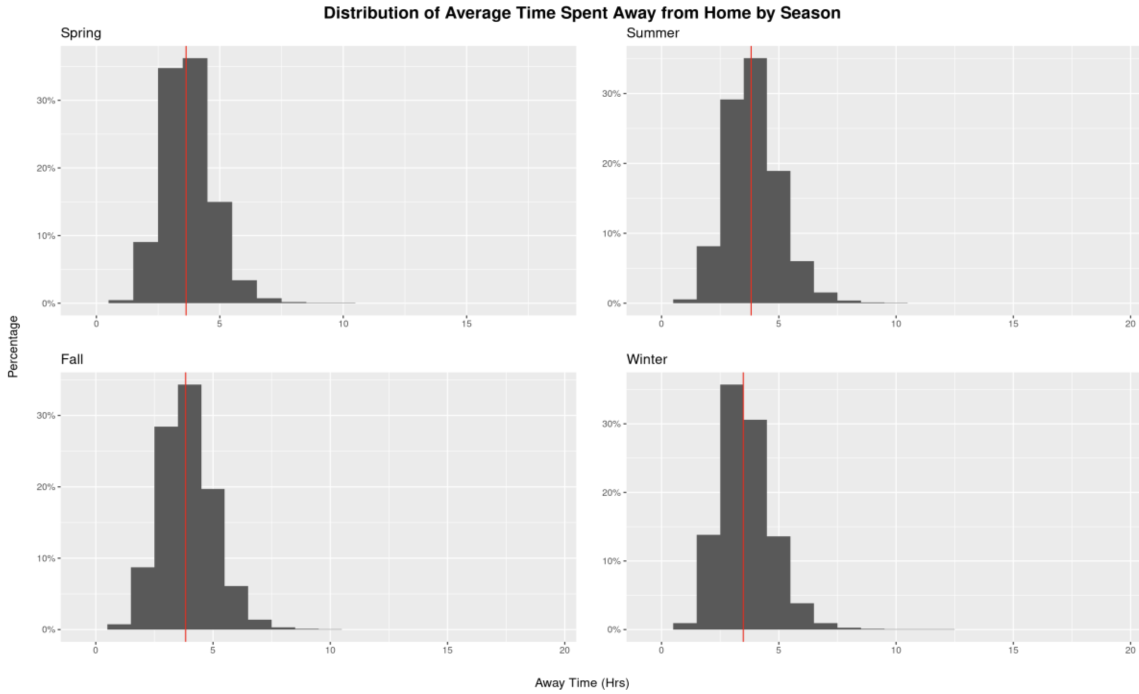
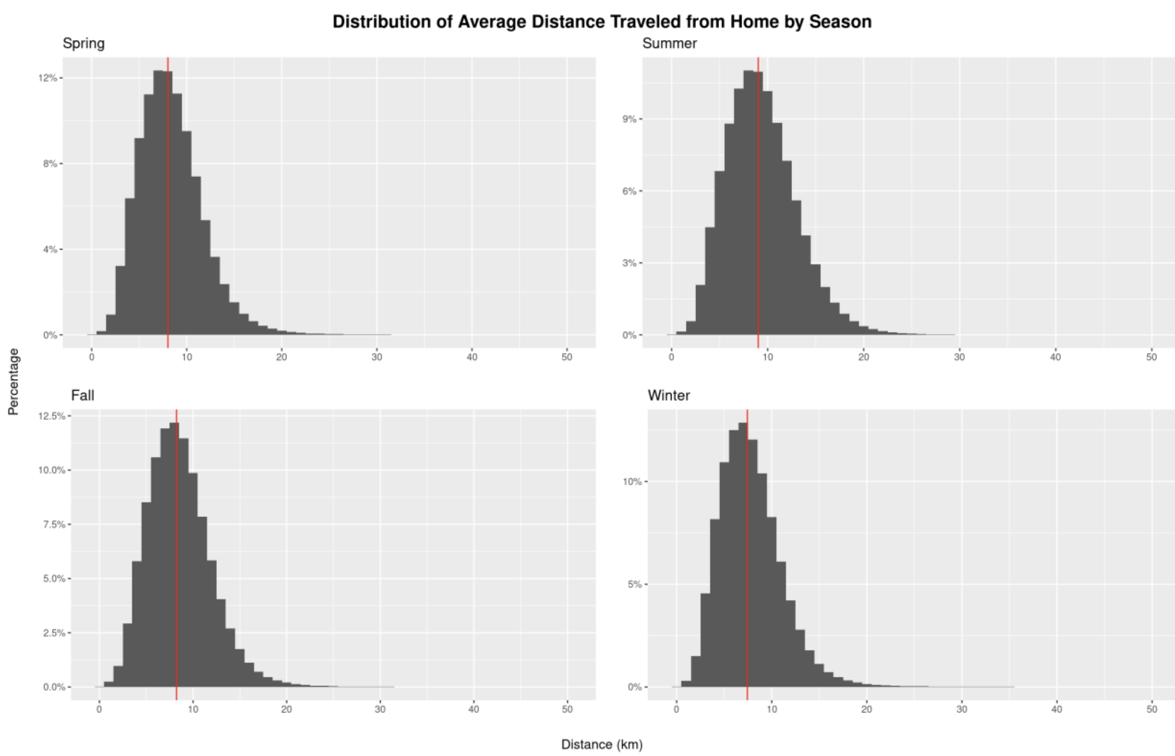


Figure C12. Seasonal distribution of distance traveled from home



APPENDIX D: DISTRIBUTION OF AQI SCALE

Figure D13. Distribution of AQI Scale

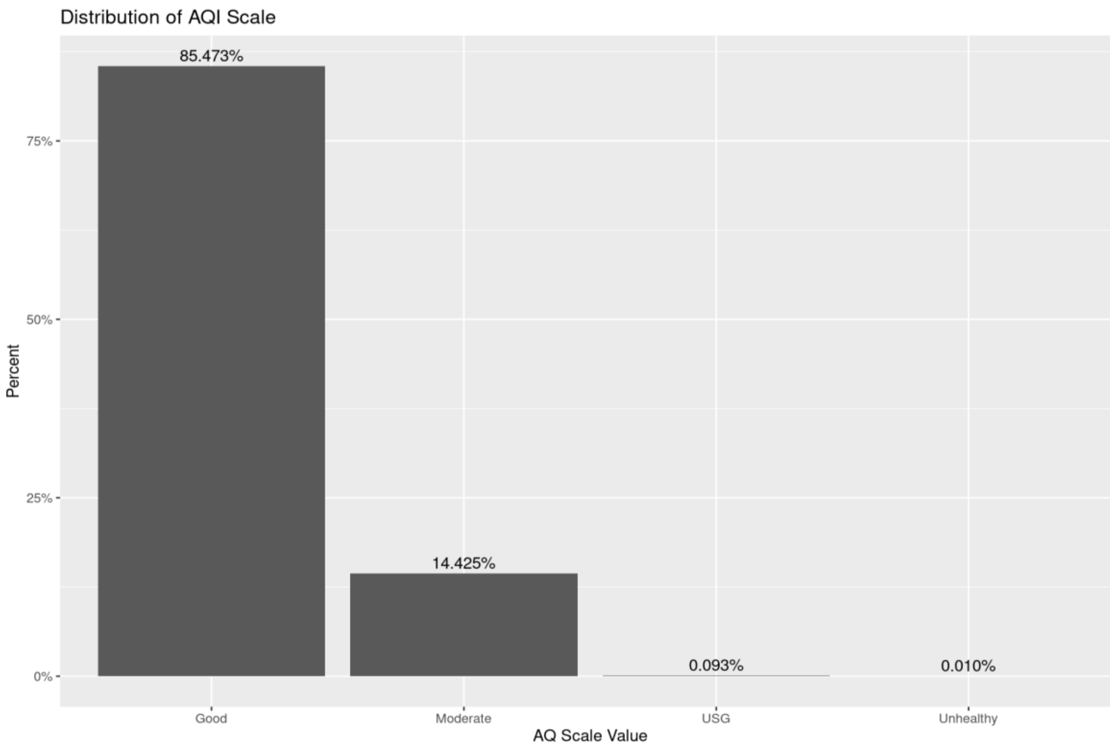


Figure D14. Regional distribution of AQI scale

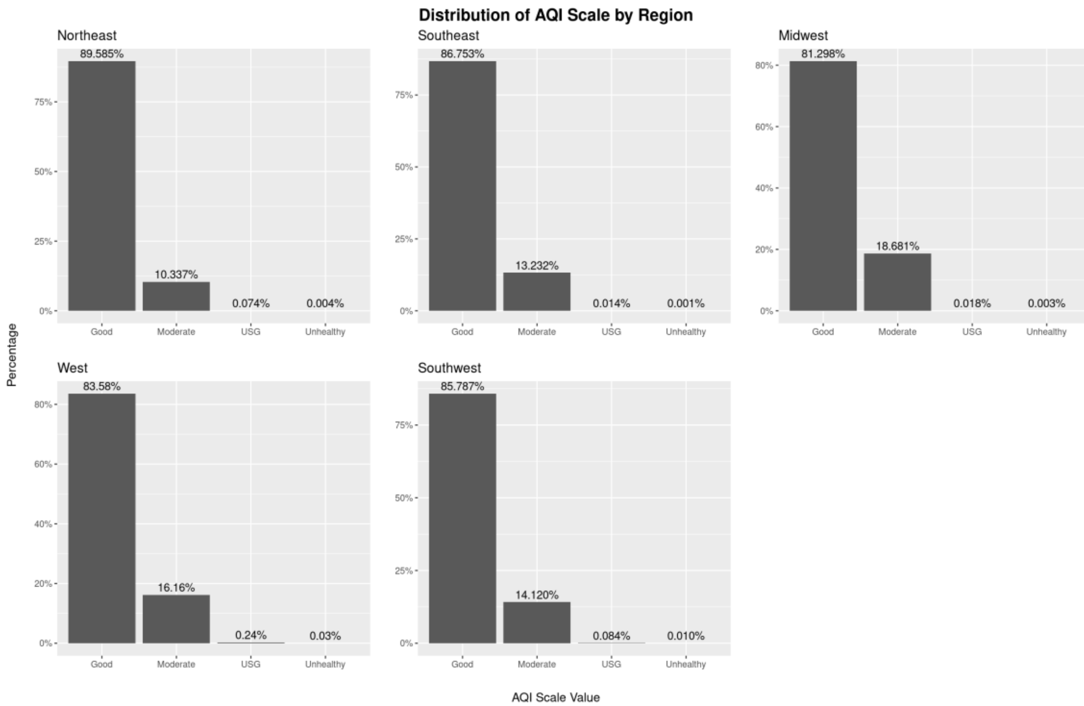


Figure D15. Seasonal distribution of AQI scale

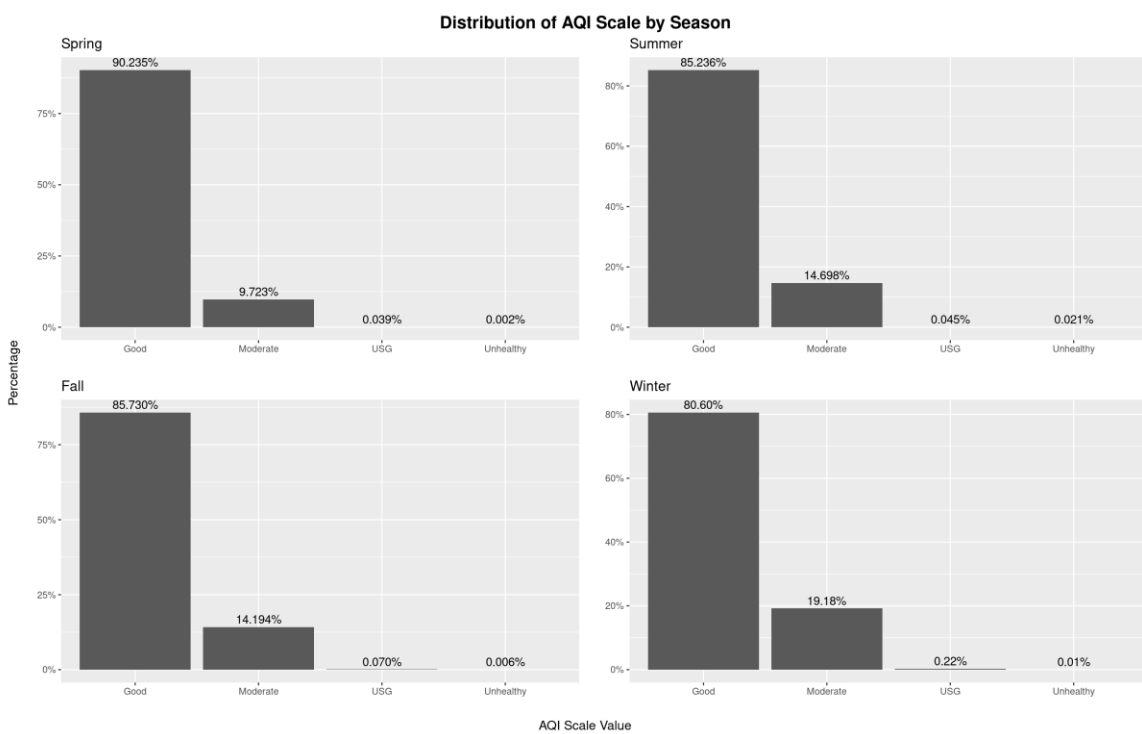


Figure E16. Distribution of air quality by month

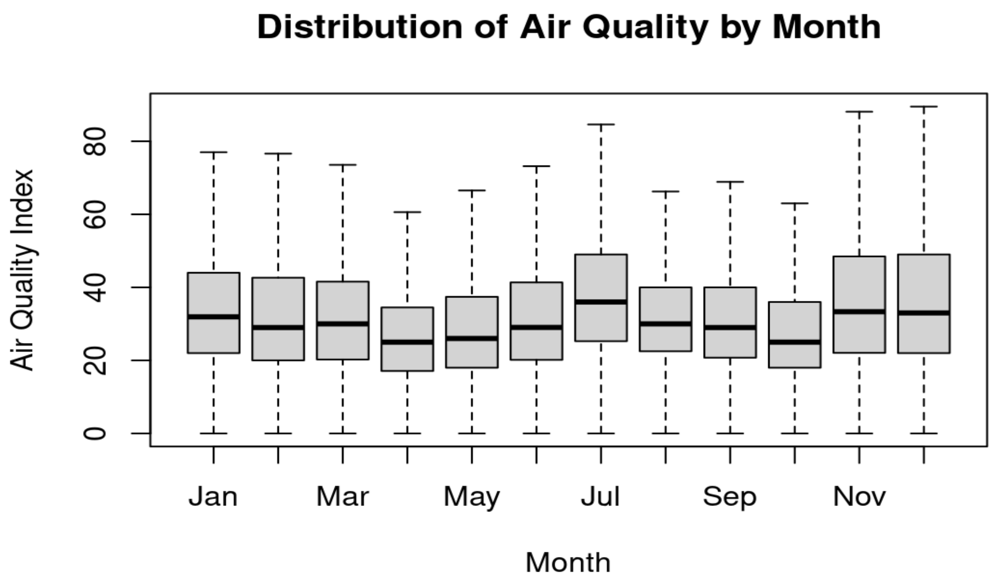
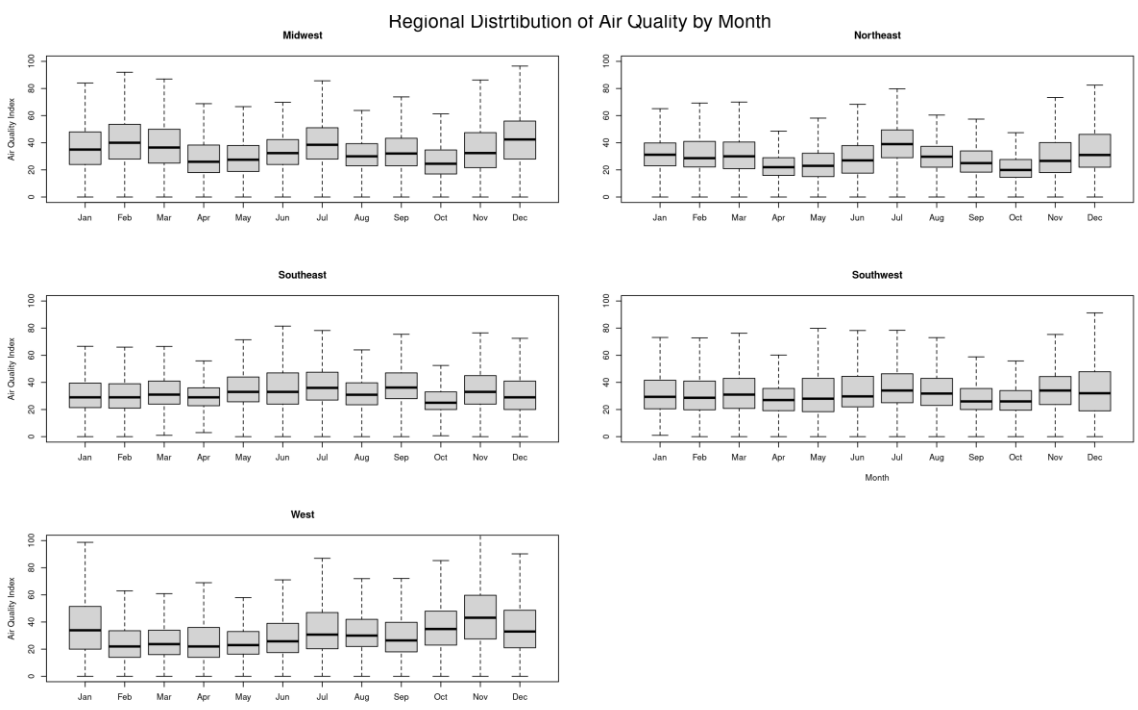


Figure E17. Regional distribution of air quality by month



APPENDIX F: BREAKDOWN OF SUMMARY STATISTICS BY REGION

Table F24. Summary statistics (Northeast)

Variable	Mean	Standard Deviation	25th Percentile	75th Percentile
Home Time (hrs)	9.71	2.02	8.55	11.05
Away Time (hrs)	3.70	1.15	2.91	4.36
Distance from Home (km)	7.78	3.62	5.19	9.71
Daily Air Quality Index	29.81	14.74	19.00	38.07
Temperature (°F)	53.59	17.95	38.84	69.35
Precipitation (in.)	3.65	8.25	0.00	2.80
Wind Speed (mph)	10.14	4.54	6.93	12.53
Shortwave Radiation (w/m^2)	183.02	92.33	105.60	262.70
Humidity (%)	64.91	12.66	57.40	73.55
N = 9,880,233				

Table F25. Summary statistics (Southeast)

Variable	Mean	Standard Deviation	25th Percentile	75th Percentile
Home Time (hrs)	9.81	1.92	8.72	11.11
Away Time (hrs)	3.93	1.11	3.19	4.57
Distance from Home (km)	9.35	3.58	6.92	11.33
Daily Air Quality Index	33.12	13.59	23.00	42.00
Temperature (°F)	67.38	15.24	56.21	80.24
Precipitation (in.)	4.26	10.91	0.00	2.60
Wind Speed (mph)	8.28	3.42	5.82	10.29
Shortwave Radiation (w/m^2)	213.38	84.11	146.50	279.40
Humidity (%)	68.90	10.52	63.25	75.65
N = 6,885,635				

Table F26. Summary statistics (Midwest)

Variable	Mean	Standard Deviation	25th Percentile	75th Percentile
Home Time (hrs)	9.69	1.99	8.54	11.03
Away Time (hrs)	3.82	1.16	3.02	4.50
Distance from Home (km)	8.24	3.58	5.73	10.18
Daily Air Quality Index	34.61	15.77	34.61	45.61
Temperature (°F)	50.16	20.47	56.21	68.27
Precipitation (in.)	3.32	7.96	0.00	2.50
Wind Speed (mph)	9.92	4.01	6.93	12.08
Shortwave Radiation (w/m^2)	173.70	87.63	93.50	248.50
Humidity (%)	69.72	11.43	62.45	77.45
N = 7,841,758				

Table F27. Summary statistics (West)

Variable	Mean	Standard Deviation	25th Percentile	75th Percentile
Home Time (hrs)	10.37	1.97	9.29	11.69
Away Time (hrs)	3.75	1.08	3.01	4.38
Distance from Home (km)	8.90	3.42	6.50	10.85
Daily Air Quality Index	31.82	17.67	18.37	43.00
Temperature (°F)	58.92	14.12	50.27	68.72
Precipitation (in.)	1.63	5.62	0.00	0.00
Wind Speed (mph)	7.14	3.36	4.70	8.72
Shortwave Radiation (w/m^2)	208.51	94.73	127.10	297.30
Humidity (%)	60.75	18.07	49.65	73.80
N = 9,418,169				

Table F28. Summary statistics (Southwest)

Variable	Mean	Standard Deviation	25th Percentile	75th Percentile
Home Time (hrs)	10.02	1.90	8.97	11.28
Away Time (hrs)	3.79	1.08	3.05	4.41
Distance from Home (km)	8.94	3.43	6.58	10.83
Daily Air Quality Index	32.14	15.18	20.97	42.00
Temperature (°F)	67.59	16.41	54.95	82.13
Precipitation (in.)	2.12	8.53	0.00	0.00
Wind Speed (mph)	8.63	3.70	6.04	10.74
Shortwave Radiation (w/m^2)	217.94	81.54	148.30	288.00
Humidity (%)	55.80	18.87	41.85	70.30
N = 3,650,340				

APPENDIX G: BREAKDOWN OF SUMMARY STATISTICS BY SEASON

Table G29. Summary statistics (Spring)

Variable	Mean	Standard Deviation	25^{th} Percentile	75^{th} Percentile
Home Time (hrs)	9.62	1.89	8.55	10.88
Away Time (hrs)	3.72	1.03	3.03	4.30
Distance from Home (km)	8.39	3.48	5.97	10.31
Daily Air Quality Index	29.14	14.45	18.23	38.00
Temperature (°F)	56.18	14.10	47.21	65.84
Precipitation (in.)	3.15	8.09	0.00	2.00
Wind Speed (mph)	9.40	4.06	6.26	11.63
Shortwave Radiation (w/m^2)	228.20	71.68	177.40	285.00
Humidity (%)	63.64	14.24	54.80	73.50
N = 9,492,599				

Table G30. Summary statistics (Summer)

Variable	Mean	Standard Deviation	25^{th} Percentile	75^{th} Percentile
Home Time (hrs)	9.82	2.01	8.68	11.20
Away Time (hrs)	3.91	1.13	3.14	4.58
Distance from Home (km)	9.38	3.72	6.71	11.61
Daily Air Quality Index	33.45	14.34	22.85	43.40
Temperature (°F)	75.35	7.78	69.80	80.96
Precipitation (in.)	3.29	8.95	0.00	1.40
Wind Speed (mph)	7.40	2.71	5.37	8.95
Shortwave Radiation (w/m^2)	287.40	54.08	258.50	326.80
Humidity (%)	65.41	14.46	59.80	75.05
N = 9,512,915				

Table G31. Summary statistics (Fall)

Variable	Mean	Standard Deviation	25th Percentile	75th Percentile
Home Time (hrs)	10.08	2.04	8.90	11.48
Away Time (hrs)	3.91	1.15	3.12	4.60
Distance from Home (km)	8.57	3.48	6.13	10.55
Daily Air Quality Index	31.63	15.68	20.00	41.47
Temperature (°F)	59.64	15.54	48.38	71.42
Precipitation (in.)	2.40	7.61	0.00	0.50
Wind Speed (mph)	8.88	4.10	6.04	10.96
Shortwave Radiation (w/m^2)	163.88	64.90	113.80	214.80
Humidity (%)	63.79	15.87	55.90	74.30
N = 9,370,602				

Table G32. Summary statistics (Winter)

Variable	Mean	Standard Deviation	25th Percentile	75th Percentile
Home Time (hrs)	10.15	1.98	9.05	11.46
Away Time (hrs)	3.61	1.15	2.83	4.25
Distance from Home (km)	7.85	3.47	5.46	9.68
Daily Air Quality Index	34.37	17.56	21.45	45.34
Temperature (°F)	40.79	15.21	30.83	51.08
Precipitation (in.)	3.32	8.42	0.00	2.10
Wind Speed (mph)	7.85	4.73	6.26	12.30
Shortwave Radiation (w/m^2)	103.54	44.32	70.30	133.10
Humidity (%)	66.03	15.32	57.15	76.70
N = 9,300,019				

Table H33. Results from regressions with county by day fixed effects

Variable	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)
Moderate AQ	-0.0010* (-2.56)	-0.0004 (-0.73)	-0.0014** (-3.09)
AQ Unsafe for Sensitive Groups	0.0066** (3.04)	-0.0106*** (-3.71)	-0.0229*** (-7.22)
Unhealthy AQ	0.0093** (2.84)	-0.0249*** (-4.56)	-0.0098 (-1.61)
Average Temperature	-0.0017*** (-9.95)	0.0009*** (3.91)	0.0041*** (21.28)
Average Temperature Squared	0.0001*** (6.31)	-0.0001*** (-3.83)	-0.0001*** (-17.45)
Precipitation	0.0001* (2.44)	-0.0003*** (-19.53)	-0.0005*** (-28.06)
Wind Speed	0.0005*** (6.71)	-0.0004*** (-4.88)	-0.0011*** (-13.50)
Shortwave Radiation	-0.0001*** (-6.83)	0.0001*** (16.54)	0.0001*** (21.65)
Humidity	-0.0001* (-2.50)	-0.0004*** (-7.82)	-0.0002*** (-4.82)
FS Residual	0.0004 (0.72)	-0.0014. (-1.88)	-0.0002 (-0.30)
N = 37,622,679			

Table H34. Results from regressions with county cluster fixed effects

Variable	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)

Moderate AQ	-0.0008* (-2.56)	-0.0014*** (-3.63)	-0.0018*** (-5.18)
AQ Unsafe for Sensitive Groups	0.0067*** (4.33)	-0.0182*** (-8.18)	-0.0190*** (-7.92)
Unhealthy AQ	0.0066* (2.47)	-0.0222*** (-4.86)	-0.0194*** (-3.78)
Average Temperature	-0.0009*** (-3.97)	0.0019*** (5.58)	0.0038*** (13.74)
Average Temperature Squared	0.0001. (1.84)	-0.0001*** (-5.27)	-0.0001*** (-9.94)
Precipitation	0.0001** (2.82)	-0.0002*** (-17.25)	-0.0002*** (-17.38)
Wind Speed	-0.0001 (-0.55)	0.0002* (2.19)	0.0004*** (4.87)
Shortwave Radiation	-0.0001*** (-4.42)	0.0001*** (8.29)	0.0001*** (10.33)
Humidity	-0.0001** (-3.00)	-0.0002*** (-4.37)	-0.0001* (-2.44)
FS Residual	0.0005 (1.36)	-0.0008 (-1.49)	-0.0012* (-2.57)
N = 37,622,679			

Table H35. Results from regressions without weights

Variable	**Log of Time Spent at Home (1)**	**Log of Time Spent Away from Home (2)**	**Log of Distance Traveled from Home (3)**
Moderate AQ	0.0003* (2.26)	-0.0012*** (-6.62)	-0.0028*** (-12.55)
AQ Unsafe for Sensitive Groups	0.0027** (2.81)	-0.0186*** (-12.24)	-0.0123*** (-6.21)
Unhealthy AQ	0.0139*** (5.79)	-0.0399*** (-9.66)	-0.0399*** (-7.73)

Average Temperature	-0.0017*** (0.00)	0.0007*** (20.50)	0.0042*** (108.21)
Average Temperature Squared	0.0001*** (40.75)	-0.0001*** (-23.64)	-0.0001*** (-76.71)
Precipitation	0.0002*** (39.59)	-0.0007*** (-110.47)	-0.0008*** (-112.81)
Wind Speed	0.0006*** (49.29)	-0.0004*** (-24.32)	-0.0014*** (-66.79)
Shortwave Radiation	-0.0001*** (-37.91)	0.0001*** (70.28)	0.0001*** (73.80)
Humidity	0.0001*** (12.24)	-0.0004*** (-56.67)	-0.0003*** (-36.59)
FS Residual	-0.0019*** (-14.92)	0.0026*** (15.29)	0.0043*** (20.91)
N = 37,622,679			

Table H36. Results from Regression with trimmed values of distance variable

Variable	Distance Traveled from Home Less Top and Bottom 1%	Distance Traveled from Home Less Top and Bottom 5%
Moderate AQ	-0.0031*** (-14.30)	-0.0032*** (-16.63)
AQ Unsafe for Sensitive Groups	-0.0134*** (-7.77)	-0.0108*** (-6.85)
Unhealthy AQ	-0.0407*** (-9.01)	-0.0363*** (-8.45)
Average Temperature	0.0037*** (76.90)	0.0028*** (66.91)
Average Temperature Squared	-0.0001*** (-55.07)	-0.0001*** (-46.09)

Precipitation	-0.0008*** (-105.94)	-0.0006*** (-99.27)
Wind Speed	-0.0012*** (-55.90)	-0.0010*** (-50.19)
Shortwave Radiation	0.0001*** (64.90)	0.0001*** (60.35)
Humidity	-0.0003*** (-28.80)	-0.0003*** (-34.49)
FS Residual	0.0047*** (21.24)	0.0047*** (24.15)
N	**36,870,225**	**33,860,411**

Table I37. Results from regressions on indoor and outdoor locations

Variable	Log of Visits to Outdoor Locations (CF)	Log of Visits to Indoor Locations (CF)
Moderate AQ	-0.0065* (-2.50)	-0.0075** (-2.92)
AQ Unsafe for Sensitive Groups	-0.0081 (-0.59)	-0.0103 (-0.76)
Unhealthy AQ	-0.1088*** (-3.67)	-0.1141*** (-3.86)
Average Temperature	0.0394*** (8.97)	0.0388*** (8.87)
Average Temperature Squared	-0.0003*** (-9.44)	-0.0003*** (-9.32)
Precipitation	-0.0024*** (-11.99)	-0.0024*** (-11.92)
Wind Speed	-0.0033*** (-7.38)	-0.0034*** (-7.73)
Shortwave Radiation	0.0004*** (11.47)	0.0004*** (11.57)
Humidity	-0.0004*** (-2.91)	-0.0004** (-2.74)
FS Residual	0.0193*** (7.10)	0.0194*** (7.18)
N	**4,948,437**	**5,002,444**

Table I38. Racial differences in the response to changes in air quality

Variable	Log of Time Spent at Home (1)	Log of Time Spent Away from Home (2)	Log of Distance Traveled from Home (3)

Moderate AQ	0.0017*** (5.07)	0.0021*** (5.90)	-0.0005 (-1.40)
Unsafe AQ	0.0036. (1.95)	-0.0110*** (-4.77)	-0.0103*** (-43.52)
Moderate AQ*Black CBG	-0.0020** (-3.18)	-0.0070*** (10.89)	-0.0098*** (-14.11)
Unsafe AQ* Black CBG	0.0032 (0.62)	-0.0325*** (-5.29)	-0.0514*** (-5.75)
Moderate AQ* Hispanic CBG	-0.0032*** (-3.56)	-0.0199*** (-4.94)	0.0022* (2.27)
Unsafe AQ*Hispanic CBG	-0.0047 (-1.46)	-0.0145*** (-3.46)	-0.0145** (-2.98)
Moderate AQ* Other Minority CBG	-0.0038*** (-4.18)	-0.0065*** (-8.69)	-0.0033*** (-4.47)
Unsafe AQ* Other Minority CBG	0.0051. (1.84)	-0.0112** (-3.00)	-0.0046 (1.03)
FS Residual	-0.0031*** (-13.53)	0.0020*** (7.10)	0.0041*** (15.93)
N = 37,622,679			

APPENDIX J: RESULTS FROM REGRESSION DISCONTINUITY MODELS

Table J39. Results from regressions using full sample

Variable	Home Time (1)	Away time (2)	Distance (3)
Alert at AQI of 100			
Alert	0.0089*** (6.83)	-0.0217*** (-11.68)	-0.0167*** (-8.04)
Centered AQI	0.0089*** (-10.03)	0.0000*** (4.50)	0.0001*** (6.69)
Average Temperature	0.0089*** (-30.84)	0.0005*** (20.68)	0.0043*** (73.12)
Average Temperature Squared	0.0089*** (19.34)	-0.0001*** (36.86)	-0.0001*** (-54.37)
Precipitation	0.0089*** (16.31)	-0.0006*** (53.04)	-0.0008*** (-99.16)
Wind Speed	0.0089*** (18.27)	-0.0004*** (69.22)	-0.0012*** (-47.15)
Shortwave Radiation	0.0089*** (-14.15)	0.0001*** (85.40)	0.0001*** (55.93)
Humidity	0.0089*** (9.49)	-0.0004*** (101.58)	-0.0003*** (-26.84)
Alert = AQI of 150			
Alert	0.0206*** (8.12)	-0.0423*** (-10.31)	-0.0450*** (-9.66)
Centered AQI	-0.0001*** (-9.91)	0.0000*** (3.82)	0.0001*** (6.32)
Average Temperature	-0.0017*** (-30.80)	0.0005*** (6.52)	0.0043*** (73.05)

Average Temperature Squared	0.0000*** (19.30)	-0.0000*** (-7.86)	-0.0000*** (-54.32)
Precipitation	0.0001*** (16.34)	-0.0006*** (-73.90)	-0.0008*** (-99.16)
Wind Speed	0.0004*** (18.35)	-0.0004*** (-12.64)	-0.0012*** (-47.29)
Shortwave Radiation	-0.0000*** (-14.18)	0.0001*** (41.16)	0.0001*** (55.98)
Humidity	0.0001*** (9.50)	-0.0004*** (-27.94)	-0.0003*** (-26.86)
N = 37,622,412			

Table J40. Results from regressions with restricted data (+ or - 20 from the cutoff)

Variable	**Home Time** (1)	**Away time** (2)	**Distance** (3)
Alert at AQI of 100			
Alert	0.0067*** (2.79)	-0.0062. (-1.83)	-0.0063 (-1.59)
Centered AQI	-0.0002 (-1.44)	-0.0002 (-1.41)	0.0003 (1.49)
Average Temperature	-0.0021*** (-5.11)	-0.0010. (-1.70)	0.0031*** (4.50)
Average Temperature Squared	0.0000 (1.32)	0.0000*** (3.10)	0.0000 (1.04)
Precipitation	0.0001 (0.68)	-0.0006* (-2.15)	-0.0010*** (-3.15)
Wind Speed	-0.0005 (-1.12)	0.0005 (0.77)	0.0013. (1.89)

Shortwave Radiation	-0.0000 (-0.72)	0.0003*** (3.92)	0.0005*** (6.60)
Humidity	-0.0000 (-0.07)	-0.0001 (-1.30)	0.0006*** (4.59)
N = 207,265			
Alert = AQI of 150			
Alert	0.0518 (1.44)	-0.0589 (-0.98)	0.1954** (2.70)
Centered AQI	-0.0009 (-0.55)	0.0044 (1.44)	-0.0117** (-3.01)
Average Temperature	0.0081 (0.75)	-0.0069 (-0.42)	-0.0605** (-2.79)
Average Temperature Squared	-0.0001 (-1.09)	0.0000 (0.20)	0.0005. (1.85)
Precipitation	-0.0200 (-1.79)	-0.0333 (-1.42)	-0.0016 (-0.08)
Wind Speed	0.0073 (1.04)	-0.0028 (-0.18)	0.0010 (0.07)
Shortwave Radiation	0.0042 (1.35)	-0.0049 (-1.08)	-0.0088 (-1.50)
Humidity	-0.0013 (-0.50)	0.0020 (0.48)	-0.0014 (-0.28)
N = 9,051			

Table J41. Results from regressions with restricted data (+ or - 10 from the cutoff)

Variable	**Home Time** **(1)**	**Away time** **(2)**	**Distance** **(3)**

Alert at AQI of 100			
Alert	0.0071 (1.70)	-0.0063 (-1.10)	-0.0123. (-1.80)
Centered AQI	-0.0005 (-1.30)	-0.0008 (-1.64)	0.0009 (1.54)
Average Temperature	-0.0030** (-2.65)	0.0013 (0.86)	0.0143*** (8.22)
Average Temperature Squared	0.0000 (0.48)	0.0000. (1.79)	-0.0000** (-3.11)
Precipitation	0.0007. (1.95)	0.0007 (0.99)	-0.0002 (-0.28)
Wind Speed	-0.0009 (-0.98)	0.0001 (0.04)	0.0038 (2.34)
Shortwave Radiation	-0.0002 (-1.41)	0.0004* (2.30)	0.0010 (4.22)
Humidity	0.0001 (0.53)	0.0003 (0.92)	0.0021 (6.09)
N = 70,477			
Alert = AQI of 150			
Alert	0.0303 (0.68)	0.0168 (0.22)	0.1338 (1.07)
Centered AQI	-0.0085** (-2.61)	0.0262*** (3.51)	-0.0299** (-3.18)
Average Temperature	-0.2722* (-2.24)	0.6283** (2.77)	-0.1743 (-0.43)
Average Temperature Squared	0.0022* (2.00)	-0.0054** (-2.79)	0.0012 (0.34)
Precipitation	0.1838 (1.55)	-0.3565* (-2.28)	-0.4449. (-1.98)
Wind Speed	-0.0882.	0.0453	-0.0322

	(-1.81)	(0.51)	(-0.25)
Shortwave Radiation	-0.0093	0.0135	-0.0062
	(-1.33)	(0.76)	(-0.28)
Humidity	0.0213.	-0.0356	-0.0127
	(1.95)	(-1.46)	(-0.45)
N = 4,094			